



Informatica®

10.4.0

# Guia de Transformação do Developer

© Copyright Informatica LLC 2009, 2020

Este software e a documentação são fornecidos somente sob um contrato de licença separado, contendo restrições sobre uso e divulgação. Não está permitida de forma alguma a reprodução ou a transmissão de qualquer parte deste documento (seja por meio eletrônico, fotocópia, gravação ou quaisquer outros meios) sem o consentimento prévio da Informatica LLC.

DIREITOS DO GOVERNO DOS ESTADOS UNIDOS Programas, softwares, bancos de dados, bem como a documentação e os dados técnicos relacionados, distribuídos a clientes do Governo dos EUA são "softwares de computador comerciais" ou "dados técnicos comerciais", de acordo com o Regulamento de Aquisição Federal aplicável e os regulamentos suplementares específicos da agência. Como tal, a utilização, duplicação, divulgação, modificação e adaptação estão sujeitas às restrições e aos termos de licença estabelecidos no contrato governamental aplicável e, na medida do que for aplicável pelos termos do contrato governamental, aos direitos adicionais estabelecidos no FAR 52.227-19, Licença de Software de Computador Comercial.

Informatica, o logotipo Informatica e PowerCenter são marcas comerciais ou marcas registradas da Informatica LLC nos Estados Unidos e em muitas jurisdições por todo o mundo. Uma lista atual das marcas comerciais da Informatica está disponível na Internet em <https://www.informatica.com/trademarks.html>. Os nomes de outras companhias e produtos podem ser nomes ou marcas comerciais de seus respectivos proprietários.

Partes deste software e/ou documentação estão sujeitas a copyright detido por terceiros. Os avisos de terceiros necessários são incluídos no produto.

As informações contidas neste documento estão sujeitas a alteração sem aviso prévio. Se você encontrar quaisquer problemas nesta documentação, informe-os em [infa\\_documentation@informatica.com](mailto:infa_documentation@informatica.com).

Os produtos Informatica apresentam garantias segundo os termos e condições dos acordos em que são fornecidos. A INFORMATICA FORNECE AS INFORMAÇÕES NESTE DOCUMENTO "COMO ESTÃO" SEM GARANTIA DE QUALQUER TIPO, EXPRESSA OU IMPLÍCITA, INCLUINDO, SEM QUAISQUER GARANTIAS DE COMERCIALIZAÇÃO, ADEQUAÇÃO A UM DETERMINADO FIM E QUALQUER GARANTIA OU CONDIÇÃO DE NÃO-VIOLAÇÃO.

Data da Publicação: 2020-02-10

# Conteúdo

<b>Prefácio.....</b>	<b>30</b>
Recursos da Informatica. . . . .	30
Rede da Informatica. . . . .	30
Base de Dados de Conhecimento da Informatica. . . . .	30
Documentação da Informatica. . . . .	31
Matrizes de Disponibilidade de Produto da Informatica. . . . .	31
Informatica Velocity. . . . .	31
Informatica Marketplace. . . . .	31
Suporte Global a Clientes da Informatica. . . . .	31
 <b>Capítulo 1: Introdução às Transformações.....</b>	<b>32</b>
Introdução às Transformações - Visão Geral. . . . .	32
Transformações Ativas. . . . .	32
Transformações passivas. . . . .	33
Transformações Desconectadas. . . . .	33
Transformações de Várias Estratégias. . . . .	33
Descrições das Transformações. . . . .	34
Transformações nos ambientes nativo e não nativo. . . . .	36
Lidando com Tipos de Dados de Transformação. . . . .	39
Tipo de Dados Decimal. . . . .	39
Registro de Data/Hora com Fuso Horário. . . . .	40
Registro de Data/Hora com Fuso Horário Local. . . . .	41
Desenvolvendo uma Transformação. . . . .	41
Transformações de vários grupos. . . . .	41
Regras e Diretrizes para Transformações de Vários Grupos. . . . .	42
Expressões nas Transformações. . . . .	42
O Editor de Expressão. . . . .	44
Nomes de Porta em uma Expressão. . . . .	44
Adicionando uma Expressão a uma Porta. . . . .	44
Comentários em uma Expressão. . . . .	44
Validação de Expressão. . . . .	44
Testar Expressões. . . . .	45
Conversão de Tipo de Dados. . . . .	45
Variáveis Locais. . . . .	46
Temporariamente armazenar dados e simplificar expressões complexas. . . . .	46
Armazenar Valores nas Linhas. . . . .	47
Capturar valores de procedimentos armazenados. . . . .	48
Diretrizes para configurar portas variáveis. . . . .	48
Inicialização de Variáveis. . . . .	49
Valores Padrão para Portas. . . . .	49

Valores Padrão Definidos pelo Usuário. . . . .	50
Valores de Entrada Padrão Definidos pelo Usuário. . . . .	52
Validação do Valor Padrão. . . . .	53
Valores de Saída Padrão Definidos pelo Usuário . . . . .	53
Regras gerais para valores padrão. . . . .	55
Validação do Valor Padrão. . . . .	56
Níveis de rastreamento. . . . .	56
Transformações Reutilizáveis. . . . .	57
Instâncias de Transformação Reutilizável e Alterações Herdadas. . . . .	57
Editando uma Transformação Reutilizável. . . . .	57
Exibições do Editor para uma Transformação Reutilizável. . . . .	57
Transformações Não Reutilizáveis. . . . .	58
Exibições do Editor para uma Transformação Não Reutilizável. . . . .	58
Criando uma Transformação. . . . .	58
<b>Capítulo 2: Portas de Transformação. . . . .</b>	<b>59</b>
Visão geral de portas de transformação. . . . .	59
Criar Portas. . . . .	59
Configurar Portas. . . . .	60
Vinculando Portas. . . . .	60
Links de Um para Um. . . . .	61
Links de Um para Muitos. . . . .	61
Vinculando Portas Manualmente. . . . .	61
Vinculando Portas Automaticamente. . . . .	61
Regras e Diretrizes para Vincular Portas. . . . .	62
Propagando Atributos de Porta. . . . .	63
Tipos de Dependência. . . . .	63
Dependências do Caminho de Link. . . . .	63
Dependências Implícitas. . . . .	64
Atributos de Porta Propagados por Transformação. . . . .	64
Copiando portas do Excel. . . . .	66
Editando transformações no Excel. . . . .	67
Copiando metadados para a Developer tool. . . . .	67
Exemplo: Editando uma transformação no Excel. . . . .	68
Regras e diretrizes para copiar do Excel. . . . .	69
<b>Capítulo 3: Caches de Transformação. . . . .</b>	<b>70</b>
Visão Geral dos Caches de Transformação. . . . .	70
Tipos de Cache. . . . .	71
Arquivos de Cache. . . . .	71
Diretório de Arquivo de Cache. . . . .	72
Tamanho do Cache. . . . .	72
Tamanho do Cache Automático. . . . .	72



Tamanho do Cache Específico. . . . .	74
Aumento do Tamanho do Cache pelo Serviço de Integração de Dados. . . . .	74
Tamanho do Cache para Caches Particionados. . . . .	75
Otimização do Tamanho do Cache. . . . .	75
Etapa 1. Defina o Nível de Rastreamento para Inicialização Detalhada. . . . .	75
Etapa 2. Executar o Mapeamento no Modo de Cache Automático. . . . .	76
Etapa 3. Analisar o Desempenho do Cache. . . . .	76
Etapa 4. Configurar os Tamanhos de Cache Específicos. . . . .	76
<b>Capítulo 4: Transformação de Validador de Endereço.....</b>	<b>78</b>
Transformação de Validador de Endereço - Visão Geral. . . . .	79
Dados de Referência de Endereço. . . . .	79
Tipos de Dados de Referência de Endereço. . . . .	80
Modos e modelos. . . . .	82
Grupos de Portas e Seleção de Porta. . . . .	82
Transformação de Validador de Endereço - Grupos de Portas de Entrada. . . . .	82
Transformação de Validador de Endereço - Grupos de Portas de Saída. . . . .	83
Portas de Várias Instâncias. . . . .	86
Projetos de Validação de Endereço. . . . .	87
Endereços Formatados e Padrões do Serviço de Correio. . . . .	88
Conclusão de Endereço Parcial . . . . .	89
Portas de Status do Validador de Endereço. . . . .	90
Definições de Código de Status do Elemento. . . . .	91
Valores da Porta de Saída do Código de Resolução de Endereço. . . . .	93
Valores de Porta de Saída do Status da Entrada do Elemento. . . . .	94
Valores de Porta de Saída da Relevância do Elemento. . . . .	94
Valores de Porta de Saída do Status do Resultado do Elemento. . . . .	95
Valores de Porta de Saída do Status do Resultado do Elemento Estendido. . . . .	97
Valores de Porta de Saída da Pontuação de Capacidade de Envio por Correio. . . . .	98
Valores de Porta de Saída do Código de Correspondência. . . . .	99
Valores de Porta de Saída do Status da Codificação Geográfica. . . . .	101
Transformação de Validador de Endereço - Configurações Gerais. . . . .	102
Propriedades de validação de endereço na janela Preferências. . . . .	103
Propriedades de dados de validação de endereço. . . . .	104
Propriedades de licença de validação de endereço. . . . .	105
Propriedades do mecanismo de validação de endereço. . . . .	105
Propriedades avançadas de validação de endereço. . . . .	106
Alias de Localidade. . . . .	107
Rua de Alias. . . . .	107
Estilo de Formatação de Maiúsculas e Minúsculas. . . . .	107
País de Origem. . . . .	108
Tipo de País. . . . .	108
País Padrão. . . . .	109

Prioridade de Endereço Duplo. . . . .	110
Abreviação de Elemento. . . . .	110
Instâncias de Execução. . . . .	110
Expansão de Intervalo Flexível. . . . .	111
Tipo de Dados de Codificação Geográfica. . . . .	112
Comprimento Máximo de Campo Global. . . . .	112
Descritor Preferencial Global. . . . .	113
Tipo de Formato de Entrada. . . . .	113
Formato de Entrada com País . . . . .	114
Separador de Linha. . . . .	114
Alternativas de Correspondência. . . . .	114
Arquivo Morto Estendido Correspondente. . . . .	115
Escopo de Correspondência. . . . .	116
Contagem Máxima de Resultados. . . . .	116
Modo. . . . .	116
Nível de Otimização. . . . .	117
Tipo de Formato de Saída. . . . .	117
Formato de Saída com País. . . . .	118
Idioma Preferencial. . . . .	118
Script Preferencial. . . . .	124
Intervalos para Expandir. . . . .	125
Padronizar Endereços Inválidos. . . . .	126
Nível de Rastreamento. . . . .	127
Relatórios de Certificação. . . . .	127
Campos de Relatório AMAS. . . . .	127
Campos de Relatório CASS. . . . .	128
Relatório SendRight. . . . .	129
Campos de Relatório SERP. . . . .	129
Configurando uma Transformação de Validador de Endereço. . . . .	130
Adicionando Portas à Transformação de Validador de Endereço. . . . .	130
Criando Modelos Definidos pelo Usuário. . . . .	131
Definindo Modelos do Validador de Endereço. . . . .	131
Definindo um Relatório de Certificação. . . . .	132
Transformação de Validador de Endereço em um ambiente não nativo. . . . .	132
Transformação de Validador de Endereço no Mecanismo Blaze. . . . .	133
Transformação de Validador de Endereço no Mecanismo Spark. . . . .	133
<b>Capítulo 5: Transformação de Agregador . . . . .</b>	<b>134</b>
Transformação de Agregador - Visão Geral. . . . .	134
Transformações de Agregador em Mapeamentos Dinâmicos. . . . .	135
Desenvolvendo uma Transformação de Agregador. . . . .	135
Portas da Transformação de Agregador. . . . .	135
Expressões Agregadas. . . . .	136

Funções de Agregação. . . . .	137
Funções de Agregação Aninhadas. . . . .	138
Cláusulas Condicionais em Expressões Agregadas. . . . .	138
Agrupar por Portas. . . . .	138
Configurar Portas de Agrupamento. . . . .	139
Parâmetros de Agrupamento. . . . .	140
Valores Padrão de Grupo por Portas . . . . .	140
Expressões Não Agregadas. . . . .	141
Caches de Agregador. . . . .	141
Entrada Classificada para uma Transformação de Agregador. . . . .	141
Condições de Entrada Classificada. . . . .	142
Classificando Dados em uma Transformação de Agregador. . . . .	142
Propriedades Avançadas da Transformação de Agregador. . . . .	143
Criando uma Transformação de Agregador Reutilizável. . . . .	144
Criando uma Transformação de Agregador Não Reutilizável. . . . .	144
Dicas para Transformações de Agregador. . . . .	145
Solucionando Problemas de Transformações de Agregador. . . . .	145
Transformação de agregador em um ambiente não nativo. . . . .	146
Transformação de agregador no mecanismo Blaze. . . . .	146
Transformação de agregador no mecanismo Spark. . . . .	147
Transformação de agregador no mecanismo Databricks Spark. . . . .	147
<b>Capítulo 6: Transformação de Associação. . . . .</b>	<b>149</b>
Transformação de Associação - Visão Geral. . . . .	149
Alocação de Memória. . . . .	150
Propriedades Avançadas da Transformação de Associação. . . . .	151
<b>Capítulo 7: Transformação de Exceção de Registro. . . . .</b>	<b>152</b>
Visão Geral da Transformação de Exceção de Registro Inválido. . . . .	152
Tipos de Registro de Saída de Exceção de Registro Inválido. . . . .	153
Fluxo do Processo de Gerenciamento de Exceção de Registro Inválido. . . . .	154
Mapeamentos de Exceção de Registro Inválido. . . . .	154
Problemas de Qualidade de Exceção de Registro Inválido. . . . .	155
Tarefas Humanas. . . . .	156
Portas de Exceção de Registro Inválido . . . . .	157
Portas de Entrada de Transformação de Exceção de Registro Inválido. . . . .	157
Saída de Transformação de Exceção de Registro Inválido. . . . .	158
Exibição Configuração de Exceção de Registro Inválido. . . . .	158
Gerando a Tabela de Registros Inválidos e a Tabela de Problemas. . . . .	160
Atribuição de Problema de Exceção de Registro Inválido . . . . .	160
Atribuindo Portas para Problemas de Qualidade. . . . .	161
Propriedades Avançadas da Transformação de Exceção. . . . .	161
Configurando uma Transformação de Exceção de Registros Inválidos. . . . .	162

Exemplo de Mapeamento de Exceção de Registro Inválido. . . . .	163
Mapplet de Exceção de Registro Inválido. . . . .	163
Grupos de Entrada de Exemplo de Exceção de Registro Inválido. . . . .	164
Configuração de Exemplo de Exceção de Registro Inválido . . . . .	164
Saída de Mapeamento de Exemplo de Exceção de Registro Inválido. . . . .	165

## **Capítulo 8: Transformação de Conversor de Maiúsculas/Minúsculas..... 168**

Transformação de Conversor de Maiúsculas/Minúsculas - Visão Geral. . . . .	168
Propriedades da Estratégia de Maiúsculas/Minúsculas. . . . .	169
Configurando uma Estratégia de Conversor de Maiúsculas/Minúsculas. . . . .	169
Propriedades Avançadas da Transformação de Conversor de Maiúsculas/Minúsculas. . . . .	170
Transformação de Conversor de maiúsculas/minúsculas em um ambiente não nativo. . . . .	170

## **Capítulo 9: Transformação de Classificador..... 171**

Transformação de Classificador - Visão Geral. . . . .	171
Modelos do Classificador. . . . .	172
Algoritmos Classificadores. . . . .	172
Opções de Transformação de Classificador. . . . .	172
Estratégias de Classificador. . . . .	173
Propriedades Avançadas da Transformação de Classificador. . . . .	173
Configurando uma Estratégia de Classificador. . . . .	174
Exemplo de Análise de Classificador. . . . .	174
Criar o Mapeamento Classificador. . . . .	175
Amostra de Dados de Entrada. . . . .	176
Configuração da Fonte de Dados. . . . .	176
Configuração da Transformação de Classificador. . . . .	176
Configuração da Transformação de Roteador. . . . .	177
Configuração de Destino de Dados. . . . .	178
Resultado do Mapeamento Classificador. . . . .	178
Transformação de Classificador em um ambiente não nativo. . . . .	179

## **Capítulo 10: Transformação de Comparação..... 180**

Transformação de Comparação - Visão Geral. . . . .	180
Estratégias de Correspondência de Campo. . . . .	180
Bigram. . . . .	181
Distância de Hamming. . . . .	181
Distância de Edição. . . . .	181
Distância de Jaro. . . . .	182
Distância de Hamming Invertida. . . . .	182
Estratégias de Correspondência de Identidade. . . . .	183
Configurando uma Estratégia de Comparação. . . . .	184
Propriedades Avançadas da Transformação de Comparação. . . . .	184
Transformação de Comparação em um ambiente não nativo. . . . .	185

<b>Capítulo 11: Transformação de Consolidação.....</b>	<b>186</b>
Transformação de Consolidação - Visão Geral. . . . .	186
Mapeamentos de Consolidação. . . . .	187
Portas da Transformação de Consolidação. . . . .	187
Exibições da Transformação de Consolidação. . . . .	187
Exibição Estratégias da Transformação de Consolidação. . . . .	188
Propriedades Avançadas da Transformação de Consolidação. . . . .	188
Tamanho do Arquivo de Cache. . . . .	189
Estratégias Simples. . . . .	190
Estratégias Baseadas em Linhas. . . . .	191
Estratégias Avançadas. . . . .	192
Funções de Consolidação Simples. . . . .	192
CONSOL_AVG. . . . .	192
CONSOL_LONGEST. . . . .	193
CONSOL_MAX. . . . .	193
CONSOL_MIN. . . . .	194
CONSOL_MOSTFREQ. . . . .	195
CONSOL_MOSTFREQ_NB. . . . .	195
CONSOL_SHORTEST. . . . .	196
Funções de Consolidação Baseadas em Linhas. . . . .	196
CONSOL_GETROWFIELD. . . . .	197
CONSOL_MODALEXACT. . . . .	197
CONSOL_MOSTDATA. . . . .	198
CONSOL_MOSTFILLED. . . . .	199
Exemplo de Mapeamento de Consolidação. . . . .	200
Dados de Entrada. . . . .	200
Transformação de Gerador de Chaves. . . . .	200
Transformação de Consolidação. . . . .	201
Saída do Mapeamento de Consolidação. . . . .	201
Configurando uma Transformação de Consolidação. . . . .	201
Transformação de Consolidação em um ambiente não nativo. . . . .	202
Transformação de Consolidação no mecanismo Blaze. . . . .	202
Transformação de Consolidação no mecanismo Spark. . . . .	202
Transformação de Consolidação no mecanismo Databricks Spark. . . . .	202
 <b>Capítulo 12: Transformação de Mascaramento de Dados.....</b>	 <b>203</b>
Transformação de Mascaramento de Dados - Visão Geral. . . . .	203
Técnicas de Mascaramento. . . . .	204
Mascaramento Aleatório. . . . .	204
Mascaramento de Expressão. . . . .	206
Mascaramento de Chave. . . . .	208
Mascaramento de Substituição. . . . .	209

Mascaramento dependente. . . . .	213
Mascaramento de Tokenização. . . . .	214
Regras de Mascaramento. . . . .	215
Formato de Máscara. . . . .	216
Caracteres da String de Origem. . . . .	217
Caracteres de Substituição da String de Resultados. . . . .	218
Intervalo. . . . .	218
Embaçamento. . . . .	219
Formatos Especiais de Máscara. . . . .	220
Mascaramento de Número de Cartão de Crédito. . . . .	220
Mascaramento de Endereço de E-mail. . . . .	220
Mascaramento de E-mail Avançado. . . . .	221
Mascaramento de Endereço IP. . . . .	222
Mascaramento de Número de Telefone. . . . .	222
Mascaramento de CPF. . . . .	222
Mascaramento de Endereço de URL. . . . .	223
Mascaramento de Número de Previdência Social. . . . .	223
Arquivo de Valor Padrão. . . . .	224
Configuração da Transformação de Mascaramento de Dados. . . . .	224
Configurar o Data Integration Service. . . . .	225
Criando uma Transformação de Mascaramento de Dados. . . . .	225
Definindo as Portas. . . . .	225
Configurando o Mascaramento de Dados para Cada Porta. . . . .	226
Visualizando os Dados Mascarados. . . . .	226
Propriedades de Tempo de Execução da Transformação de Mascaramento de Dados. . . . .	226
Exemplo de Mascaramento de Dados. . . . .	227
Dados Read_Customer. . . . .	228
Transformação de Mascaramento de Dados de Clientes. . . . .	228
Resultados de Dados de Teste de Clientes. . . . .	229
Propriedades Avançadas da Transformação de Mascaramento de Dados. . . . .	230
Transformação de Mascaramento de dados em um ambiente não nativo. . . . .	230
Transformação de Mascaramento de dados no mecanismo Blaze. . . . .	230
Transformação de Mascaramento de dados no mecanismo Spark. . . . .	231
<b>Capítulo 13: Transformação do Processador de Dados. . . . .</b>	<b>233</b>
Visão geral da Transformação do Processador de Dados. . . . .	233
Exibições da Transformação do Processador de Dados. . . . .	234
Portas da Transformação do Processador de Dados. . . . .	235
Portas de Entrada da Transformação do Processador de Dados. . . . .	235
Portas de Saída de Transformação do Processador de Dados. . . . .	236
Portas de Passagem. . . . .	237
Componente de Inicialização. . . . .	237
Referências. . . . .	238

Configurações da Transformação do Processador de Dados. . . . .	239
Codificação de Caracteres. . . . .	239
Regras e Diretrizes da Codificação de Caracteres. . . . .	241
Configurações de Saída. . . . .	241
Configurações de Processamento. . . . .	243
Configurações de XMap. . . . .	244
Configuração de Saída XML. . . . .	244
Eventos. . . . .	246
Tipos de Eventos. . . . .	246
Exibição de Eventos do Processador de Dados. . . . .	247
Logs. . . . .	247
Log de Eventos de Tempo de Design. . . . .	248
Log de Eventos de Tempo de Execução. . . . .	248
Exibindo um Log de Eventos na Exibição de Eventos do Processador de Dados. . . . .	249
Log do Usuário. . . . .	249
Desenvolvimento da Transformação do Processador de Dados. . . . .	250
Criar a Transformação do Processador de Dados. . . . .	250
Selecionar os Objetos de Esquema . . . . .	251
Criar Objetos em uma Transformação do Processador de Dados em Branco. . . . .	251
Criar as Portas. . . . .	253
Testando a Transformação. . . . .	254
Importação e Exportação da Transformação do Processador de Dados. . . . .	254
Exportando a Transformação do Processador de Dados como um Serviço. . . . .	254
Importando vários serviços do Data Transformation. . . . .	255
Importando um Serviço de Data Transformation . . . . .	255
Exportando um Mapeamento com uma Transformação do Processador de Dados para o PowerCenter. . . . .	256
Transformação de Processador de Dados em um ambiente não nativo. . . . .	257
<b>Capítulo 14: Transformação de Decisão.....</b>	<b>258</b>
Transformação de Decisão - Visão Geral. . . . .	258
Transformação de Decisão - Funções. . . . .	259
<b>Transformação de Decisão - Instruções Condicionais</b> . . . . .	261
Transformação de Decisão - Operadores. . . . .	262
Tratamento NULL da Transformação de Decisão. . . . .	263
Configurando uma Estratégia de Decisão . . . . .	263
Propriedades Avançadas da Transformação de Decisão. . . . .	264
Transformação de Decisão em um ambiente não nativo. . . . .	264
<b>Capítulo 15: Transformação de Exceção de Registro.....</b>	<b>265</b>
Visão Geral da Transformação de Exceção de Registro Duplicado. . . . .	265
Fluxo do Processo de Exceção de Registro Duplicado. . . . .	266
Exceções de Registro Duplicado. . . . .	266

Exibição de Configuração de Exceção de Registro Duplicado . . . . .	267
Gerando uma Tabela de Registros Duplicados. . . . .	268
Portas. . . . .	269
Portas de Entrada da Transformação de Exceção de Registro Duplicado. . . . .	269
Portas de Saída da Transformação de Exceção de Registro Duplicado. . . . .	270
Criando portas . . . . .	271
Propriedades Avançadas da Transformação de Exceção de Registro Duplicado. . . . .	271
Exemplo de Mapeamento de Exceção de Registro Duplicado. . . . .	272
Mapeamento de Exceção de Registro Duplicado. . . . .	272
Transformação de Correspondência . . . . .	273
Grupos de Entrada de Exceção de Registro Duplicado. . . . .	273
Exemplo de Exibição de Configuração de Exceção de Registro Duplicado. . . . .	274
Tabela de Registros de Saída Padrão. . . . .	275
Saída de Cluster. . . . .	276
Criando uma Transformação de Exceção de Registro Duplicado. . . . .	278
<b>Capítulo 16: Transformação de Expressão. . . . .</b>	<b>279</b>
Visão Geral da Transformação de Expressão. . . . .	279
Portas de Transformação de Expressão. . . . .	280
Testar Expressões. . . . .	281
Cadeias de Formato de Data para Dados de Amostra. . . . .	282
Testando uma Expressão. . . . .	282
Seletores de Portas. . . . .	283
Configuração do seletor de portas. . . . .	283
Regras de Seleção. . . . .	284
Criando um Seletor de Portas. . . . .	285
Definição de Janelas. . . . .	286
Configuração da definição de janelas. . . . .	286
Expressões Dinâmicas. . . . .	290
Configurações da Porta de Saída. . . . .	291
Criando uma Expressão Dinâmica . . . . .	292
Transformação de Expressão - Propriedades Avançadas. . . . .	295
Transformação de Expressão em um ambiente não nativo. . . . .	295
Transformação de Expressão no mecanismo Blaze. . . . .	295
Transformação de Expressão no mecanismo Spark. . . . .	296
Transformação de Expressão no mecanismo Databricks Spark. . . . .	296
<b>Capítulo 17: Transformação de Filtro . . . . .</b>	<b>297</b>
Transformação de Filtro - Visão Geral. . . . .	297
Transformações de Filtro em Mapeamentos Dinâmicos. . . . .	298
Condição de Filtro. . . . .	299
Parametrizar a Condição de Filtro. . . . .	299
Filtrando Linhas com Valores Nulos. . . . .	301



Transformação de Filtro - Propriedades Avançadas. . . . .	301
Transformação de Filtro - Dicas de Desempenho. . . . .	301
Transformação de Filtro em um ambiente não nativo. . . . .	301
Transformação de Filtro no mecanismo Blaze. . . . .	302
<b>Capítulo 18: Transformação Hierárquica para Relacional. . . . .</b>	<b>303</b>
Visão geral da transformação de Hierárquica para Relacional. . . . .	303
Exemplo - transformação de Hierárquica para Relacional. . . . .	304
Portas relacionais de saída e a exibição Visão Geral. . . . .	305
Portas da Transformação Hierárquica para Relacional. . . . .	306
Referências de esquema. . . . .	307
Configuração de portas. . . . .	307
Desenvolvimento da Transformação Hierárquica para Relacional. . . . .	308
Criando a transformação de Hierárquica para Relacional . . . . .	308
Configurando as Portas e o Mapeamento. . . . .	308
Testando a Transformação. . . . .	309
<b>Capítulo 19: Transformação Java. . . . .</b>	<b>310</b>
Visão Geral de Transformações Java. . . . .	310
Transformações Java Reutilizáveis e Não Reutilizáveis. . . . .	311
Transformações Java Ativas e Passivas. . . . .	311
Conversão de Tipo de Dados. . . . .	311
Conversão de tipo de dados complex no mecanismo Spark. . . . .	312
Projetando uma Transformação Java. . . . .	314
Portas da Transformação Java. . . . .	314
Criando portas. . . . .	315
Definindo Valores Padrão para Portas. . . . .	315
Transformação Java - Propriedades Avançadas. . . . .	315
Configurando o Classpath para a ferramenta Developer do . . . . .	317
Configurando o Classpath para o Serviço de Integração de Dados. . . . .	318
Desenvolvendo Código Java. . . . .	318
Criando trechos de código Java. . . . .	320
Importando pacotes Java. . . . .	321
Definindo Passiva. . . . .	321
Propriedades Java da Transformação Java. . . . .	322
Guia Importações. . . . .	323
Guia Auxiliares. . . . .	323
Guia Na Entrada. . . . .	323
Guia No Final. . . . .	324
Guia Funções. . . . .	324
Guia Código Completo. . . . .	325
Otimização de Filtro com a Transformação Java. . . . .	325
Otimização de Seleção Antecipada com a Transformação Java. . . . .	325

Otimização de Envio com a Transformação Java. . . . .	327
Criando uma Transformação Java. . . . .	328
Criando uma Transformação Java Reutilizável. . . . .	328
Criando uma Transformação Java Não Reutilizável. . . . .	328
Compilando uma transformação Java. . . . .	329
Solucionando Problemas de uma Transformação Java. . . . .	329
Localizando a Origem de Erros de Compilação. . . . .	330
Identificando a origem de erros de compilação. . . . .	330
Convertendo para exemplo de dados struct. . . . .	331
Transformação Java em um ambiente não nativo. . . . .	334
Transformação Java no mecanismo Blaze. . . . .	334
Transformação Java no mecanismo Spark. . . . .	335
 <b>Capítulo 20: Referência da API da transformação Java. . . . .</b>	 <b>337</b>
Visão geral dos métodos API da transformação Java. . . . .	337
defineJExpression. . . . .	338
failSession. . . . .	339
generateRow. . . . .	339
getInRowType. . . . .	340
getMetadata. . . . .	341
incrementErrorCount. . . . .	341
invokeJExpression. . . . .	342
isNull. . . . .	343
logError. . . . .	344
logInfo. . . . .	344
resetNotification. . . . .	345
setNull. . . . .	345
storeMetadata. . . . .	346
 <b>Capítulo 21: Expressões Java. . . . .</b>	 <b>348</b>
Visão Geral das Expressões Java. . . . .	348
Tipos de Função de Expressão. . . . .	349
Usando a caixa de diálogo Definir Função para definir uma expressão. . . . .	349
Etapa 1. Configurar a Função. . . . .	350
Etapa 2. Criar e Validar a Expressão. . . . .	350
Etapa 3. Gerar o Código Java para a Expressão. . . . .	350
Criando uma expressão e gerando código Java usando a caixa de diálogo Definir Função.. . . .	350
Modelos da expressão Java. . . . .	351
Trabalhando com a Interface Simples. . . . .	351
invokeJExpression. . . . .	351
Exemplo de Interface Simples. . . . .	352
Trabalhando com a Interface Avançada. . . . .	353
Invocando uma expressão com a Interface Avançada. . . . .	353

Regras e Diretrizes para Trabalhar com a Interface Avançada. . . . .	353
Classe EDataType. . . . .	354
Classe JExprParamMetadata. . . . .	354
defineJExpression. . . . .	355
Classe JExpression. . . . .	356
Exemplo de Interface Avançada. . . . .	356
Referência de API da classe JExpression. . . . .	357
getBytes. . . . .	357
getDouble. . . . .	358
getInt. . . . .	358
getLong. . . . .	358
getResultDataType. . . . .	358
getResultMetadata. . . . .	358
getStringBuffer. . . . .	359
invoke. . . . .	359
isResultNull. . . . .	359
<b>Capítulo 22: Transformação de Associador . . . . .</b>	<b>361</b>
Transformação de Associador - Visão Geral. . . . .	361
Propriedades Avançadas da Transformação de Associador. . . . .	362
Caches de Associador. . . . .	363
Portas de Transformação de Associador. . . . .	364
Transformações de Associador em Mapeamentos Dinâmicos. . . . .	365
Seletores de Portas em uma Transformação de Associador. . . . .	365
Regras de Seleção . . . . .	366
Criando um Seletor de Portas. . . . .	367
Definindo uma Condição de Associação. . . . .	368
Tipo de Condição Simples. . . . .	369
Tipo de Condição Avançada. . . . .	369
Seletores de Portas em Condições de Associação. . . . .	370
Portas Dinâmicas em Condições de Associação. . . . .	371
Parâmetro de Expressão. . . . .	371
Tipos de Associação. . . . .	372
Associação Normal. . . . .	372
Associação Externa Mestra. . . . .	373
Associação Externa Detalhada. . . . .	374
Associação Externa Completa. . . . .	374
Entrada Classificada para uma Transformação de Associador. . . . .	375
Configurando a Ordem de Classificação. . . . .	375
Adicionando Transformações ao Mapeamento. . . . .	375
Regras e Diretrizes para Condições de Associação. . . . .	376
Exemplo de uma Condição de Associação e Ordem de Classificação. . . . .	377
Associando Dados da Mesma Origem. . . . .	379

Associando Duas Ramificações do Mesmo Pipeline. . . . .	379
Associando Duas Instâncias da Mesma Origem. . . . .	380
Diretrizes para Associar Dados da Mesma Origem. . . . .	381
Bloqueando os Pipelines de Origem. . . . .	381
Transformação de Associador Não Classificada. . . . .	382
Transformação de Associador Classificada. . . . .	382
Transformação de Associador - Dicas de Desempenho. . . . .	382
Regras e Diretrizes para uma Transformação de Associador. . . . .	383
Transformação de Associador em um ambiente não nativo . . . . .	384
Transformação de Associador no mecanismo Blaze. . . . .	384
Transformação de Associador no mecanismo Spark. . . . .	384
Transformação de Associador no mecanismo Databricks Spark. . . . .	385
<b>Capítulo 23: Transformação de Gerador de Chaves.....</b>	<b>386</b>
Transformação de Gerador de Chaves - Visão Geral. . . . .	386
Estratégia de Soundex. . . . .	387
Propriedades da Estratégia de Soundex. . . . .	387
Estratégia de String. . . . .	388
Propriedades da Estratégia de String. . . . .	388
Estratégia de NYSIIS. . . . .	388
Portas de Saída do Gerador de Chaves. . . . .	389
Configurando uma Estratégia de Agrupamento. . . . .	389
Propriedades de Criação de Chave. . . . .	390
Propriedades Avançadas da Transformação do Gerador de Chaves. . . . .	390
Transformação de Gerador de Chaves em um ambiente não nativo. . . . .	391
<b>Capítulo 24: Transformação de Rotulador.....</b>	<b>392</b>
Transformação de Rotulador - Visão Geral. . . . .	392
Quando Usar uma Transformação de Rotulador. . . . .	393
Uso de Dados de Referência na Transformação de Rotulador. . . . .	394
Conjuntos de Caracteres. . . . .	395
Modelos Probabilísticos. . . . .	395
Tabelas de Referência. . . . .	395
Expressões Regulares. . . . .	395
Conjuntos de Tokens. . . . .	396
Estratégias de Transformação de Rotulador. . . . .	396
Operações de Rotulagem de Caractere. . . . .	396
Operações de Rotulagem de Token. . . . .	397
Portas da Transformação de Rotulador. . . . .	397
Propriedades de Rotulagem de Caractere. . . . .	397
Propriedades Gerais. . . . .	398
Propriedades da Tabela de Referência. . . . .	398
Propriedades do Conjunto de Caracteres. . . . .	399

Propriedades de Filtro. . . . .	399
Propriedades de Rotulagem de Token. . . . .	400
Propriedades Gerais. . . . .	400
Propriedades do Conjunto de Tokens. . . . .	401
Propriedades de Rótulo Personalizadas. . . . .	401
Propriedades da Correspondência Probabilística. . . . .	402
Propriedades da Tabela de Referência. . . . .	402
Configurando uma Estratégia de Rotulagem de Caractere. . . . .	403
Configurando uma Estratégia de Rotulagem de Token. . . . .	403
Propriedades Avançadas da Transformação de Rotulador. . . . .	404
Transformação de Rotulador em um ambiente não nativo. . . . .	404
 <b>Capítulo 25: Transformação de Pesquisa. . . . .</b>	<b>405</b>
Transformação de Pesquisa - Visão Geral. . . . .	405
Pesquisas Conectadas e Não Conectadas. . . . .	406
Pesquisas Conectadas. . . . .	407
Pesquisas Não Conectadas. . . . .	408
Desenvolvendo uma Transformação de Pesquisa. . . . .	409
Consulta de Pesquisa. . . . .	409
Consulta de Pesquisa Padrão. . . . .	409
Substituição SQL para uma Consulta de Pesquisa. . . . .	409
Parâmetros em uma Consulta de Substituição SQL. . . . .	410
Palavras reservadas. . . . .	411
Orientações para Substituir a Consulta de Pesquisa. . . . .	411
Substituindo a Consulta de Pesquisa. . . . .	411
Filtro da Origem de Pesquisa. . . . .	412
Filtrando Linhas de Origem em uma Pesquisa. . . . .	412
Condição de Pesquisa. . . . .	412
Configurar a Condição de Pesquisa. . . . .	414
Regras e Diretrizes para Condições de Transformação de Pesquisa. . . . .	414
Cache de Pesquisa. . . . .	415
Propriedades de Consulta. . . . .	416
Transformações de Pesquisa em Mapeamentos Dinâmicos. . . . .	416
Definir Portas Dinâmicas. . . . .	417
Alterar a Origem da Pesquisa. . . . .	417
Parametrizar a Origem de Pesquisa. . . . .	418
Origens de Pesquisa que Contêm Parâmetros . . . . .	420
Configurar Parâmetros em um Objeto de Dados Duplicado. . . . .	420
Seletores de Portas. . . . .	422
Configuração do seletor de portas. . . . .	423
Regras de Seleção. . . . .	424
Parametrizar a Condição de Pesquisa. . . . .	425
Criando um Seletor de Portas. . . . .	426

Propriedades de Tempo de Execução. . . . .	428
Propriedades avançadas. . . . .	429
Criando uma Transformação de Pesquisa Reutilizável. . . . .	430
Criando uma Transformação de Pesquisa Não Reutilizável. . . . .	431
Criando uma Transformação de Pesquisa Não Conectada. . . . .	432
Exemplo de Pesquisa Não Conectada. . . . .	433
Transformação de Pesquisa em um ambiente não nativo. . . . .	435
Transformação de Pesquisa no mecanismo Blaze. . . . .	435
Transformação de Pesquisa no mecanismo Spark. . . . .	435
Transformação de Pesquisa no mecanismo Databricks Spark. . . . .	437
<b>Capítulo 26: Caches de Pesquisa. . . . .</b>	<b>438</b>
Visão geral de Caches de Pesquisa. . . . .	438
Tipos de Cache de Pesquisa. . . . .	439
Pesquisa sem Cache. . . . .	440
Cache de Pesquisa Estático. . . . .	440
Cache de Pesquisa Persistente. . . . .	441
Recriando um Cache de Pesquisa Persistente. . . . .	441
Cache de Pesquisa Dinâmica. . . . .	442
Cache de Pesquisa Compartilhado. . . . .	442
Regras e Diretrizes para Compartilhar um Cache de Pesquisa. . . . .	443
Comparação de Caches. . . . .	444
Particionamento do Cache para Pesquisas. . . . .	444
<b>Capítulo 27: Cache de Pesquisa Dinâmica. . . . .</b>	<b>445</b>
Visão geral do Cache de Pesquisa Dinâmica. . . . .	445
Usa para um Cache de Pesquisa Dinâmica. . . . .	446
Propriedades do Cache de Pesquisa Dinâmica. . . . .	446
Cache de Pesquisa Dinâmica e Valores de Saída. . . . .	448
Valores de Transformação Pesquisa. . . . .	448
Exemplo de Valores de Transformação de Pesquisa. . . . .	449
Substituição SQL e Cache de Pesquisa Dinâmica. . . . .	451
Configuração de Mapeamento para um Cache de Pesquisa Dinâmica. . . . .	452
Inserir ou Atualizar. . . . .	452
Atualizar ou Inserir. . . . .	453
Cache de Pesquisa Dinâmica e Destino de Sincronização. . . . .	454
Atualizações de Cache de Pesquisa Dinâmica Condicional. . . . .	454
Processamento de Cache Dinâmico Condicional. . . . .	455
Configurando um Cache de Pesquisa Dinâmica Condicional. . . . .	455
Atualização de cache dinâmico com resultados de expressões. . . . .	455
Valores nulos da expressão. . . . .	456
Processamento de Expressão. . . . .	456
Configurando uma Expressão para Atualizações de Cache Dinâmico. . . . .	456

Exemplo de Cache de Pesquisa Dinâmica. . . . .	457
Regras e Diretrizes para Cache de Pesquisa Dinâmica. . . . .	458
<b>Capítulo 28: Transformação de Correspondência. . . . .</b>	<b>459</b>
Transformação de Correspondência - Visão Geral. . . . .	459
Análise de Correspondência. . . . .	460
Análise de Colunas. . . . .	460
Análise de Origem Única e Análise de Origem Dupla. . . . .	461
Análise de Correspondência de Campos e Análise de Correspondência de Identidade. . . . .	461
Grupos na Análise de Correspondência. . . . .	462
Pares de Correspondência e Clusters. . . . .	463
Cálculos de Pontuação de Correspondência. . . . .	464
Pontuações Ponderadas. . . . .	464
Pontuações de Correspondência Nulas. . . . .	464
Opções de Saída de Cluster. . . . .	465
Pontuações de Controle e Pontuações de Vínculo na Análise de Cluster. . . . .	466
Análise de Dados Mestre. . . . .	467
Reutilização de Mapeamentos. . . . .	468
Análise de Correspondência de Identidade e Dados de Indexação Persistente. . . . .	469
Regras e Diretrizes para Dados de Índice Persistentes. . . . .	469
Desempenho do Mapeamento de Correspondência. . . . .	470
Exibindo Dados de Análise de Cluster de Correspondência. . . . .	471
Exibindo Dados de Análise de Desempenho de Correspondência. . . . .	471
Desempenho de Correspondência na Análise de Identidade. . . . .	472
Criando um Armazenamento de Dados para Dados de Índice de Identidade. . . . .	473
Usando o Armazenamento de Dados de Índice na Análise de Origem Única. . . . .	474
Exibições de Transformação de Correspondência. . . . .	475
Portas de Transformação de Correspondência. . . . .	476
Portas de Entrada da Transformação de Correspondência. . . . .	477
Portas de Saída da Transformação de Correspondência. . . . .	477
Códigos de Status de Persistência e Descrições de Status de Persistência. . . . .	478
Portas de Saída e Seleção de Saída de Correspondência. . . . .	481
Mapplets de Correspondência. . . . .	481
Criando um Maplet de Correspondência. . . . .	482
Usando um Maplet de Correspondência. . . . .	482
Configurando uma Operação de Análise de Correspondência. . . . .	483
Transformação de Correspondência em um ambiente não nativo. . . . .	484
Transformação de Correspondência no mecanismo Blaze. . . . .	484
Transformação de Correspondência no mecanismo Spark. . . . .	484
<b>Capítulo 29: Transformações de Correspondência na Análise de Campos. . . . .</b>	<b>485</b>
Análise de Correspondência de Campos. . . . .	485
Fluxo de Processos para Análise de Correspondência de Campos. . . . .	486

Opções de Tipo de Correspondência de Campos. . . . .	486
Estratégias de Correspondência de Campos. . . . .	487
Algoritmos de Correspondência de Campos. . . . .	487
Propriedades da Estratégia de Correspondência de Campos. . . . .	489
Opções de Saída de Correspondência de Campos. . . . .	490
Tipos de Saída de Correspondência. . . . .	490
Propriedades da Saída de Correspondência. . . . .	491
Propriedades Avançadas de Correspondência de Campos. . . . .	492
Exemplo de Análise de Correspondência de Campos. . . . .	493
Criar o Mapeamento. . . . .	493
Amostra de Dados de Entrada. . . . .	494
Configuração da Transformação de Gerador de Chaves. . . . .	494
Configuração da Transformação de Correspondência. . . . .	494
Executar o Visualizador de Dados. . . . .	497
Conclusão. . . . .	497

## **Capítulo 30: Transformações de Correspondência na Análise de Identidade 498**

Análise de Correspondência de Identidade. . . . .	498
Fluxo de Processos para Análise de Correspondência de Identidade. . . . .	499
Propriedades do Tipo de Correspondência de Identidade. . . . .	500
Propriedades do Diretório de Índice e do Diretório de Cache. . . . .	502
Parâmetros de Métodos de Persistência. . . . .	503
Estratégias de Correspondência de Identidade. . . . .	503
Algoritmos de Correspondência de Identidade. . . . .	504
Propriedades da Estratégia de Correspondência de Identidade. . . . .	505
Opções de Saída de Correspondência de Identidade. . . . .	506
Tipos de Saída de Correspondência. . . . .	506
Propriedades da Saída de Correspondência. . . . .	507
Propriedades Avançadas de Correspondência de Identidade. . . . .	509
Estudo de Caso do Índice Persistente. . . . .	510
Exemplo de Análise de Correspondência de Identidade. . . . .	512
Criar o Mapeamento. . . . .	513
Amostra de Dados de Entrada. . . . .	513
Configuração da Transformação de Expressão. . . . .	514
Configuração da Transformação de Correspondência. . . . .	514
Executar o Visualizador de Dados. . . . .	518
Conclusão. . . . .	518

## **Capítulo 31: Transformação Normalizador..... 520**

Visão Geral de Transformação do Normalizador. . . . .	520
Campos de Ocorrência Múltipla. . . . .	521
ID da Coluna Gerada. . . . .	521
Registros de Ocorrência Múltipla. . . . .	522



Definição de Hierarquia de Entrada. . . . .	523
Transformação Normalizador Portas de Entrada. . . . .	524
Mesclar Campos. . . . .	525
Nivelar campos. . . . .	526
Grupos e Portas de Saída de transformação de Normalizador. . . . .	531
Criar um Grupo de Saída. . . . .	532
Atualizar um Grupo de Saída. . . . .	534
Geração de Chave para Grupos de Saída. . . . .	535
Propriedades Avançadas da Transformação de Normalizador. . . . .	535
Gerar Grupos de Saída de Primeiro Nível. . . . .	536
Criando uma Transformação de Normalizador. . . . .	536
Criando uma Transformação de Normalizador de uma Origem Upstream. . . . .	537
Exemplo de Mapeamento de Normalizador. . . . .	537
Exemplo de Mapeamento do Normalizador. . . . .	538
Definição de Exemplo de Normalizador. . . . .	538
Exemplo de Grupos de Entrada e Saída do Normalizador. . . . .	539
Saída de Mapeamento de Exemplo de Normalizador. . . . .	540
Transformação de Normalizador em um ambiente não nativo. . . . .	541
<b>Capítulo 32: Transformação de Mesclagem. . . . .</b>	<b>542</b>
Transformação de Mesclagem - Visão Geral. . . . .	542
Configurando uma Estratégia de Mesclagem. . . . .	542
Propriedades Avançadas da Transformação de Mesclagem. . . . .	543
Transformação de Mesclagem em um ambiente não nativo. . . . .	543
<b>Capítulo 33: Transformação de Analisador. . . . .</b>	<b>544</b>
Transformação de Analisador - Visão Geral. . . . .	544
Modos de Transformação de Analisador. . . . .	545
Quando Usar uma Transformação de Analisador. . . . .	545
Dados de Referência Usados na Transformação de Analisador. . . . .	546
Conjuntos de Padrões. . . . .	547
Modelos Probabilísticos. . . . .	547
Tabelas de Referência. . . . .	548
Expressões Regulares. . . . .	548
Conjuntos de Tokens. . . . .	548
Operações de Análise de Token. . . . .	548
Portas de Análise de Token. . . . .	549
Propriedades de Análise de Token. . . . .	550
Propriedades Gerais. . . . .	550
Propriedades do Modelo Probabilístico. . . . .	551
Propriedades da Tabela de Referência. . . . .	551
Propriedades do Conjunto de Tokens. . . . .	551
Modo de Análise Baseada em Padrões. . . . .	552

Portas de Análise Baseada em Padrões. . . . .	553
Configurando uma Estratégia de Análise de Token. . . . .	553
Configurando uma Estratégia de Análise de Padrão. . . . .	554
Propriedades Avançadas da Transformação de Analisador. . . . .	555
Transformação de Analisador em um ambiente não nativo. . . . .	555
<b>Capítulo 34: Transformação Python. . . . .</b>	<b>556</b>
<b>Capítulo 35: Transformação de Classificação. . . . .</b>	<b>557</b>
Transformação de Classificação - Visão Geral. . . . .	557
Classificando Valores de String. . . . .	558
Propriedades da Transformação de Classificação. . . . .	558
Transformações de Classificação em Mapeamentos Dinâmicos. . . . .	558
Transformação de Classificação - Portas. . . . .	559
Índice de Classificação. . . . .	559
Porta de Classificação. . . . .	560
Definir Portas de Agrupamento. . . . .	560
Parâmetros de Agrupamento. . . . .	561
Caches de Classificação. . . . .	562
Transformação de Classificação - Propriedades Avançadas. . . . .	562
Transformação de Classificação em um ambiente não nativo. . . . .	563
Transformação de Classificação no mecanismo Blaze. . . . .	563
Transformação de Classificação no mecanismo Spark. . . . .	564
Transformação de Classificação no mecanismo Databricks Spark. . . . .	564
<b>Capítulo 36: Transformação de Leitura. . . . .</b>	<b>565</b>
Visão Geral da Transformação de Leitura. . . . .	565
Propriedades da Transformação de Leitura. . . . .	566
Propriedades Gerais. . . . .	567
Propriedades do Objeto de Dados. . . . .	567
Propriedades de Consulta. . . . .	567
Propriedades de Tempo de Execução. . . . .	568
Propriedades de Origens. . . . .	568
Propriedades Avançadas. . . . .	568
Sincronizar Objetos de Dados Relacionais. . . . .	569
Alterar o Objeto de Dados de Origem. . . . .	570
Parametrizar a transformação de Leitura. . . . .	571
Parâmetros da Transformação de Leitura. . . . .	572
Restrições. . . . .	572
Criar uma Transformação de Leitura. . . . .	573
Criando uma Transformação de Leitura no Editor de Mapeamento. . . . .	573

<b>Capítulo 37: Transformação de Relacional para Hierárquica.....</b>	<b>575</b>
Visão geral da transformação de Relacional para Hierarquia. . . . .	575
Exemplo - transformação de Relacional para Hierárquica. . . . .	576
Portas relacionais de entrada e a exibição Visão Geral. . . . .	578
Portas da transformação de Relacional para Hierarquia. . . . .	579
Referências de esquema. . . . .	579
Desenvolvimento de uma transformação de Relacional para Hierárquica. . . . .	579
Criando a transformação de Relacional para Hierárquica. . . . .	579
Criando as portas. . . . .	580
 <b>Capítulo 38: Transformação de Consumidor de Serviço da Web REST.....</b>	 <b>581</b>
Visão Geral da Transformação de Consumidor de Serviço da Web REST. . . . .	581
Processo da Transformação de Consumidor de Serviço da Web REST. . . . .	583
Configuração da Transformação de Consumidor de Serviço da Web REST. . . . .	583
Configuração de Mensagens. . . . .	583
Identificação do Recurso. . . . .	584
Métodos HTTP. . . . .	585
Método Get HTTP. . . . .	585
Método Post HTTP. . . . .	586
Método Put HTTP. . . . .	586
Método Delete HTTP. . . . .	587
Portas da Transformação de Consumidor de Serviço da Web REST. . . . .	588
Portas de Entrada. . . . .	588
Portas de Saída. . . . .	588
Portas de Passagem. . . . .	588
Portas de Argumentos. . . . .	589
Portas de URL. . . . .	589
Portas de Cabeçalho HTTP. . . . .	589
Portas para Cookies. . . . .	590
Portas de Saída XML. . . . .	590
Portas de Código de Resposta. . . . .	590
Mapeamento de Entradas da Transformação de Consumidor de Serviço da Web REST. . . . .	591
Regras e Diretrizes para Mapear Portas de Entrada para Elementos. . . . .	591
Mapeando Portas de Entrada para a Entrada do Método. . . . .	592
Mapeamento da Saída da Transformação de Consumidor de Serviço da Web REST. . . . .	593
Regras e Diretrizes para Mapear Elementos para Portas de Saída. . . . .	594
Personalizar Opções de Exibição. . . . .	594
Mapeando a Saída do Método para Portas de Saída. . . . .	594
Propriedades Avançadas da Transformação de Consumidor de Serviço da Web REST. . . . .	595
Criação da Transformação de Consumidor de Serviço da Web REST. . . . .	596
Criando uma Transformação Consumidor de Serviço da Web REST. . . . .	596
Analisando uma mensagem de resposta JSON que contém matrizes. . . . .	597

Exemplo de mensagem de resposta JSON. . . . .	597
Matrizes não nomeados em uma mensagem de resposta. . . . .	598
<b>Capítulo 39: Transformação de Roteador. . . . .</b>	<b>599</b>
Transformação de Roteador - Visão Geral. . . . .	599
Transformações de Roteador em Mapeamentos Dinâmicos. . . . .	600
Trabalhando com Grupos. . . . .	601
Grupo de Entrada. . . . .	601
Grupos de Saída. . . . .	601
Usando Condições de Filtro de Grupos. . . . .	601
Portas Dinâmicas em Condições de Filtro de Grupos. . . . .	603
Parametrizar o Filtro de Grupo. . . . .	604
Adicionando Grupos. . . . .	604
Trabalhando com Portas. . . . .	604
Conectando Transformações de Roteador em um Mapeamento. . . . .	605
Transformação de Roteador - Propriedades Avançadas. . . . .	605
Transformação de Roteador em um ambiente não nativo. . . . .	606
<b>Capítulo 40: Transformação de Gerador de Sequência. . . . .</b>	<b>607</b>
Visão Geral da Transformação de Gerador de Sequência. . . . .	607
Portas do Gerador de Sequência. . . . .	607
Portas de Passagem. . . . .	608
Porta NEXTVAL. . . . .	608
Propriedades da Transformação Gerador de Sequência. . . . .	610
Valor Inicial. . . . .	610
Valor final. . . . .	610
Valor de Incremento. . . . .	611
Percorrer um Intervalo de Valores. . . . .	611
Redefinir. . . . .	611
Manter Ordem das Linhas. . . . .	611
Objeto de Dados de Sequência. . . . .	612
Criando um objeto de dados de sequência. . . . .	612
Criando uma Transformação de Gerador de Sequência. . . . .	615
Perguntas frequentes. . . . .	616
Transformação de Gerador de Sequência em um ambiente não nativo. . . . .	617
Transformação de Gerador de Sequência no mecanismo Blaze. . . . .	617
Transformação de Gerador de Sequência no mecanismo Spark. . . . .	617
<b>Capítulo 41: Transformação de Ordenador . . . . .</b>	<b>618</b>
Transformação de Classificador - Visão Geral. . . . .	618
Transformações do Classificador em Mapeamentos Dinâmicos. . . . .	619
Desenvolvendo uma Transformação de Ordenador. . . . .	619
Portas de Transformação do Classificador. . . . .	620

Guia Classificação. . . . .	620
Configurar Chaves de Classificação. . . . .	620
Parametrizar as Chaves de Classificação. . . . .	621
Transformação do Classificador - Propriedades Avançadas. . . . .	623
Cache de classificador. . . . .	624
Otimizando o cache de classificador. . . . .	624
Criando uma Transformação de Ordenador. . . . .	624
Criando uma Transformação do Classificador Reutilizável. . . . .	625
Criando uma Transformação de Ordenador Não Reutilizável. . . . .	625
Exemplo de Transformação de Ordenador. . . . .	625
Transformação de Classificador em um ambiente não nativo. . . . .	627
Transformação de Classificador no mecanismo Blaze. . . . .	627
Transformação de Classificador no mecanismo Spark. . . . .	628
Transformação de Classificador no mecanismo Databricks Spark. . . . .	629
<b>Capítulo 42: Transformação SQL . . . . .</b>	<b>630</b>
Transformação SQL - Visão Geral. . . . .	630
Portas da Transformação SQL. . . . .	631
Portas de Entrada. . . . .	632
Portas de Saída. . . . .	632
Portas de Passagem . . . . .	633
Porta SQLException. . . . .	634
Número de Linhas Afetadas. . . . .	635
Propriedades Avançadas da Transformação SQL. . . . .	635
Consulta da Transformação SQL. . . . .	637
Definir a Consulta SQL. . . . .	638
Cardinalidade de Linha de Entrada para Linha de Saída. . . . .	639
Processamento de Instruções de Consulta. . . . .	640
Configuração de Porta. . . . .	641
Contagem Máxima de Linhas de Saída. . . . .	641
Linhas de Erro. . . . .	641
Continuar com Erro de SQL. . . . .	643
Otimização de Filtro com a Transformação SQL. . . . .	643
Otimização de Seleção Antecipada com a Transformação SQL. . . . .	644
Otimização de Envio com a Transformação SQL. . . . .	644
Exemplo de transformação SQL com uma Consulta SQL. . . . .	645
Mapeamento de Objeto de Dados Lógicos. . . . .	645
Tabela Salário. . . . .	646
Tabela Funcionário . . . . .	646
Transformação SQL . . . . .	646
Saída. . . . .	648
Procedimentos Armazenados. . . . .	649
Portas da Transformação SQL para Procedimentos Armazenados. . . . .	649

Conjuntos de Resultados de Procedimentos Armazenados. . . . .	651
Exemplo de Procedimento Armazenado. . . . .	653
Conexão de Transformação SQL. . . . .	654
Criando um Parâmetro de Nome da Conexão. . . . .	654
Criando Manualmente uma Transformação SQL. . . . .	655
Criando uma Transformação SQL a partir de um Procedimento Armazenado. . . . .	656
<b>Capítulo 43: Transformação de Padronizador. . . . .</b>	<b>657</b>
Transformação de Padronizador - Visão Geral. . . . .	657
Estratégias de Padronização. . . . .	657
Propriedades de Padronização. . . . .	658
Configurando uma Estratégia de Padronização. . . . .	659
Propriedades Avançadas da Transformação de Padronizador. . . . .	659
Transformação de Padronizador em um Ambiente Não Nativo. . . . .	660
<b>Capítulo 44: Transformação de União. . . . .</b>	<b>661</b>
Transformação de União - Visão Geral. . . . .	661
Grupos e Portas. . . . .	662
Transformação de União - Propriedades Avançadas. . . . .	662
Transformação de União - Processamento. . . . .	663
Criando uma Transformação de União. . . . .	663
Criando uma Transformação de União Reutilizável. . . . .	663
Criando uma Transformação de União Não Reutilizável. . . . .	663
Transformação de União em um ambiente não nativo. . . . .	664
Transformação da União em um mapeamento de fluxo. . . . .	664
Transformação de União no mecanismo Databricks Spark. . . . .	664
<b>Capítulo 45: Transformação de Estratégia de Atualização. . . . .</b>	<b>665</b>
Transformação de Estratégia de Atualização - Visão Geral. . . . .	665
Definindo a Estratégia de Atualização. . . . .	666
Transformações de Estratégia de Atualização em Mapeamentos Dinâmicos. . . . .	666
Sinalizando Linhas Dentro de um Mapeamento. . . . .	666
Expressões de Estratégia de Atualização. . . . .	667
Transformação de Estratégia de Atualização - Propriedades Avançadas. . . . .	667
Transformações de Agregador e de Estratégia de Atualização. . . . .	668
Especificando Opções de Atualização para Destinos Individuais. . . . .	668
Transformação de Estratégia de Atualização em um ambiente não nativo. . . . .	669
Transformação de Estratégia de Atualização no mecanismo Blaze. . . . .	669
Transformação de Estratégia de Atualização no mecanismo Spark. . . . .	670
<b>Capítulo 46: Transformação de Consumidor de Serviço da Web. . . . .</b>	<b>672</b>
Visão Geral da Transformação de Consumidor de Serviço da Web. . . . .	672
Mensagens SOAP. . . . .	673

Arquivos WSDL. . . . .	673
Operações. . . . .	674
Segurança de Serviços da Web. . . . .	674
Seleção WSDL. . . . .	675
Portas de Transformação de Consumidor de Serviço da Web. . . . .	676
Portas de Entrada de Cabeçalho HTTP. . . . .	676
Outras Portas de Entrada. . . . .	677
Mapeamento de Entrada da Transformação de Consumidor de Serviço da Web. . . . .	678
Regras e Diretrizes para Mapear Portas de Entrada para Nós. . . . .	679
Personalizar Opções de Exibição. . . . .	679
Mapeando Portas de Entrada para a Entrada de Operação. . . . .	679
Transformação de Consumidor de Serviço da Web - Mapeamento de Saída. . . . .	681
Regras e Diretrizes para Mapear Nós para Portas de Saída. . . . .	682
Mapeando a Mensagem SOAP como XML . . . . .	682
Personalizar Opções de Exibição. . . . .	682
Mapeando a Saída de Operação para Portas de Saída. . . . .	683
Propriedades Avançadas da Transformação de Consumidor de Serviço da Web. . . . .	684
Tratamento de Erros de Serviços da Web. . . . .	686
Compactação de Mensagens . . . . .	687
Simultaneidade. . . . .	688
Otimização de Filtro. . . . .	689
Habilitando a Otimização de Seleção Antecipada com a Transformação de Consumidor de Serviço da Web. . . . .	689
Otimização de Envio com a Transformação de Consumidor de Serviço da Web. . . . .	689
Criando uma Transformação de Consumidor de Serviço da Web. . . . .	691
Exemplo de Transformação de Consumidor de Serviço da Web. . . . .	693
Arquivo de Entrada. . . . .	693
Modelo de Objeto de Dados Lógicos. . . . .	693
Mapeamento de Objeto de Dados Lógicos. . . . .	694
Transformação de Consumidor de Serviço da Web. . . . .	694

## **Capítulo 47: Analisando Mensagens SOAP de Serviços da Web..... 697**

Analisando Mensagens SOAP de Serviços da Web - Visão Geral. . . . .	697
Interface do Usuário de Transformações. . . . .	698
Configuração da Saída de Ocorrência Múltipla. . . . .	699
Saída Relacional Normalizada. . . . .	699
Chaves Geradas. . . . .	699
Saída Relacional Desnormalizada. . . . .	700
Saída Relacional Dinamizada. . . . .	701
Analisando Elementos anyType. . . . .	701
Analisando Tipos Derivados. . . . .	702
Analisando Elementos QName. . . . .	703
Analisando Grupos de Substituição. . . . .	704

Analizando Construções XML em Mensagens SOAP. . . . .	704
Elemento de Opção. . . . .	704
Elemento de Lista. . . . .	704
Elemento de União. . . . .	705

## **Capítulo 48: Gerando Mensagens SOAP de Serviços da Web..... 706**

Gerando Visão Geral de Mensagens SOAP de Serviços da Web. . . . .	706
Interface do Usuário de Transformações. . . . .	707
Área Portas de Entrada. . . . .	707
Área Operação. . . . .	708
Relacionamentos de Níveis de Hierarquia e Portas . . . . .	708
Chaves. . . . .	709
Mapear Portas. . . . .	711
Mapear uma Porta . . . . .	712
Mapear um Grupo. . . . .	712
Mapear Várias Portas. . . . .	713
Dinamizando Portas de Ocorrência Múltipla . . . . .	713
Mapear Dados Desnormalizados. . . . .	714
Tipos Derivados e Substituição de Elementos. . . . .	715
Gerando Tipos Derivados. . . . .	716
Gerando Elementos e Atributos anyType. . . . .	716
Gerando Grupos de Substituição. . . . .	717
Gerando Construções XML em Mensagens SOAP. . . . .	717
Elemento de Opção. . . . .	717
Elemento de Lista. . . . .	718
Elemento de União. . . . .	718

## **Capítulo 49: Transformação de Média Ponderada..... 720**

Transformação de Média Ponderada - Visão Geral. . . . .	720
Configurando uma Transformação de Média Ponderada. . . . .	720
Exemplo de Pontuações de Correspondência Ponderada. . . . .	721
Propriedades Avançadas da Transformação de Média Ponderada. . . . .	721
Transformação de média ponderada em um ambiente não nativo. . . . .	722

## **Capítulo 50: Transformação de Janela..... 723**

## **Capítulo 51: Transformação de Gravação..... 724**

Visão Geral da Transformação de Gravação. . . . .	724
Propriedades da transformação de Gravação . . . . .	724
Propriedades Gerais. . . . .	725
Propriedades do Objeto de Dados. . . . .	726
Propriedades de Portas. . . . .	726
Propriedades de Tempo de Execução. . . . .	727



Propriedades de Vinculação em Tempo de Execução. . . . .	727
Propriedades Avançadas. . . . .	728
Criar uma Transformação de Gravação. . . . .	730
Criando uma Transformação de Gravação a partir de um Objeto de Dados. . . . .	731
Criando uma Transformação de Gravação de um Fluxo de Mapeamento. . . . .	731
Criando uma Transformação de Gravação de um Parâmetro. . . . .	732
Criando uma Transformação de Gravação de uma Transformação Existente . . . . .	733
<b>Apêndice A: Delimitadores de Transformação. . . . .</b>	<b>735</b>
Delimitadores de Transformação - Visão Geral. . . . .	735
<b>Índice. . . . .</b>	<b>736</b>

# Prefácio

Use o *Guia do Informatica® Developer Transformation* para aprender sobre a configuração, as diretrizes, o uso e o comportamento em tempo de execução das transformações da Informatica. Entenda como usar as transformações em casos de uso apropriados, dependendo do ambiente e do mecanismo de tempo de execução. Visualize o suporte para cada transformação, dependendo de onde e como você planeja executar o mapeamento.

## Recursos da Informatica

A Informatica oferece uma variedade de recursos de produtos através da Rede da Informatica e outros portais on-line. Use os recursos para obter o máximo de seus produtos e soluções da Informatica e para aprender com outros usuários da Informatica e especialistas no assunto.

### Rede da Informatica

A Rede da Informatica é a porta de entrada para muitos recursos, incluindo a Base de Dados de Conhecimento da Informatica e o Suporte Global a Clientes da Informatica. Para acessar a Rede da Informatica, visite <https://network.informatica.com>.

Como membro da Rede da Informatica, você tem as seguintes opções:

- Pesquisar por recursos do produto na Base de Dados de Conhecimento.
- Visualizar informações sobre disponibilidade de produtos.
- Criar e revisar seus casos de suporte.
- Encontrar a sua Rede de Grupo de Usuários da Informatica local e colaborar com seus colegas.

### Base de Dados de Conhecimento da Informatica

Use a Base de Dados de Conhecimento da Informatica para encontrar recursos de produtos, como artigos de instruções, práticas recomendadas, tutoriais em vídeo e respostas a perguntas frequentes.

Para pesquisar na Base de Dados de Conhecimento, visite <https://search.informatica.com>. Em caso de dúvidas, comentários ou ideias sobre a Base de Dados de Conhecimento, entre em contato com a equipe da Base de Dados de Conhecimento da Informatica em [KB\\_Feedback@informatica.com](mailto:KB_Feedback@informatica.com).

## Documentação da Informatica

Use o Portal de Documentação da Informatica para explorar uma extensa biblioteca de documentação para versões de produtos atuais e recentes. Para explorar o Portal de Documentação, visite <https://docs.informatica.com>.

Em caso de dúvidas, comentários ou ideias sobre a documentação do produto, entre em contato com a equipe da Documentação da Informatica em [infa\\_documentation@informatica.com](mailto:infa_documentation@informatica.com).

## Matrizes de Disponibilidade de Produto da Informatica

As Matrizes de Disponibilidade de Produto (PAMs) indicam as versões dos sistemas operacionais, os bancos de dados e tipos de fontes e destinos de dados com os quais uma versão de produto é compatível. Veja as PAMs da Informatica em <https://network.informatica.com/community/informatica-network/product-availability-matrices>.

## Informatica Velocity

O Informatica Velocity é uma coleção de dicas e práticas recomendadas desenvolvidas pelos Serviços Profissionais da Informatica e baseada em experiências reais de centenas de projetos de gerenciamento de dados. O Informatica Velocity representa o conhecimento coletivo dos consultores da Informatica que trabalham com organizações em todo o mundo para planejar, desenvolver, implantar e manter soluções de gerenciamento de dados bem-sucedidas.

Encontre os recursos do Informatica Velocity em <http://velocity.informatica.com>. Se você tiver dúvidas, comentários ou ideias sobre o Informatica Velocity, entre em contato com os Serviços Profissionais da Informatica em [ips@informatica.com](mailto:ips@informatica.com).

## Informatica Marketplace

O Informatica Marketplace é um fórum onde você pode encontrar soluções que ampliam e aprimoram suas implementações da Informatica. Aproveite as centenas de soluções dos desenvolvedores e parceiros da Informatica no Marketplace para melhorar sua produtividade e agilizar o tempo de implementação em seus projetos. Encontre o Informatica Marketplace em <https://marketplace.informatica.com>.

## Suporte Global a Clientes da Informatica

Você pode entrar em contato com um Centro de Suporte Global por telefone ou por meio da Rede da Informatica.

Para descobrir o número de telefone local do Suporte Global a Clientes da Informatica, visite o site da Informatica no seguinte link: <https://www.informatica.com/services-and-training/customer-success-services/contact-us.html>.

Para encontrar recursos de suporte on-line na Rede da Informatica, visite <https://network.informatica.com> e selecione a opção eSupport.

# CAPÍTULO 1

## Introdução às Transformações

Este capítulo inclui os seguintes tópicos:

- [Introdução às Transformações - Visão Geral, 32](#)
- [Transformações nos ambientes nativo e não nativo, 36](#)
- [Lidando com Tipos de Dados de Transformação, 39](#)
- [Desenvolvendo uma Transformação, 41](#)
- [Transformações de vários grupos, 41](#)
- [Expressões nas Transformações, 42](#)
- [Variáveis Locais, 46](#)
- [Valores Padrão para Portas, 49](#)
- [Níveis de rastreamento, 56](#)
- [Transformações Reutilizáveis, 57](#)
- [Transformações Não Reutilizáveis, 58](#)
- [Criando uma Transformação, 58](#)

## Introdução às Transformações - Visão Geral

Uma transformação é um objeto que gera, modifica ou passa dados.

O Informatica Developer fornece um conjunto de transformações que executam funções específicas. Por exemplo, uma transformação Agregador realiza cálculos em grupos de dados.

As transformações em um mapeamento representam as operações executadas pelo Serviço de Integração de Dados nos dados. Os dados passam através de portas de transformação vinculadas a um mapeamento ou mapplet.

As transformações podem ser ativas ou passivas. As transformações podem ser conectadas ou desconectadas do fluxo de dados.

### Transformações Ativas

Uma transformação ativa altera o número de linhas que passam através de uma transformação. Ou altera o tipo de linha.

Por exemplo, a transformação de Filtro está ativa porque remove linhas que não atendem à condição do filtro. A transformação de Estratégia de Atualização é ativa porque sinaliza linhas para inserir, excluir, atualizar ou rejeitar.

Não é possível conectar várias transformações ativas ou uma transformação ativa e uma passiva à mesma transformação downstream ou grupo de entrada de transformação, porque o Data Integration Service pode não conseguir concatenar as linhas passadas por transformações ativas.

Por exemplo, uma ramificação em um mapeamento contém uma transformação de Estratégia de Atualização que sinaliza uma linha para exclusão. Outra ramificação contém uma transformação de Estratégia de Atualização que sinaliza uma linha para inserção. Se você conectar essas transformações a um único grupo de entrada de transformação, o Data Integration Service não poderá combinar as operações de exclusão e inserção da linha.

## Transformações passivas

Uma transformação passiva não altera o número de linhas que passam pela transformação e mantém o limite de transação e o tipo de linha.

Você pode conectar várias transformações à mesma transformação downstream ou ao mesmo grupo de entrada de transformação quando todas as transformações nas ramificações upstream são passivas. A transformação que origina a ramificação pode ser ativa ou passiva.

## Transformações Desconectadas

As transformações podem ser conectadas ou desconectadas do fluxo de dados. Uma transformação desconectada não está conectada a outras transformações no mapeamento. Uma transformação desconectada é chamada de outra transformação e retorna um valor para essa transformação.

## Transformações de Várias Estratégias

Uma estratégia é um conjunto de uma ou mais operações que uma transformação pode realizar nos dados. Você pode atribuir um conjunto diferente de portas de entrada e de saída a cada estratégia em uma transformação. A transformação armazena as estratégias definidas por você em um único objeto de transformação.

Use a exibição **Dependências** para exibir as portas usadas por cada estratégia.

Você pode definir várias estratégias nas seguintes transformações:

- Conversor de Maiúsculas/Minúsculas
- Classificador
- Decisão
- Gerador de Chaves
- Rotulador
- Correspondência
- Mesclagem
- Analisador
  - Para usar várias estratégias em uma transformação do Analisador, configure a transformação para analisar tokens.
- Padronizador

## Descrições das Transformações

A Developer tool contém transformações comuns e de qualidade de dados. As transformações comuns estão disponíveis no Informatica Data Quality e no Informatica Data Services. Transformações de qualidade de dados estão disponíveis no Informatica Data Quality.

A seguinte tabela descreve cada transformação:

Transformação	Tipo	Descrição
Validador de Endereço	Ativa ou Passiva/ Conectada	Verifica e aprimora a precisão dos registros de endereço postal, e adiciona informações que ajudam os usuários a selecionar os destinatários de e-mail e a entregar a correspondência.
Associação	Ativa/ Conectada	Cria links entre os registros duplicados que uma transformação de Correspondência atribui a diferentes clusters.
Agregador	Ativa/ Conectada	Executa cálculos de agregação.
Exceção de Registro Inválido	Ativa/ Conectada	Identifica os registros que podem conter erros de dados e carrega os registros para tabelas que um usuário da ferramenta Analyst pode analisar e atualizar.
Conversor de Maiúsculas/Minúsculas	Passiva/ Conectada	Padroniza a distinção entre maiúsculas e minúsculas das cadeias.
Classificador	Passiva/ Conectada	Grava rótulos que resumem as informações em campos de porta de entrada. Use quando os campos contêm quantidades significativas de texto.
Comparação	Passiva/ Conectada	Gera pontuações numéricas que indicam o grau de similaridade entre pares de cadeias de entrada.
Consolidação	Ativa/ Conectada	Cria um registro consolidado dos registros identificados como duplicatas pela transformação de Correspondência.
Mascaramento de Dados	Passiva/ Conectada ou Não Conectada	Substitui dados de produção confidenciais por dados de teste realistas para ambientes que não são de produção.
Processador de Dados	Ativa/ Conectada	Processa formatos de arquivo não estruturados e semiestruturados em um mapeamento.
Decisão	Passiva/ Conectada	Avalia as condições nos dados de entrada e cria saídas com base nos resultados dessas condições.
Exceção de Registro Duplicado	Ativa/ Conectada	Identifica os registros que podem conter informações duplicadas e carrega os registros para tabelas que um usuário da ferramenta Analyst pode analisar e atualizar.
Expressão	Passiva/ Conectada	Calcula um valor.
Filtro	Ativa/ Conectada	Filtra dados.

<b>Transformação</b>	<b>Tipo</b>	<b>Descrição</b>
Hierárquico para Relacional	Ativa/ Conectada	Processa a entrada hierárquica e a transforma em saída relacional.
Java	Ativa ou Passiva/ Conectada	Executa uma lógica de usuário codificada em Java. O repositório armazena o código de bytes para a lógica do usuário.
Associador	Ativa/ Conectada	Reúne dados de diferentes bancos de dados ou de sistemas de arquivos simples.
Gerador de Chave	Ativa/ Conectada	Atribui registros a grupos com base nos valores de dados de uma coluna que você selecionar.
Rotulador	Passiva/ Conectada	Grava rótulos que descrevem os caracteres ou as cadeias em um campo de porta de entrada.
Pesquisa	Ativa ou Passiva/ Conectada ou Não Conectada	Pesquisa e retorna dados de um arquivo simples, um objeto de dados lógicos, uma tabela de referência, uma tabela relacional, uma exibição ou um sinônimo.
Correspondência	Ativa/ Conectada	Gera pontuações que indicam os graus de semelhança entre registros de entrada.
Mesclar	Passiva/ Conectada	Lê os valores de dados de várias colunas de entrada e cria uma única coluna de saída.
Normalizador	Ativa/ Conectada	Processa linhas de origem que contêm dados de ocorrência múltipla e retorna uma linha de destino para cada instância desses dados.
Saída	Passiva/ Conectada	Define linhas de saída de mapplet.
Analizador	Passiva/ Conectada	Analisa os valores de uma porta de entrada em portas de saída separadas com base nos tipos de informações que eles contêm.
Classificação	Ativa/ Conectada	Limita registros a uma faixa superior ou inferior.
Leitura	Passiva/ Conectada	Lê dados de uma origem.
Relacional para Hierárquica	Ativa/ Conectada	Processa a entrada relacional e a transforma em uma saída hierárquica.
Consumidor de Serviço da Web REST	Ativa/ Conectada	Conecta-se a um serviço da Web REST como um cliente de serviços da Web para acessar ou transformar dados.
Roteador	Ativa/ Conectada	Roteia dados em várias transformações com base nas condições do grupo.
Gerador de Sequência	Passiva/ Conectada	Gera uma sequência numérica de valores.

Transformação	Tipo	Descrição
Classificador	Ativa/ Conectada	Classifica dados com base em uma chave de classificação.
SQL	Ativa ou Passiva/ Conectada	Executa consultas SQL em um banco de dados.
Padronizador	Passiva/ Conectada	Gera versões padronizadas de cadeias de entrada.
União	Ativa/ Conectada	Mescla dados de diferentes bancos de dados ou sistemas de arquivos simples.
Estratégia de Atualização	Ativa/ Conectada	Determina se é para inserir, excluir, atualizar ou rejeitar linhas.
Consumidor de Serviço da Web	Ativa/ Conectada	Conecta-se a um serviço da Web como um cliente de serviços da Web para acessar ou transformar dados.
Média ponderada	Passiva/ Conectada	Lê as pontuações de correspondência que uma transformação de Correspondência gera para registros em um conjunto de dados, e calcula uma pontuação média para cada par de registros. Você pode aplicar diferentes pesos para as pontuações que a transformação gera para cada par de registros.
Gravação	Passiva/ Conectada	Representa um destino no qual o mapeamento grava dados.

## Transformações nos ambientes nativo e não nativo

Os mapeamentos executados no ambiente não nativo podem retornar resultados diferentes dos mapeamentos executados no ambiente nativo.

Considere as seguintes diferenças de processamento:

- O ambiente não nativo usa processamento distribuído e processa dados em diferentes nós. Cada nó não tem acesso aos dados que estão sendo processados em outros nós. Como resultado, o mecanismo de tempo de execução pode não conseguir determinar a ordem na qual os dados foram originados. Portanto, quando você executa um mapeamento em um ambiente não nativo e executa o mesmo mapeamento no ambiente nativo, os dois mapeamentos retornam resultados corretos, mas os resultados podem não ser idênticos.
- Cada um dos mecanismos de tempo de execução no ambiente não nativo pode processar a lógica de mapeamento de maneira diferente. No ambiente não nativo, as transformações da Informatica podem ser totalmente suportadas, suportadas com restrições ou não suportadas. Da mesma forma, no ambiente nativo, algumas transformações e comportamentos de transformação da Informatica podem não ser suportados.



A tabela a seguir lista as transformações e o suporte para diferentes mecanismos em um ambiente não nativo:

Transformação	Mecanismos suportados
<i>Transformações não listadas nesta tabela não são suportadas em um ambiente não nativo.</i>	
Validador de Endereço	<ul style="list-style-type: none"> <li>- Blaze</li> <li>- Spark</li> </ul>
Agregador	<ul style="list-style-type: none"> <li>- Blaze</li> <li>- Spark*</li> <li>- Databricks Spark</li> </ul>
Conversor de Maiúsculas/Minúsculas	<ul style="list-style-type: none"> <li>- Blaze</li> <li>- Spark</li> </ul>
Classificador	<ul style="list-style-type: none"> <li>- Blaze</li> <li>- Spark</li> </ul>
Comparação	<ul style="list-style-type: none"> <li>- Blaze</li> <li>- Spark</li> </ul>
Consolidação	<ul style="list-style-type: none"> <li>- Blaze</li> <li>- Spark</li> </ul>
Mascaramento de dados	<ul style="list-style-type: none"> <li>- Blaze</li> <li>- Spark*</li> </ul>
Processador de Dados	<ul style="list-style-type: none"> <li>- Blaze</li> </ul>
Decisão	<ul style="list-style-type: none"> <li>- Blaze</li> <li>- Spark</li> </ul>
Expressão	<ul style="list-style-type: none"> <li>- Blaze</li> <li>- Spark*</li> <li>- Databricks Spark</li> </ul>
Filtro	<ul style="list-style-type: none"> <li>- Blaze</li> <li>- Spark*</li> <li>- Databricks Spark</li> </ul>
Java	<ul style="list-style-type: none"> <li>- Blaze</li> <li>- Spark*</li> </ul>
Associador	<ul style="list-style-type: none"> <li>- Blaze</li> <li>- Spark*</li> <li>- Databricks Spark</li> </ul>
Gerador de Chave	<ul style="list-style-type: none"> <li>- Blaze</li> <li>- Spark</li> </ul>
Rotulador	<ul style="list-style-type: none"> <li>- Blaze</li> <li>- Spark</li> </ul>
Pesquisa	<ul style="list-style-type: none"> <li>- Blaze</li> <li>- Spark*</li> <li>- Databricks Spark</li> </ul>

Transformação	Mecanismos suportados
Correspondência	<ul style="list-style-type: none"> <li>- Blaze</li> <li>- Spark</li> </ul>
Mesclar	<ul style="list-style-type: none"> <li>- Blaze</li> <li>- Spark</li> </ul>
Normalizador	<ul style="list-style-type: none"> <li>- Blaze</li> <li>- Spark*</li> <li>- Databricks Spark</li> </ul>
Analisador	<ul style="list-style-type: none"> <li>- Blaze</li> <li>- Spark</li> </ul>
Python	<ul style="list-style-type: none"> <li>- Spark*</li> <li>- Databricks Spark</li> </ul>
Classificação	<ul style="list-style-type: none"> <li>- Blaze</li> <li>- Spark*</li> <li>- Databricks Spark</li> </ul>
Roteador	<ul style="list-style-type: none"> <li>- Blaze</li> <li>- Spark*</li> <li>- Databricks Spark</li> </ul>
Gerador de Sequência	<ul style="list-style-type: none"> <li>- Blaze</li> <li>- Spark</li> </ul>
Classificador	<ul style="list-style-type: none"> <li>- Blaze</li> <li>- Spark*</li> <li>- Databricks Spark</li> </ul>
Padronizador	<ul style="list-style-type: none"> <li>- Blaze</li> <li>- Spark</li> </ul>
União	<ul style="list-style-type: none"> <li>- Blaze</li> <li>- Spark*</li> <li>- Databricks Spark</li> </ul>
Atualizar Estratégia	<ul style="list-style-type: none"> <li>- Blaze</li> <li>- Spark</li> </ul>
Média ponderada	<ul style="list-style-type: none"> <li>- Blaze</li> <li>- Spark</li> </ul>
Janela	<ul style="list-style-type: none"> <li>- Spark**</li> </ul>
<p>* Com suporte para mapeamentos de fluxo e em lote.</p> <p>** Com suporte apenas para mapeamentos de fluxo. Para obter mais informações, consulte o Guia do Usuário do Data Engineering Streaming.</p>	

# Lidando com Tipos de Dados de Transformação

Transformações podem processar funções específicas de tipos de dados ou permitir a passagem de dados sem o processamento desses dados. Com base em transformações, o Serviço de Integração de Dados processa alguns tipos de dados, como Decimal, Registro de Data/Hora com Fuso Horário e Registro de Data/Hora com Fuso Horário Local.

## Tipo de Dados Decimal

Você pode usar o tipo de dados Decimal para ler e gravar dados em arquivos simples e bancos de dados com suporte, como o Oracle, o Microsoft SQL Server, o IBM DB2 e o ODBC.

Para transformações que oferecem suporte a uma precisão de até 38 dígitos, a precisão é de 1 a 38 dígitos, e a escala é de 0 a 38.

### Transformações com Suporte para o Tipo de Dados Decimal

As seguintes transformações oferecem suporte ao tipo de dados Decimal com uma precisão de até 38 dígitos e podem realizar cálculos nos dados:

- Agregador
- Mascaramento de Dados
- Expressão
- Filtro
- Java
- Associador
- Pesquisa
- Normalizador
- Classificação
- Roteador
- Gerador de Sequência
- Classificador
- União
- Estratégia de Atualização

### Transformações com Suporte de Passagem para o Tipo de Dados Decimal

Algumas transformações só podem transmitir dados decimais com uma precisão de até 38 dígitos através da transformação. A transformação não pode realizar cálculos nos dados. Quando você usa um tipo de dados Decimal com precisão de até 38 dígitos para realizar cálculos nas transformações, o Serviço de Integração de Dados processa esse tipo de dados como Duplo.

As seguintes transformações têm suporte de passagem para o tipo de dados Decimal com precisão de até 38 dígitos:

- Processador de Dados
- Hierárquico para Relacional
- Transformação de Consumidor de Serviço da Web REST
- SQL

- Consumidor de Serviços da Web

### Transformações sem Suporte para o Tipo de Dados Decimal

Algumas transformações não oferecem suporte ao tipo de dados Decimal, como transformações de Qualidade de Dados.

A seguinte lista aplicável de transformações de Qualidade de Dados não oferece suporte ao tipo de dados Decimal com precisão de até 38 dígitos:

- Validador de Endereço
- Associação
- Conversor de Maiúsculas/Minúsculas
- Classificador
- Comparação
- Consolidação
- Decisão
- Gerador de Chave
- Rotulador
- Correspondência
- Mesclar
- Analisador
- Padronizador
- Média Ponderada

Para transformações sem suporte para o tipo de dados Decimal 38, quando um tipo de dados Decimal tem precisão maior que 28 dígitos, o Serviço de Integração de Dados converte valores decimais em duplos no modo de alta precisão.

## Registro de Data/Hora com Fuso Horário

O Registro de Data/Hora com Fuso Horário é uma variante do tipo de dados Registro de Data/Hora que inclui um deslocamento de fuso horário ou uma região de fuso horário.

As seguintes transformações oferecem suporte ao tipo de dados Registro de Data/Hora com Fuso Horário:

- Agregador
- Expressão
- Filtro
- Java
- Associador
- Pesquisa
- Normalizador
- Classificação
- Roteador
- Gerador de Sequência
- Classificador

- União
- Estratégia de Atualização

O suporte a passagens implica que é possível transmitir os dados pelas transformações, mas não é possível executar funções no tipo de dados Registro de Data/Hora com Fuso Horário.

As seguintes transformações oferecem suporte a passagens para o tipo de dados Registro de Data/Hora com Fuso Horário:

- Mascaramento de Dados
- Processador de Dados
- Hierárquico para Relacional
- SQL

## Registro de Data/Hora com Fuso Horário Local

Registro de Data/Hora com Fuso Horário Local tem suporte implícito por transformações, pois sua funcionalidade é equivalente a Registro de Data/Hora.

# Desenvolvendo uma Transformação

Ao criar um mapeamento, você adiciona transformações e as configura para lidar com os dados de acordo com uma finalidade comercial.

Conclua as seguintes tarefas para desenvolver uma transformação e incorporá-la em um mapeamento:

1. Adicione uma transformação não reutilizável a um mapeamento ou um mapplet. Ou crie uma transformação reutilizável que pode ser adicionada a vários mapeamentos ou mapplets.
2. Configure a transformação. Cada tipo de transformação tem um conjunto exclusivo de opções que você pode configurar.
3. Se a transformação for reutilizável, adicione-a ao mapeamento ou mapplet.
4. Vincule a transformação a outros objetos no mapeamento ou mapplet.

Arraste portas de objetos upstream para as portas de entrada da transformação. Arraste portas de saída da transformação para portas em objetos downstream. Algumas transformações usam portas predefinidas que você pode selecionar.

**Nota:** Se criar uma transformação reutilizável, adicione as portas de entrada e saída necessárias antes de vincular a transformação a outros objetos. Não é possível adicionar portas à instância de transformação na tela do mapplet ou mapeamento. Para atualizar as portas em uma transformação reutilizável, abra o objeto de transformação do projeto do repositório e adicione as portas.

## Transformações de vários grupos

Uma transformação pode ter vários grupos de entrada e de saída. Um grupo é um conjunto de portas que define uma linha de dados de entrada ou de saída.

Um grupo é análogo a uma tabela em uma definição de origem ou de destino relacional. A maioria das transformações tem um grupo de entrada e um de saída. No entanto, algumas têm vários grupos de entrada, de saída ou ambos. Um grupo é a representação de uma linha de dados que entram ou saem de uma transformação.

Todas as transformações de vários grupos são transformações ativas. Você não pode conectar várias transformações ativas ou uma transformação ativa e uma passiva à mesma transformação downstream ou ao mesmo grupo de entrada da transformação.

Algumas transformações de vários grupos de entrada exigem que o Serviço de Integração bloqueie os dados em um grupo de entrada enquanto espera por uma linha de um grupo de entrada diferente. Uma transformação de bloqueio é uma transformação de vários grupos de entrada que bloqueia dados de entrada. As seguintes transformações são transformações de bloqueio:

- Transformação Personalizar com a propriedade Entradas Podem Bloquear ativada
- Transformação Separador configurada para a entrada não classificada

Quando você salva ou valida um mapeamento, alguns dos que contêm transformações ativas ou de bloqueio podem não ser válidos.

## Regras e Diretrizes para Transformações de Vários Grupos

Ao conectar as transformações em um mapeamento, você deve considerar algumas regras e diretrizes para a conexão de transformações de vários grupos.

Considere as seguintes regras e diretrizes para transformações de vários grupos:

- Você pode conectar um grupo a uma transformação ou um destino.
- Você pode conectar uma ou mais portas de saída em um grupo aos destinos ou transformações múltiplas.
- Você não pode conectar campos de vários grupos de saída em uma transformação ao mesmo grupo de entrada em outra transformação.
- Você não pode conectar os campos de vários grupos de saída em diferentes transformações ao mesmo grupo de entrada em outra transformação, a menos que cada transformação entre a origem e a transformação seja uma transformação passiva.
- Você não pode conectar campos de vários grupos de saída em uma transformação ao mesmo grupo de entrada em outra transformação, a menos que a outra transformação seja uma transformação de bloqueio.
- Você não pode conectar um campo de saída a vários campos de entrada no mesmo grupo de entrada, a menos que o grupo esteja em uma transformação de Normalizador.

## Expressões nas Transformações

Você pode inserir expressões no **Editor de Expressão** em algumas transformações. As expressões modificam dados ou testam se os dados correspondem às condições.

Crie expressões que usam funções de linguagem de transformação. As funções de linguagem de transformação são funções semelhantes ao SQL que transformam dados.

Insira uma expressão em uma porta que use o valor dos dados de uma porta de entrada ou de entrada e saída. Por exemplo, você tem uma transformação com uma porta de entrada IN\_SALARY que contém os salários de todos os funcionários. Você pode usar os valores da coluna IN\_SALARY posteriormente no

mapeamento. Pode também usar a transformação para calcular o total e a média de salários. A Developer tool exige que você crie uma porta de saída separada para cada valor calculado.

A tabela a seguir lista as transformações nas quais você pode inserir expressões:

Transformação	Expressão	Valor de Retorno
Agregador	Realiza um cálculo agregado com base em todos os dados que passam pela transformação. Você também pode especificar um filtro para registros no cálculo agregado de forma a excluir certos tipos de registro. Por exemplo, pode encontrar o número total e o salário médio de todos os funcionários de uma filial usando essa transformação.	Resultado de um cálculo agregado para uma porta.
Expressão	Realiza um cálculo com base nos valores em uma única linha. Por exemplo, com base no preço e na quantidade de um determinado item, é possível calcular o preço de compra total desse item de linha em um pedido.	Resultado de um cálculo no nível de linha para uma porta.
Filtro	Especifica uma condição usada para filtrar linhas que passam por essa transformação. Por exemplo, se quiser gravar dados de clientes com saldo pendente na tabela BAD_DEBT, você poderá usar a transformação de Filtro para filtrar dados do cliente.	TRUE ou FALSE, dependendo de a linha atender à condição especificada ou não. O Serviço de Integração de Dados transmite linhas que retornam TRUE por essa transformação. A transformação aplica esse valor a cada linha que passa por ela.
Associador	Especifica uma condição avançada usada para associar dados de origem não classificados. Por exemplo, você pode concatenar as portas mestras de nome e sobrenome e depois correspondê-las com a porta detalhada de nome completo.	TRUE ou FALSE, dependendo de a linha atender à condição especificada ou não. Dependendo do tipo de associação selecionada, o Serviço de Integração de Dados adiciona a linha ao conjunto de resultados ou a descarta.
Classificação	Define condições para as linhas incluídas em uma classificação. Por exemplo, você pode classificar os 10 melhores vendedores empregados na organização.	Resultado de um cálculo ou condição para uma porta.
Roteador	Roteia dados em várias transformações com base em uma expressão de grupo. Por exemplo, use essa transformação para comparar salários de funcionários em três níveis diferentes de pagamento. Para tanto, é possível criar três grupos na transformação de Roteador. Por exemplo, crie uma expressão de grupo para cada faixa salarial.	TRUE ou FALSE, dependendo de a linha atender à expressão de grupo selecionada ou não. O Serviço de Integração de Dados transmite linhas que retornam TRUE através de cada grupo definido pelo usuário nessa transformação. Linhas que retornam FALSE passam pelo grupo padrão.
Estratégia de Atualização	Sinaliza uma linha para atualização, inserção, exclusão ou rejeição. Essa transformação é usada quando se deseja controlar atualizações em um destino com base em condições que você aplicada. Por exemplo, você pode usar a transformação de Estratégia de Atualização para sinalizar todas as linhas de clientes para atualização quando o endereço de correspondência tiver sido alterado. Ou, você pode sinalizar todas as linhas de funcionários para rejeição para pessoas que não trabalham mais para a organização.	Código numérico para atualização, inserção, exclusão ou rejeição. A transformação aplica esse valor a cada linha que passa por ela.

## O Editor de Expressão

Use o **Editor de Expressão** para compilar instruções semelhantes ao SQL.

Você pode inserir uma expressão manualmente ou usar o método de apontar e clicar. Selecione funções, portas, variáveis e operadores na interface de apontar e clicar para minimizar erros ao construir expressões. O número máximo de caracteres que você pode incluir em uma expressão é 32.767.

## Nomes de Porta em uma Expressão

Você pode inserir nomes de porta da transformação em uma expressão.

Para transformações conectadas, se você usar nomes de porta em uma expressão, a ferramenta Developer atualizará essa expressão quando você alterar os nomes de porta na transformação. Por exemplo, você escreve uma expressão que determina a diferença entre duas datas, Date\_Promised e Date\_Delivered. Se você alterar o nome da porta Date\_Promised para Due\_Date, a ferramenta Developer alterará o nome da porta Date\_Promised para Due\_Date na expressão.

**Nota:** Você pode propagar o nome Due\_Date para outras transformações não reutilizáveis que dependem desta porta no mapeamento.

## Adicionando uma Expressão a uma Porta

Você pode adicionar uma expressão a uma porta de saída.

1. Na transformação, selecione a porta e abra o **Editor de Expressão**.
2. Insira a expressão. Use as guias Funções e Portas e as teclas de operação.  
**Nota:** Não é possível usar um caractere de escape em uma expressão. Se você incluir um caractere de escape na expressão, a Developer tool poderá exibir um erro de análise.
3. Opcionalmente, adicione comentários à expressão.  
Use indicadores de comentários -- ou //.
4. Clique no botão Validar para validar a expressão.
5. Clique em **OK**.
6. Se a expressão não for válida, corrija os erros de validação e valide a expressão novamente.
7. Quando a expressão for válida, clique em **OK** para fechar o **Editor de Expressão**.

## Comentários em uma Expressão

Você pode adicionar comentários a uma expressão para descrevê-la ou especificar uma URL válida para acessar a documentação comercial sobre a expressão.

Para adicionar comentários dentro da expressão, use os indicadores de comentário -- ou //.

## Validação de Expressão

É necessário validar uma expressão para executar um mapeamento ou visualizar a saída do mapplet.

Use o botão Validar no **Editor de Expressão** para validar uma expressão. Se você não validar a expressão, a ferramenta Developer o fará quando você fechar o **Editor de Expressão**. Se a expressão for inválida, a ferramenta Developer exibirá um aviso. Você poderá salvar a expressão inválida ou modificá-la.



## Testar Expressões

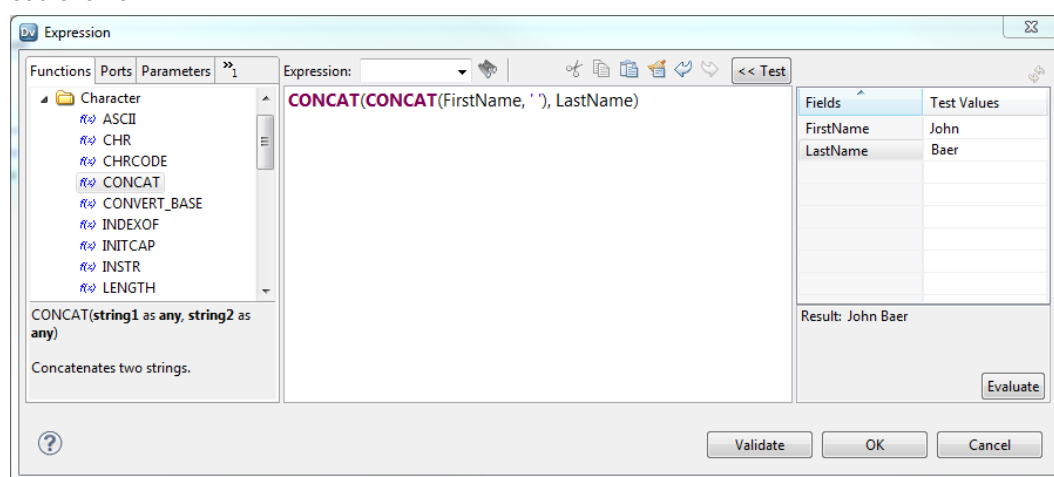
Você pode testar algumas expressões configuradas no Editor de Expressões. Ao testar uma expressão, você insere dados de amostra e, em seguida, avalia a expressão.

Você pode testar expressões ao configurar expressões das seguintes maneiras:

- Em uma porta de saída ou variável na transformação de Expressão
- Na exibição Saídas de Mapeamento de uma transformação de Expressão, após a adição da transformação a um mapeamento

Por exemplo, depois de configurar uma expressão que concatena o nome, um espaço e o sobrenome, você pode inserir dados de amostra para as portas e, em seguida, avaliar a expressão para verificar o resultado.

A seguinte imagem mostra os resultados de uma expressão que concatena uma amostra de nome e sobrenome:



## Cadeias de Formato de Data para Dados de Amostra

Ao testar uma expressão que utiliza uma porta com o tipo de dados Data/Hora ou Registro de Data/Hora com Fuso Horário, você deve inserir os dados de amostra da porta usando a cadeia de formato de data necessária.

Para inserir dados de amostra para uma porta com o tipo de dados Data/Hora, use o formato MM/DD/YYYY HH24:MI:SS. Quando a expressão for avaliada, o Editor de Expressão exibirá o resultado utilizando o formato especificado na expressão. Se você omitir a cadeia de formato na expressão, o Editor de Expressão exibirá o resultado usando o mesmo formato MM/DD/YYYY HH24:MI:SS.

Para inserir dados de amostra para uma porta com o tipo de dados Registro de Data/Hora com Fuso Horário, use o formato MM/DD/YYYY HH24:MI:SS TZR. Quando a expressão for avaliada, o Editor de Expressão exibirá o resultado utilizando o formato YYYY-MM-DD HH24:MI:SS.NS TZR.

## Conversão de Tipo de Dados

Várias funções de expressão e agregação podem gerar dados de um tipo de dados que é diferente dos dados de entrada.

Por exemplo, quando você multiplica dois números decimais com uma precisão de 18 dígitos, o tipo de dados resultante pode ser um decimal com precisão de 28 dígitos.

Para o tipo de dados de entrada Decimal com precisão de 38 dígitos, o resultado de certas operações pode produzir dados que talvez não se encaixem no tipo de dados resultante. Portanto, o usuário pode receber uma exceção de estouro.

As seguintes funções podem exigir uma conversão de tipo de dados para acomodar um aumento no tamanho dos dados quando comparados com os tipos de dados de entrada:

- avg
- cume
- divide
- median
- movingavg
- movingsum
- multiply
- Percentil
- Soma

Por exemplo, quando os dados de entrada são do tipo de dados Integer e você utiliza uma operação de multiplicação, o tipo de dados resultante pode ser bigint. Da mesma forma, quando o tipo de dados de entrada é Decimal com precisão de 18 dígitos, o resultado da operação de multiplicação pode ser grande e no dentro do tipo de dados Decimal com precisão de 28 dígitos.

## Variáveis Locais

Use variáveis locais em transformações de Agregador, Expressão e Classificação para melhorar o desempenho. Você pode fazer referência a variáveis em uma expressão ou usá-las para armazenar dados temporariamente.

Você pode usar variáveis para realizar as seguintes tarefas:

- Armazenar dados temporariamente.
- Simplificar expressões complexas.
- Armazenar valores de linhas anteriores.
- Capturar vários valores de retorno de um procedimento armazenado.
- Comparar valores.
- Armazenar os resultados de uma transformação Pesquisa não concatenada.

## Temporariamente armazenar dados e simplificar expressões complexas

As variáveis aumentam o desempenho quando você insere várias expressões relacionadas na mesma transformação. Você pode definir componentes como variáveis, em vez de analisar e validar os componentes da mesma expressão várias vezes na transformação.

Por exemplo, se uma transformação de Agregador usar a mesma condição de filtro antes de calcular somas e médias, você poderá definir essa condição como variável e reutilizá-la nos cálculos agregados.

Você pode simplificar as expressões complexas. Se um Agregador incluir o mesmo cálculo em diversas expressões, você poderá aumentar o desempenho criando uma variável para armazenar os resultados dos cálculos.

Por exemplo, você pode criar as seguintes expressões para encontrar o salário médio e o salário total usando os mesmos dados:

```
AVG( SALARY, ( ( JOB_STATUS = 'Full-time' ) AND (OFFICE_ID = 1000 ) ) )  
SUM( SALARY, ( ( JOB_STATUS = 'Full-time' ) AND (OFFICE_ID = 1000 ) ) )
```

Em vez de inserir os mesmos argumentos para os dois cálculos, você pode criar uma porta variável para cada condição no cálculo e modificar a expressão para usar as variáveis.

A tabela a seguir mostra como usar variáveis para simplificar expressões complexas e armazenar dados temporariamente:

Porta	Valor
V_CONDITION1	JOB_STATUS = 'Full-time'
V_CONDITION2	OFFICE_ID = 1000
AVG_SALARY	AVG(SALARY, (V_CONDITION1 E V_CONDITION2) )
SUM_SALARY	SUM(SALARY, (V_CONDITION1 AND V_CONDITION2) )

## Armazenar Valores nas Linhas

Você pode configurar variáveis em transformações para armazenar dados de linhas de origem. Você pode utilizar as variáveis em expressões de transformação.

Por exemplo, um arquivo de origem contendo as seguintes linhas:

```
California  
California  
California  
Hawaii  
Hawaii  
New Mexico  
New Mexico  
New Mexico
```

Cada linha contém um estado. Você deve contar o número de linhas e retornar a contagem de linhas para cada estado:

```
California,3  
Hawaii,2  
New Mexico,3
```

Você pode configurar uma transformação do Agregador para agrupar as linhas de origem por estado e contar o número de linhas em cada grupo. Configure uma variável na transformação do Agregador para armazenar a contagem de linhas. Defina outra variável para armazenar o nome do estado da linha anterior.

A transformação do Agregador tem as seguintes portas:

Porta	Tipo de Porta	Expressão	Descrição
Estado	Passagem	n/a	O nome de um estado. As linhas de origem são agrupadas por nome do estado. A transformação de Agregador retorna uma linha para cada estado.
State_Count	Variável	IIF (PREVIOUS_STATE = STATE, STATE_COUNT +1, 1)	O número de linhas do Estado atual. Quando o valor da coluna do Estado atual for igual ao da coluna Previous_State, o Serviço de Integração incrementará State_Count. Caso contrário, ele redefinirá State_Count como 1.
Previous_State	Variável	Estado	O valor da coluna Estado na linha anterior. Quando o Serviço de Integração processa uma linha, ele transfere o valor de Estado para Previous_State.
State_Counter	Saída	State_Count	O número de linhas processadas pela transformação de Agregador para um estado. O Serviço de Integração retorna State_Counter uma vez para cada estado.

## Capturar valores de procedimentos armazenados

As variáveis fornecem uma maneira de capturar várias colunas de valores de retorno em procedimentos armazenados.

## Diretrizes para configurar portas variáveis

Considere os seguintes fatores quando configurar portas variáveis em uma transformação:

- **Ordem da porta.** O Serviço de Integração avalia as portas por dependência. A ordem das portas em uma transformação deve corresponder à ordem de avaliação: portas de entrada, portas variáveis, portas de saída.
- **Tipo de dados.** O tipo de dados que você escolhe reflete o valor retornado da expressão inserida.
- **Inicialização de variáveis.** O Serviço de Integração configura valores iniciais em portas variáveis, onde você pode criar contadores.

### Ordem de portas

O Serviço de Integração avalia primeiro as portas de entrada. Em seguida, ele avalia as portas variáveis e, por último, as portas de saída.

O Serviço de Integração avalia portas na seguinte ordem:

1. **Portas de entrada.** O Serviço de Integração avaliará todas as portas antes, uma vez que não depende de nenhuma outra. Portanto, você pode criar portas de entrada em qualquer ordem. O Serviço de Integração não classifica as portas de entrada porque elas não fazem referência a outras portas.
2. **Portas variáveis.** Portas variáveis podem fazer referência a portas de entrada e portas variáveis, mas não a portas de saída. Como as portas variáveis podem fazer referência a portas de entrada, o Serviço de Integração as avalia após as portas de entrada. As variáveis podem fazer referência a outras

variáveis, portanto, a ordem de exibição das portas variáveis é a mesma na qual o Serviço de Integração avalia cada variável.

Por exemplo, se você calcular o valor original de uma construção e ajustá-lo por depreciação, poderá criar o cálculo do valor original como uma porta variável. Essa porta variável deve aparecer antes da porta que ajusta por depreciação.

3. **Portas de saída.** O Serviço de Integração avalia as portas de saída por último, porque elas podem fazer referência a portas de entrada e portas variáveis. A ordem de exibição das portas de saída não importa, já que elas não podem fazer referência a outras portas de saída. Certifique-se de que as portas de saída são exibidas no final da lista de portas.

## Tipo de Dados

Ao configurar uma porta como variável, é possível inserir qualquer expressão ou condição nela. O tipo de dados que você escolhe para a porta reflete o valor de retorno da expressão inserida. Se você especificar uma condição por meio da porta variável, qualquer tipo de dados numérico retornará os valores para TRUE (não zero) e FALSE (zero).

## Inicialização de Variáveis

O Serviço de Integração não define o valor inicial das variáveis como NULL.

Ele usa as seguintes diretrizes para definir os valores iniciais das variáveis:

- zero para portas numéricas
- strings vazias para portas de string
- 01/01/0001 para portas de Data/Hora

Portanto, use variáveis como contadores, que precisam de valor inicial. Por exemplo, você pode criar uma variável numérica com a seguinte expressão:

```
VAR1 + 1
```

Essa expressão conta o número de linhas na porta VAR1. Se o valor inicial da variável fosse definido como NULL, a expressão seria sempre avaliada como NULL. Por isso o valor inicial é definido como zero.

## Valores Padrão para Portas

Todas as transformações usam valores padrão que determinam como o Serviço de Integração trata os valores nulos de entrada e os erros de transformação de saída.

As portas de entrada, saída e entrada/saída têm um valor padrão do sistema que, às vezes, você pode substituir por um valor padrão definido pelo usuário. Os valores padrão contêm diferentes funções em diferentes tipos de portas:

- **Porta de entrada.** O valor padrão do sistema para portas de entrada nula é NULL. O valor padrão aparece como uma lacuna na transformação. Se um valor de entrada for NULL, o Serviço de Integração irá deixá-lo como NULL.
- **Porta de saída.** O valor padrão do sistema para erros de transformação de saída é ERROR. O valor padrão aparece na transformação como ERROR('transformation error'). Se ocorrer um erro de transformação, o Serviço de Integração ignorará a linha. O Serviço de Integração anota todas as linhas de entrada que a função ERROR ignora no arquivo de log.

Os seguintes erros são erros de transformação:

- Erros de conversão de dados, como passar um número para uma função de dados.
- Erros de avaliação de expressão, como dividir por zero.
- Chamadas para uma função ERROR.
- **Porta de passagem.** O valor padrão do sistema para a entrada nula é o mesmo que para as portas de entrada, NULL. O valor padrão do sistema aparece como uma lacuna na transformação. O valor padrão dos erros de transformação de saída é o mesmo que das portas de saída. O valor padrão dos erros de transformação de saída não aparece na transformação.

**Nota:** A Transformação Java converte tipos de dados PowerCenter® em tipos de dados Java, com base no tipo da porta da Transformação Java. Os valores padrão para a entrada nula variam de acordo com o tipo de dados Java.

A tabela a seguir mostra os valores padrão do sistema para portas em transformações conectadas:

Tipo de porta	Valor padrão	Comportamento do Serviço de Integração	Valor padrão com suporte definido pelo usuário
Entrada, Passagem	NULL	O Serviço de Integração transmite todos os valores de entrada nulos como NULL.	Entrada, Entrada/Saída
Saída, Passagem	ERROR	O Serviço de Integração chama a função ERROR para erros de transformação de porta de saída. O Serviço de Integração ignora linhas com erros e grava os dados de entrada e a mensagem de erro no arquivo do log.	Saída

Portas com variáveis não suportam valores padrão. O Serviço de Integração inicializa as portas com variáveis de acordo com o tipo de dados.

Você pode substituir alguns dos valores padrão para alterar o modo como o Serviço de Integração se comporta quando encontra valores de entrada nulos e erros de transformação de saída.

## Valores Padrão Definidos pelo Usuário

Você pode substituir os valores padrão do sistema por valores padrão definidos pelo usuário para as portas compatíveis de entrada, de passagem e de saída em uma transformação conectada.

Use as seguinte regras e diretrizes para substituir os valores padrão do sistema por portas:

- **Portas de entrada.** Você pode inserir valores padrão definidos pelo usuário para as portas de entrada se não desejar que o Serviço de Integração trate os valores nulos como NULL. Se NULL for passado para a porta de entrada, o Serviço de Integração substituirá NULL pelo valor padrão.
- **Portas de saída.** Você poderá inserir valores padrão definidos pelo usuário para as portas de saída se não desejar que o Serviço de Integração ignore a linha ou se desejar que ele grave uma mensagem específica com a linha ignorada no log. Se você definir um valor padrão na porta de saída, o Serviço de Integração substituirá a linha pelo valor padrão quando a porta de saída tiver um erro de transformação.
- **Portas de passagem.** Você pode inserir valores padrão definidos pelo usuário para as portas de passagem se não desejar que o Serviço de Integração trate os valores nulos como NULL. Você não pode inserir valores padrão definidos pelo usuário para os erros de transformação de saída em uma porta de passagem.

**Nota:** O Serviço de Integração ignora valores padrão definidos pelo usuário para as transformações desconectadas. Por exemplo, se você chamar uma transformação de Pesquisa ou de Procedimento Armazenado por meio de uma expressão, o Serviço de Integração ignorará qualquer valor padrão definido pelo usuário e aplicará o valor padrão do sistema.

Use as seguintes opções para inserir valores padrão definidos pelo usuário:

- **Valor constante.** Use qualquer constante (numérica ou texto), incluindo NULL.
- **Expressão constante.** Você pode incluir uma função de transformação com parâmetros constantes.
- **ERROR.** Gere um erro de transformação. Grave a linha e uma mensagem no log do mapeamento ou no log de erro de linha.
- **ABORT.** Anule o mapeamento.

## Valores Constantes

Você pode inserir qualquer valor constante como um valor padrão. O valor constante deve corresponder ao tipo de dados da porta.

Por exemplo, um valor padrão para uma porta numérica deve ser uma constante numérica. Dentre alguns valores constantes, temos:

```
0
9999
NULL
'Unknown Value'
'Null input data'
```

## Expressões Constantes

Uma expressão constante é qualquer expressão que utilize funções de transformação (à exceção de funções de agregação) para registrar expressões constantes. Não é possível usar valores de portas de entrada, entrada/saída ou variáveis em uma expressão constante.

Eis algumas expressões constantes válidas:

```
500 * 1.75
TO_DATE('January 1, 1998, 12:05 AM','MONTH DD, YYYY, HH:MI AM')
ERROR('Null not allowed')
ABORT('Null not allowed')
SESSSTARTTIME
```

Você não pode usar valores de portas na expressão, pois o Serviço de Integração atribui valores padrão para todo o mapeamento ao iniciá-lo.

Os seguintes exemplos não são válidos porque usam valores de portas:

```
AVG(IN_SALARY)
IN_PRICE * IN_QUANTITY
:LKP(LKP_DATES, DATE_SHIPPED)
```

**Nota:** Não é possível chamar um procedimento armazenado ou tabela de pesquisa a partir de uma expressão de valor padrão.

## Funções ERROR e ABORT

Use as funções ERROR e ABORT para os valores padrão da porta de entrada e saída, e os valores de entrada para as portas de entrada/saída. O Serviço de Integração ignora a linha quando encontra a função ERROR. Ele anula o mapeamento quando encontra a função ABORT.

## Valores de Entrada Padrão Definidos pelo Usuário

Você poderá inserir um valor de entrada padrão definido pelo usuário se não desejar que o Serviço de Integração trate os valores nulos como NULL.

Para substituir valores nulos, conclua uma das seguintes tarefas:

- Substitua o valor nulo por um valor constante ou expressão constante.
- Ignore o valor nulo com uma função ERROR.
- Anule o mapeamento com a função ABORT.

A tabela a seguir resume como o Serviço de Integração manipula a entrada nula para portas de entrada e entrada/saída:

Valor Padrão	Tipo do Valor Padrão	Descrição
NULL (exibe em branco)	Sistema	O Serviço de Integração transmite NULL.
Constante ou expressão Constante	Definidos pelo usuário	O Serviço de Integração substitui o valor nulo pelo valor da constante ou da expressão constante.
ERROR	Definidos pelo usuário	O Serviço de Integração trata isto como um erro de transformação: <ul style="list-style-type: none"><li>- Aumenta a contagem de erros de transformação em 1.</li><li>- Ignora a linha e grava a mensagem de erro no arquivo de log ou no log de erros de linha.</li></ul> O Serviço de Integração não grava linhas no arquivo rejeitado.
ABORT	Definidos pelo usuário	O mapeamento é anulado quando o Serviço de Integração encontra um valor de entrada nulo. O Serviço de Integração não aumenta a contagem de erros ou grava linhas no arquivo rejeitado.

### Substituir Valores Nulos

Use um valor ou uma expressão constante para substituir um valor especificado para valores nulos em uma porta.

Por exemplo, se o nome de uma porta de string de entrada for DEPT\_NAME e você quiser substituir os valores nulos pela string 'UNKNOWN DEPT', defina o valor padrão como 'UNKNOWN DEPT'. Dependendo da transformação, o Serviço de Integração passa 'UNKNOWN DEPT' para uma expressão ou variável dentro da transformação ou para a próxima transformação no fluxo de dados.

Por exemplo, o Serviço de Integração substitui todos os valores nulos em uma porta pela string 'UNKNOWN DEPT'.

DEPT_NAME	REPLACED VALUE
Housewares	Housewares
NULL	UNKNOWN DEPT
Produce	Produce

### Ignorar Registros Nulos

Utilize a função ERROR como valor padrão quando não quiser que valores nulos passem para uma transformação. Por exemplo, talvez você queira ignorar uma linha quando o valor de entrada de DEPT\_NAME for NULO. Você poderia usar a seguinte expressão como valor padrão:

```
ERROR('Error. DEPT is NULL')
```



Quando você usa a função ERROR como um valor padrão, o Serviço de Integração ignora a linha com o valor nulo. O Serviço de Integração grava todas as linhas ignoradas pela função ERROR no arquivo de log. Ele não grava essas linhas no arquivo rejeitado.

DEPT_NAME	RETURN VALUE
Housewares	Housewares
NULL	'Error. DEPT is NULL' (Row is skipped)
Produce	Produce

O seguinte log mostra onde o Serviço de Integração ignora a linha com o valor nulo:

```
TE_11019 Port [DEPT_NAME]: Default value is: ERROR(<<Transformation Error>> [error]:
Error. DEPT is NULL
... error('Error. DEPT is NULL')
).
CMN_1053 EXPTRANS: : ERROR: NULL input column DEPT_NAME: Current Input data:
CMN_1053 Input row from SRCTRANS: Rowdata: ( RowType=4 Src Rowid=2 Targ Rowid=2
DEPT_ID (DEPT_ID:Int:): "2"
DEPT_NAME (DEPT_NAME:Char.25:): "NULL"
MANAGER_ID (MANAGER_ID:Int:): "1"
)
```

## Anular o Mapeamento

Use a função ABORT para anular o mapeamento quando o Serviço de Integração encontrar valores de entrada nulos.

## Validação do Valor Padrão

A ferramenta Developer valida os valores padrão à medida que você os insere.

A ferramenta Developer valida os valores padrão quando você salva um mapeamento. Se você inserir um valor padrão que não é válido, a ferramenta Developer marcará o mapeamento como não válido.

## Valores de Saída Padrão Definidos pelo Usuário

Você pode criar valores padrão definidos pelo usuário para substituir os valores padrão do sistema para portas de saída.

Você poderá inserir valores padrão definidos pelo usuário para as portas de saída se não desejar que o Serviço de Integração ignore linhas com erros ou se desejar que ele grave uma mensagem específica com a linha ignorada no log. Você pode inserir valores padrão para concluir as seguintes funções quando o Serviço de Integração encontrar erros na transformação de saída:

- Substituir o erro por um valor constante ou expressão constante. O Serviço de Integração não ignora a linha.
- Anule o mapeamento com a função ABORT.
- Gravar mensagens específicas no log para os erros de transformação.

Você não pode inserir valores de saída padrão definidos pelo usuário, para as portas de entrada/saída.

A tabela a seguir resume como o Serviço de Integração manipula os erros de transformação da porta de saída e os valores padrão nas transformações:

Valor Padrão	Tipo do Valor Padrão	Descrição
Erro de Transformação	Sistema	Quando um erro de transformação ocorrer e você não tiver substituído o valor padrão, o Serviço de Integração executará as seguintes tarefas: <ul style="list-style-type: none"><li>- Aumenta a contagem de erros de transformação em 1.</li><li>- Ignora a linha e grava o erro e a linha de entrada no arquivo de log da sessão ou no log de erro de linha, dependendo da configuração da sessão.</li></ul> O Serviço de Integração não grava a linha no arquivo rejeitado.
Constante ou Expressão Constante	Definidos pelo usuário	O Serviço de Integração substitui o erro pelo valor padrão. O Serviço de Integração não aumenta a contagem de erros e nem grava uma mensagem no log.
ABORT	Definidos pelo usuário	O mapeamento é anulado e o Serviço de Integração grava uma mensagem no log. O Serviço de Integração não aumenta a contagem de erros ou grava linhas no arquivo rejeitado.

## Substituir Erros

Se você não quiser que o Serviço de Integração ignore uma linha quando ocorrer um erro de transformação, use uma constante ou uma expressão constante como valor padrão de uma porta de saída.

Por exemplo, se você tiver uma porta de saída numérica chamada NET\_SALARY e quiser usar o valor constante '9999' quando ocorrer um erro de transformação, atribua o valor padrão 9999 à porta NET\_SALARY. Se houver um erro de transformação (como uma divisão por zero) durante a computação do valor de NET\_SALARY, o Serviço de Integração usará o valor padrão 9999.

## Anular o Mapeamento

Use a função ABORT para interromper a sessão quando o Serviço de Integração encontrar valores de entrada nulos.

## Gravar Mensagens no Log de Mapeamento

Você pode configurar um valor padrão definido pelo usuário em uma porta de saída se desejar que o Serviço de Integração grave uma mensagem específica no log de mapeamento para a linha ignorada. O padrão do sistema é ERROR ('erro de transformação') e o Serviço de Integração grava a mensagem 'erro de transformação' no log junto com a linha ignorada. Se quiser gravar outra mensagem, você poderá substituir a mensagem 'transformation error'.

## Funções ERROR em Expressões da Porta de Saída

Se você inserir uma expressão que usa a função ERROR, o valor padrão definido pelo usuário para a porta de saída poderia substituir essa função.

Por exemplo, insira a expressão a seguir que instrui o Serviço de Integração a usar o valor 'Venda Negativa' quando ele encontrar um erro:

```
IF( TOTAL_SALES>0, TOTAL_SALES, ERROR ('Negative Sale'))
```

Os exemplos a seguir mostram como os valores padrão definidos pelo usuário podem substituir a função ERROR na expressão:

- **Valor de constante ou expressão.** O valor de constante ou expressão substitui a função ERROR na expressão da porta de saída.

Por exemplo, se você inserir '0' como o valor padrão, o Serviço de Integração substitui a função ERROR na expressão da porta de saída. Ele passa o valor 0 quando encontra um erro. Ele não ignora a linha nem grava 'Venda Negativa' no log.

- **ABORT.** A função ABORT substitui a função ERROR na expressão de porta de saída.

Se você usar a função ABORT como o valor padrão, o Serviço de Integração será anulado quando ocorrer um erro de transformação. A função ABORT substitui a função ERROR na expressão de porta de saída.

- **ERROR.** Se você usar a função ERROR como o valor padrão, o Serviço de Integração incluirá a seguinte informação no log:

- Mensagem de erro do valor padrão
- Mensagem de erro indicada na função ERROR na expressão de porta de saída
- Linha ignorada

Por exemplo, você pode substituir o valor padrão pela função ERROR a seguir:

```
ERROR('No default value')
```

O Serviço de Integração ignora a linha e inclui ambas as mensagens de erro no log.

```
TE_7007 Transformation Evaluation Error; current row skipped...
TE_7007 [<<Transformation Error>> [error]: Negative Sale
... error('Negative Sale')
]
Sun Sep 20 13:57:28 1998
TE_11019 Port [OUT_SALES]: Default value is: ERROR(<<Transformation Error>> [error]:
No default value
... error('No default value')
```

## Regras gerais para valores padrão

Use as seguintes regras e diretrizes ao criar valores padrão:

- O valor padrão deve ser um NULL, um valor constante, uma expressão constante, uma função ERROR ou uma função ABORT.
- Para portas de entrada/saída, o Serviço de Integração usa os valores padrão para manipular os valores de entrada nulos. O valor padrão de saída das portas de entrada/saída é sempre ERROR('Transformation Error').
- Portas variáveis não usam valores padrão.
- Você pode atribuir valores padrão ao grupo por portas, nas transformações de Agregador e Classificação.
- Nem todos os tipos de porta em todas as transformações usam os valores padrão definidos pelo usuário. Se uma porta não permite valores padrão definidos pelo usuário, o campo do valor padrão é desativado.
- Nem todas as transformações permitem valores padrão definidos pelo usuário.
- Se uma transformação não for conectada ao fluxo de dados do mapeamento, o Serviço de Integração ignora os valores padrão definidos pelo usuário.
- Se qualquer porta de entrada estiver desconectada, presume-se que o seu valor seja NULL e o Serviço de Integração usa o valor padrão para essa porta de entrada.
- Se um valor padrão da porta de entrada contiver a função ABORT e o valor de entrada for NULL, o Serviço de Integração interromperá o mapeamento imediatamente. Use a função ABORT como valor padrão para restringir valores de entrada nulos. O primeiro valor nulo em uma porta de entrada interrompe o mapeamento

- Se o valor padrão da porta de saída contiver a função ABORT e algum erro de transformação ocorrer para essa porta, o mapeamento será interrompido imediatamente. Use a função ABORT como valor padrão para reforçar as regras estritas para os erros de transformação. O primeiro erro de transformação para essa porta interrompe o mapeamento.
- A função ABORT, os valores constantes e as expressões constantes substituem as funções ERROR configuradas nas expressões da porta de saída.

## Validação do Valor Padrão

A ferramenta Developer valida os valores padrão à medida que você os insere.

A ferramenta Developer valida os valores padrão quando você salva um mapeamento. Se você inserir um valor padrão que não é válido, a ferramenta Developer marcará o mapeamento como não válido.

## Níveis de rastreamento

Quando você configura uma transformação, pode configurar a quantidade de detalhes que o Serviço de Integração de Dados grava no log.

Por padrão, o nível de rastreamento para cada transformação é Normal. Altere o nível de rastreamento para uma configuração Detalhada apenas quando precisar solucionar problemas de uma transformação que não está se comportando conforme o esperado. Defina o nível de rastreamento para Conciso quando desejar que a quantidade mínima de detalhes apareça no log.

Configure a seguinte propriedade na guia Avançado:

### Nível de Rastreamento

Quantidade de detalhes que aparece no log para uma transformação.

A seguinte tabela descreve os níveis de rastreamento:

Nível de Rastreamento	Descrição
Conciso	Registra informações de inicialização, mensagens de erro e notificação de dados rejeitados.
Normal	Registra informações de inicialização e status, erros encontrados e linhas ignoradas devido a erros na linha da transformação. Resume os resultados do mapeamento, mas não no nível das linhas individuais. O padrão é normal.
Inicialização Detalhada	Além do rastreamento normal, registra detalhes adicionais da inicialização, nomes de arquivos de dados e índice usados, e estatísticas de transformação detalhadas.
Dados Detalhados	Além do rastreamento da inicialização detalhado, registra cada linha que passa para o mapeamento. Também observa onde os dados de strings foram truncados para ajustar a precisão de uma coluna e fornece estatísticas detalhadas de transformação. Quando você configura esse nível de rastreamento, os dados de todas as linhas em um bloco são gravados no log quando uma transformação é processada.

# Transformações Reutilizáveis

Transformações reutilizáveis são aquelas que você pode usar em vários mapeamentos ou mapplets.

Por exemplo, você pode criar uma transformação de Expressão que calcule o imposto sobre valor agregado das vendas no Canadá para analisar o custo de fazer negócios nesse país. Em vez de fazer o mesmo trabalho toda vez, você pode criar uma transformação reutilizável. Quando precisar incorporar a transformação a um mapeamento, você adicionará uma instância dela ao mapeamento. Se você mudar a definição da transformação, todas as instâncias dela herdarão as alterações.

A ferramenta Developer armazena cada transformação reutilizável como metadados separados de qualquer mapeamento ou mapplet que usa a transformação. Ela armazena as transformações reutilizáveis em um projeto ou pasta.

Quando você adiciona instâncias de uma transformação reutilizável a mapeamentos, as alterações feitas na transformação podem invalidar o mapeamento ou gerar dados inesperados.

## Instâncias de Transformação Reutilizável e Alterações Herdadas

Ao adicionar uma transformação reutilizável a um mapeamento ou mapplet, você adiciona uma instância da transformação. A definição da transformação ainda existe fora do mapeamento ou mapplet, enquanto uma instância da transformação aparece dentro do mapeamento ou mapplet.

Quando você altera a transformação, as instâncias da transformação refletem essas alterações. Em vez de atualizar a mesma transformação em cada mapeamento que a usa, você pode atualizar a transformação reutilizável uma única vez, e todas as instâncias da transformação herdam a mudança. As instâncias herdam as alterações feitas em portas, expressões, propriedades e no nome da transformação.

## Editando uma Transformação Reutilizável

Quando você edita uma transformação reutilizável, todas as instâncias dessa transformação herdam as alterações. Algumas alterações podem invalidar os mapeamentos que usam a transformação reutilizável.

Você pode abrir a transformação no editor para editar uma transformação reutilizável. Não é possível editar uma instância da transformação em um mapeamento. No entanto, você pode editar as propriedades de tempo de execução da transformação.

Se você fizer alguma das seguintes alterações em uma transformação reutilizável, os mapeamentos que usarem instâncias dela poderão não ser válidos:

- Ao excluir uma ou mais portas de uma transformação, você desconecta a instância de parte ou de todo o fluxo de dados através do mapeamento.
- Quando você altera um tipo de dados da porta, impossibilita o mapeamento de dados dessa porta para outra que use um tipo de dados incompatível.
- Quando você altera um nome de porta, as expressões que se referem à porta deixam de ser válidas.
- Se você inserir uma expressão que não for válida na transformação reutilizável, os mapeamentos que usarem a transformação não serão mais válidos. O Data Integration Service não pode executar mapeamentos que não são válidos.

## Exibições do Editor para uma Transformação Reutilizável

Você define propriedades e cria código Java para uma transformação Java reutilizável em exibições no editor.

Para transformações reutilizáveis, as seguintes exibições estão disponíveis:

### Visão Geral

Insira o nome e a descrição da transformação, e crie e configure portas de entrada e saída.

### Avançado

Defina as propriedades avançadas da transformação.

## Transformações Não Reutilizáveis

Uma transformação não reutilizável é uma transformação criada em um mapeamento específico. Você não pode usar a transformação em nenhum outro mapeamento.

Por exemplo, você pode criar um mapeamento que contenha várias transformações. Cada transformação realiza cálculos nos dados de origem. Você cria uma transformação de Agregador não reutilizável no final do mapeamento para processar os resultados. Ao criar uma transformação não reutilizável, você pode arrastar as portas de uma transformação para outra no mapeamento e criar as portas de entrada.

A ferramenta Developer armazena a transformação de Agregador não reutilizável como metadados que ela mantém com o mapeamento.

## Exibições do Editor para uma Transformação Não Reutilizável

Defina as propriedades de uma transformação não reutilizável nas exibições do editor.

As seguintes exibições são mostradas para as transformações não reutilizáveis:

### Geral

Insira o nome e a descrição da transformação.

### Portas

Crie e configure portas de entrada e saída.

### Avançado

Defina as propriedades avançadas da transformação.

## Criando uma Transformação

Você pode criar uma transformação reutilizável para reutilizar em vários mapeamentos ou mapplets. Ou pode criar uma transformação não reutilizável para usar uma única vez em um mapeamento ou mapplet.

Para criar uma transformação reutilizável, clique em **Arquivo > Novo > Transformação** e conclua o assistente.

Para criar uma transformação não reutilizável em um mapeamento ou mapplet, selecione uma transformação da paleta de Transformação e arraste a transformação para o editor.

Determinadas transformações exigem que você escolha um modo ou execute uma configuração adicional quando criar a transformação. Por exemplo, quando você cria uma transformação de Pesquisa, deve escolher um objeto de dados para usar como uma origem de pesquisa.

Após criar uma transformação, ela aparece no editor. Algumas transformações contêm portas e grupos predefinidos. Outras transformações estão vazias.

## CAPÍTULO 2

# Portas de Transformação

Este capítulo inclui os seguintes tópicos:

- [Visão geral de portas de transformação, 59](#)
- [Criar Portas, 59](#)
- [Configurar Portas, 60](#)
- [Vinculando Portas, 60](#)
- [Propagando Atributos de Porta, 63](#)
- [Copiando portas do Excel, 66](#)

## Visão geral de portas de transformação

Portas de transformação são as colunas individuais de dados que você conecta em um mapeamento ou mapplet. Transformações recebem os dados de portas de entrada e os enviam usando portas de saída. Portas de entrada/saída recebem dados e os transferem inalterados.

Cada objeto de entrada, objeto de saída, mapplet e transformação contém um conjunto de portas. Objetos de entrada fornecem dados e, portanto, contêm apenas portas de saída. Objetos de saída recebem dados e, portanto, contêm apenas portas de entrada. Mapplets contêm apenas portas de entrada e saída. Transformações contêm uma combinação de portas de entrada, saída e entrada/saída, dependendo da transformação e sua aplicação.

Uma porta dinâmica pode receber uma ou mais colunas de uma transformação upstream. Portas dinâmicas recebem colunas novas ou alteradas com base no fluxo de dados.

Depois de criar uma transformação em um mapeamento, crie as portas e defina suas propriedades. Conclua o mapeamento vinculando-a destinos e outras transformações através das portas. Propague atributos de porta para transmitir atributos alterados a uma porta em um mapeamento.

## Criar Portas

Ao criar algumas transformações, você não precisa criar todas as portas manualmente. Por exemplo, você pode criar uma transformação de Pesquisa e fazer referência a uma tabela de pesquisa. Se você exibir as portas de transformação, poderá ver que a transformação tem uma porta de saída para cada coluna na tabela que você referenciou. Você não precisa definir essas portas.

Crie uma porta das maneiras a seguir:

- **Arraste uma porta de outra transformação.** Quando você arrasta uma porta de outra transformação, o Designer criará uma porta com as mesmas propriedades e vincula as duas portas.
- **Clique no botão Adicionar na guia Portas.** O Designer criará uma porta vazia que você pode configurar.

## Configurar Portas

Ao definir as portas de transformação, você define as propriedades de porta. As propriedades de porta incluem o nome da porta, o tipo de dados, o tipo de porta e o valor padrão.

Configure as seguintes propriedades de porta:

- **Nome da porta.** O nome da porta. Use as seguintes convenções ao nomear portas:
  - Comece com uma letra de bytes simples ou duplos, ou sublinhado de bytes duplos (\_).
  - Os nomes de porta podem conter um dos seguintes caracteres de bytes simples ou duplos: uma letra, um número, o sublinhado (\_), \$, #, ou @.
- **Tipo de dados, precisão e escala.** Se você planejar inserir uma expressão ou uma condição, verifique se o tipo de dados corresponde ao valor de retorno da expressão.
- **Tipo de porta.** As transformações podem conter uma combinação de tipos de porta de entrada, de saída, de entrada/saída e variável.
- **Valor padrão.** Atribua um valor padrão a uma porta que contenha valores nulos ou um erro de transformação de saída. Você pode substituir o valor padrão em algumas portas.
- **Descrição.** Uma descrição da porta.
- **Outras propriedades.** Algumas transformações têm propriedades específicas, como expressões ou propriedades de agrupamento.

## Vinculando Portas

Depois de adicionar e configurar objetos de entrada, saída, transformação e mapplet em um mapeamento, conclua esse mapeamento vinculando portas entre os objetos de mapeamento.

A ferramenta Developer cria a conexão somente quando esta atende aos requisitos de concatenação e validação de links.

Você pode deixar as portas desconectadas. O Data Integration Service ignora portas não conectadas.

Ao vincular portas entre objetos de entrada, transformações, mapplets e objetos de saída, você pode criar os seguintes tipos de link:

- Um para um
- Um para muitos

Você pode vincular portas manual ou automaticamente.



## Links de Um para Um

Vincule uma porta em um objeto de entrada ou transformação a uma porta em um objeto de saída ou transformação.

## Links de Um para Muitos

Quando você deseja usar os mesmos dados para diferentes finalidades, é possível vincular a porta que fornece esses dados a várias portas no mapeamento.

Você pode criar um link de um para muitos das seguintes maneiras:

- Vincule uma porta a várias transformações ou objetos de saída.
- Vincule várias portas em uma transformação a várias transformações ou objetos de saída.

Por exemplo, você deseja usar informações salariais para calcular o salário médio em uma agência bancária por meio da transformação de Agregador. É possível usar as mesmas informações em uma transformação de Expressão configurada para calcular o pagamento mensal de cada funcionário.

## Vinculando Portas Manualmente

É possível vincular manualmente uma ou várias portas.

Arraste uma porta de um objeto de entrada ou transformação até a porta de um objeto de saída ou transformação.

Use a tecla Ctrl ou Shift para selecionar várias portas a serem vinculadas a outra transformação ou objeto de saída. A ferramenta Developer vincula as portas, começando com o par na parte superior. Ela vincula todas as portas que atendem aos requisitos de validação.

Quando você arrasta uma porta até outra vazia, a ferramenta Developer copia essa porta e cria um link.

## Vinculando Portas Automaticamente

Ao vincular portas automaticamente, você pode vincular por posição ou por nome.

Quando você vincula portas automaticamente por nome, é possível especificar um prefixo ou sufixo por meio do qual criar esse link. Use prefixos ou sufixos para indicar onde as portas ocorrem em um mapeamento.

### Vinculando Portas por Nome

Quando você vincula portas por nome, a ferramenta Developer adiciona links entre as portas de entrada e saída que apresentam o mesmo nome. Vincule por nome quando usar os mesmos nomes de portas em transformações.

É possível vincular portas com base em prefixos e sufixos que você define. Use prefixos ou sufixos para indicar onde as portas ocorrem em um mapeamento. Vincule por nome e prefixo ou sufixo quando usar prefixos ou sufixos em nomes de porta para distinguir quando eles ocorrem no mapeamento ou no mapplet.

A vinculação por nome não diferencia maiúsculas de minúsculas.

1. Clique em **Mapeamento > Vinculação Automática**.  
A caixa de diálogo **Vinculação Automática** é exibida.
2. Selecione um objeto na janela **De** como a origem do link.
3. Selecione um objeto na janela **Para** como o destino do link.

4. Selecione **Nome**.
5. Opcionalmente, clique em **Mostrar Avançado** para vincular portas com base em prefixos e sufixos.
6. Clique em **OK**.

## Vinculando Portas por Posição

Quando você vincula por posição, a ferramenta Developer vincula cada porta de saída à porta de entrada correspondente. Por exemplo, a primeira porta de saída é vinculada à primeira porta de entrada, a segunda porta de saída é vinculada à segunda porta de entrada e assim por diante. Vincule por posição quando você criar transformações com portas relacionadas na mesma ordem.

1. Clique em **Mapeamento > Vinculação Automática**.

A caixa de diálogo **Vinculação Automática** é exibida.

2. Selecione um objeto na janela **De** como a origem do link.
3. Selecione um objeto na janela **Para** como o destino do link.
4. Selecione **Posição** e clique em **OK**.

A ferramenta Developer vincula cada porta de saída à porta de entrada correspondente. Por exemplo, a primeira porta de saída é vinculada à primeira porta de entrada, a segunda porta de saída é vinculada à segunda porta de entrada e assim por diante.

## Regras e Diretrizes para Vincular Portas

Certas regras e diretrizes são aplicáveis quando você vincula portas.

Considere as seguintes regras e diretrizes quando conectar objetos de mapeamento:

- Se a ferramenta Developer detectar um erro quando você tentar vincular portas entre dois objetos de mapeamento, ela exibirá um símbolo indicando que as portas não podem ser vinculadas.
- Siga a lógica do fluxo de dados no mapeamento. É possível vincular os seguintes tipos de portas:
  - A porta de recepção deve ser uma porta de entrada ou de entrada/saída.
  - A porta de origem deve ser uma porta de saída ou de entrada/saída.
  - Não é possível vincular portas de entrada a portas de entrada ou portas de saída a portas de saída.
- Você deve vincular pelo menos uma porta de um grupo de entrada a uma transformação ascendente.
- Você deve vincular pelo menos uma porta de um grupo de saída a uma transformação descendente.
- É possível vincular portas de uma transformação ativa ou de um grupo de saída de uma transformação ativa a um grupo de entrada de outra transformação.
- Não é possível conectar uma transformação ativa e uma transformação passiva à mesma transformação descendente ou grupo de entrada de transformação.
- Não é possível conectar mais de uma transformação ativa à mesma transformação descendente ou grupo de entrada de transformação.
- É possível conectar qualquer número de transformações passivas à mesma transformação descendente, grupo de entrada de transformação ou destino.
- É possível vincular portas de dois grupos de saída na mesma transformação a uma transformação de Associador configurada para dados classificados, quando os dados de ambos os grupos de saída estão classificados.

- Apenas é possível vincular portas a tipos de dados compatíveis. A ferramenta Developer verificará se consegue mapear entre os dois tipos de dados antes de os vincular. O Data Integration Service não pode transformar dados entre portas com tipos de dados incompatíveis.
- A ferramenta Developer marca alguns mapeamentos como não válidos quando eles violam a validação do fluxo de dados.

## Propagando Atributos de Porta

Propague atributos de porta para transmitir atributos alterados a uma porta em um mapeamento.

1. No editor, selecione uma porta em uma transformação.
2. Clique em **Mapeamento > Propagar Atributos**.  
A caixa de diálogo **Propagar Atributos** é exibida.
3. Selecione uma direção para propagar atributos.
4. Selecione os atributos que você deseja propagar.
5. Opcionalmente, visualize os resultados.
6. Clique em **Aplicar**.  
A ferramenta Developer propaga os atributos de porta.

## Tipos de Dependência

Quando você propaga atributos de porta, a ferramenta Developer atualiza dependências.

A ferramenta Developer pode atualizar as seguintes dependências:

- Dependências do caminho de link
- Dependências implícitas

## Dependências do Caminho de Link

Uma dependência de caminho de link é uma dependência entre uma porta propagada e as portas em seu caminho de link.

Quando você propaga dependências em um caminho de link, a ferramenta Developer atualiza todas as portas de entrada e de entrada/saída em seu caminho de link progressivo e todas as portas de saída e de entrada/saída em seu caminho de link retroativo. A ferramenta Developer realiza as seguintes atualizações:

- Atualiza o nome da porta, o tipo de dados, a precisão, a escala e a descrição de todas as portas no caminho do link da porta propagada.
- Atualiza todas as expressões ou condições que fazem referência à porta propagada com o nome de porta alterado.
- Atualiza a propriedade da porta associada em uma transformação de Pesquisa dinâmica quando o nome da porta associada é alterado.

## Dependências Implícitas

Uma dependência implícita é uma dependência dentro de uma transformação entre duas portas com base em uma expressão ou condição.

É possível propagar tipos de dados, precisão, escala e descrição para portas com dependências implícitas. Também é possível analisar condições e expressões para identificar as dependências implícitas da porta propagada. Todas as portas com dependências implícitas são portas de saída ou de entrada/saída.

Quando você inclui condições, a ferramenta Developer atualiza as seguintes dependências:

- Dependências do caminho de link
- Portas de saída usadas na mesma condição de pesquisa que a porta propagada
- Portas associadas em transformações de Pesquisa dinâmicas associadas à porta propagada
- Portas mestres usadas na mesma condição de junção que a porta de detalhes

Quando você inclui expressões, a ferramenta Developer atualiza as seguintes dependências:

- Dependências do caminho de link
- Portas de saída contendo uma expressão que usa a porta propagada

A ferramenta Developer não propaga para dependências implícitas na mesma transformação. Você deve propagar os atributos alterados a partir de outra transformação. Por exemplo, quando você altera o tipo de dados de uma porta usada em uma condição de pesquisa e propaga essa alteração a partir da transformação de Pesquisa, a ferramenta Developer não propaga a alteração para a outra porta que depende da condição na mesma transformação de Pesquisa.

## Atributos de Porta Propagados por Transformação

A Developer tool propaga dependências e atributos para cada transformação.

A tabela a seguir descreve as dependências e os atributos que a Developer tool propaga para cada transformação:

Transformação	Dependência	Atributos Propagados
Validador de Endereço	Nenhum.	Nenhum. Essa transformação tem nomes de portas e tipos de dados predefinidos.
Agregador	<ul style="list-style-type: none"><li>- Portas no caminho de link</li><li>- Expressão</li><li>- Dependências implícitas</li></ul>	<ul style="list-style-type: none"><li>- Nome de porta, tipo de dados, precisão, escala, descrição</li><li>- Nome da porta</li><li>- Tipo de dados, precisão, escala</li></ul>
Associação	<ul style="list-style-type: none"><li>- Portas no caminho de link</li></ul>	<ul style="list-style-type: none"><li>- Nome de porta, tipo de dados, precisão, escala, descrição</li></ul>
Conversor de Maiúsculas/Minúsculas	<ul style="list-style-type: none"><li>- Portas no caminho de link</li></ul>	<ul style="list-style-type: none"><li>- Nome de porta, tipo de dados, precisão, escala, descrição</li></ul>
Classificador	<ul style="list-style-type: none"><li>- Portas no caminho de link</li></ul>	<ul style="list-style-type: none"><li>- Nome de porta, tipo de dados, precisão, escala, descrição</li></ul>
Comparação	<ul style="list-style-type: none"><li>- Portas no caminho de link</li></ul>	<ul style="list-style-type: none"><li>- Nome de porta, tipo de dados, precisão, escala, descrição</li></ul>

<b>Transformação</b>	<b>Dependência</b>	<b>Atributos Propagados</b>
Consolidador	Nenhum.	Nenhum. Essa transformação tem nomes de portas e tipos de dados predefinidos.
Mascaramento de Dados	- Portas no caminho de link	- Nome de porta, tipo de dados, precisão, escala, descrição
Processador de Dados	- Portas no caminho de link	- Nome de porta, tipo de dados, precisão, escala, descrição
Decisão	- Portas de downstream no caminho de link	- Nome de porta, tipo de dados, precisão, escala, descrição
Expressão	- Portas no caminho de link - Expressão - Dependências implícitas	- Nome de porta, tipo de dados, precisão, escala, descrição - Nome da porta - Tipo de dados, precisão, escala
Filtro	- Portas no caminho de link - Condição	- Nome de porta, tipo de dados, precisão, escala, descrição - Nome da porta
Hierárquico para Relacional	- Portas no caminho de link	- Nome de porta, tipo de dados, precisão, escala, descrição
Associador	- Portas no caminho de link - Condição - Dependências Implícitas	- Nome de porta, tipo de dados, precisão, escala, descrição - Nome da porta - Tipo de dados, precisão, escala
Gerador de Chave	- Portas no caminho de link	- Nome de porta, tipo de dados, precisão, escala, descrição
Rotulador	- Portas no caminho de link	- Nome de porta, tipo de dados, precisão, escala, descrição
Pesquisa	- Portas no caminho de link - Condição - Portas associadas (pesquisa dinâmica) - Dependências Implícitas	- Nome de porta, tipo de dados, precisão, escala, descrição - Nome da porta - Nome da porta - Tipo de dados, precisão, escala
Correspondência	- Portas no caminho de link	- Nome de porta, tipo de dados, precisão, escala, descrição
Mesclar	- Portas no caminho de link	- Nome de porta, tipo de dados, precisão, escala, descrição
Normalizador	- Portas no caminho de link	- Nome da porta
Analisador	- Portas no caminho de link	- Nome de porta, tipo de dados, precisão, escala, descrição
Classificação	- Portas no caminho de link - Expressão - Dependências implícitas	- Nome de porta, tipo de dados, precisão, escala, descrição - Nome da porta - Tipo de dados, precisão, escala

Transformação	Dependência	Atributos Propagados
Leitura		
Consumidor de Serviço da Web REST	- Portas no caminho de link	- Nome de porta, tipo de dados, precisão, escala, descrição
Roteador	- Portas no caminho de link - Condição	- Nome de porta, tipo de dados, precisão, escala, descrição - Nome da porta
Gerador de Sequência	- Portas no caminho de link	- Nome de porta, tipo de dados, precisão, escala, descrição
Classificador	- Portas no caminho de link	- Nome de porta, tipo de dados, precisão, escala, descrição
SQL	- Portas no caminho de link	- Nome de porta, tipo de dados, precisão, escala, descrição
Padronizador	- Portas no caminho de link	- Nome de porta, tipo de dados, precisão, escala, descrição
União	- Portas no caminho de link - Dependências implícitas	- Nome de porta, tipo de dados, precisão, escala, descrição - Tipo de dados, precisão, escala
Estratégia de Atualização	- Portas no caminho de link - Expressão - Dependências implícitas	- Nome de porta, tipo de dados, precisão, escala, descrição - Nome da porta - Tipo de dados, precisão, escala
Média Ponderada	- Portas no caminho de link	- Nome de porta, tipo de dados, precisão, escala, descrição
Gravação		

## Copiando portas do Excel

Você pode configurar portas e propriedades de portas no Excel e copiá-las para portas de transformação na Developer tool. As propriedades de porta incluem o nome da coluna, o tipo de dados, a precisão e a escala. Talvez você queira fazer isso quando precisar desenvolver ou editar uma transformação com muitas portas.

Você pode copiar metadados para os seguintes tipos de transformação:

- Agregador
- Expressão
- Filtro
- Java
- Associador
- Pesquisa

- Normalizador
- Classificação
- Leitura
- Roteador
- Sequência
- Classificador
- SQL
- União
- Estratégia de Atualização
- Consumidor de Serviço da Web
- Janela
- Gravação

## Editando transformações no Excel

Quando você precisa editar uma grande parte de uma transformação, você não precisa alterar cada valor na Developer tool. Em vez disso, você pode copiar as portas de transformação para o Excel, alterar todos os valores simultaneamente usando o Preenchimento Automático e depois **Colar (Substituir)** as portas de transformação novamente na Developer tool.

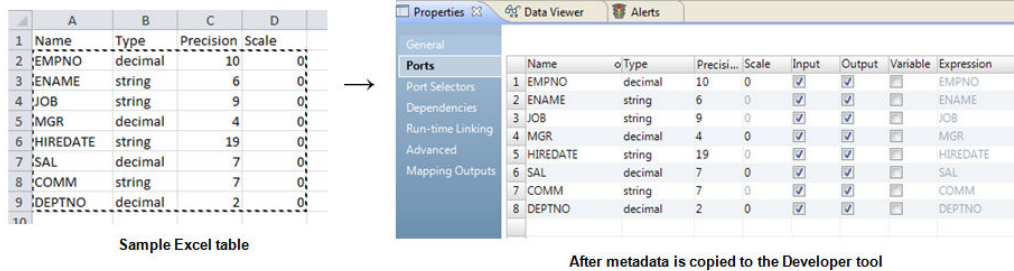
1. No editor de mapeamento da Developer tool, selecione a transformação da qual você deseja copiar as portas.
2. Para copiar a transformação original da Developer tool, clique com o botão direito do mouse em Portas e clique em **Selecionar tudo**.
3. Copie as portas para uma planilha do Excel.
4. Faça alterações na planilha do Excel. Se estiver alterando grandes volumes de metadados, você pode usar o recurso de Preenchimento Automático no Excel. Isso permite preencher dados com base em células adjacentes arrastando a alça de preenchimento. Consulte [“Exemplo: Editando uma transformação no Excel” na página 68](#) para obter mais informações.
5. Copie os metadados do Excel.
6. Para atualizar a transformação com as alterações, clique com o botão direito do mouse em Portas e clique em **Colar (Substituir)**.

## Copiando metadados para a Developer tool

Você pode criar portas de transformação no Excel e copiá-las para a Developer tool.

1. Crie um mapeamento na Developer tool com as transformações necessárias.
2. Defina metadados para uma transformação no Excel.
3. Copie os metadados do Excel.
4. Para mover os metadados para a transformação na Developer tool, clique com o botão direito do mouse em Portas e clique em **Colar (Substituir)**.

A seguinte imagem mostra uma tabela de exemplo do Excel e a transformação resultante após a cópia dos metadados para a Developer tool:



**Nota:** Você deve confirmar que os valores em cada célula são válidos antes de copiá-los para uma transformação. Por exemplo, um tipo de cadeia não pode ter um valor de escala diferente de "0". Valores de precisão não podem ser palavras e valores de tipo não podem ser números. Se os metadados estiverem incorretos, uma mensagem de erro será exibida.

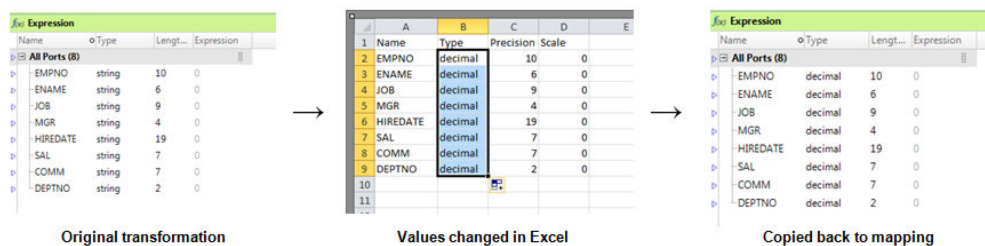
## Exemplo: Editando uma transformação no Excel

Você está desenvolvendo uma transformação e precisa alterar todos os tipos de dados de cadeia para decimais. Em vez de alterar cada campo individualmente, você pode fazer alterações globais no Excel e copiá-las para a Developer tool.

1. Clique com o botão direito do mouse em Portas, clique em **Selecionar tudo** e cole os metadados no Excel.
2. Altere o primeiro valor de tipo de dados de "cadeia" para "decimal" e, em seguida, use a alça de preenchimento para alterar automaticamente as células restantes na coluna.
3. Para atualizar a transformação com as alterações, copie os metadados do Excel, clique com o botão direito do mouse em Portas e clique em **Colar (Substituir)**.

Usando Excel, é possível evitar a necessidade de alterar cada campo individualmente.

A seguinte imagem mostra o processo de mover uma transformação para o Excel usando o Preenchimento Automático para alterar determinados valores e, em seguida, copiando a transformação novamente na Developer tool:





## Regras e diretrizes para copiar do Excel

Considere as seguintes regras e diretrizes ao copiar metadados do Excel para a Developer tool:

- Você deve confirmar que os valores em cada célula são válidos antes de copiá-los para uma transformação. Por exemplo, um tipo de cadeia não pode ter um valor de escala diferente de "0". Valores de precisão não podem ser palavras e valores de tipo não podem ser números. Se os metadados estiverem incorretos, uma mensagem de erro será exibida.
- O local para onde você pode copiar metadados depende do tipo de transformação que você está atualizando. Por exemplo, a opção Colar (Substituir) pode estar indisponível quando você clica com o botão direito do mouse na exibição Propriedades de uma transformação. No entanto, a opção ainda estará disponível quando você clicar com o botão direito do mouse diretamente nas portas de transformação do editor.

## CAPÍTULO 3

# Caches de Transformação

Este capítulo inclui os seguintes tópicos:

- [Visão Geral dos Caches de Transformação, 70](#)
- [Tipos de Cache, 71](#)
- [Arquivos de Cache, 71](#)
- [Tamanho do Cache, 72](#)
- [Aumento do Tamanho do Cache pelo Serviço de Integração de Dados, 74](#)
- [Tamanho do Cache para Caches Particionados, 75](#)
- [Otimização do Tamanho do Cache, 75](#)

## Visão Geral dos Caches de Transformação

O Serviço de Integração de Dados aloca a memória em cache das transformações de Agregador, de Associador, de Pesquisa, de Classificação e de Classificador em um mapeamento. O Serviço de Integração de Dados cria caches de índice e de dados para as transformações de Agregador, de Associador, de Pesquisa e de Classificação. O Serviço de Integração de Dados cria um cache para a transformação de Classificador.

Você pode configurar os tamanhos do cache para essas transformações. O tamanho do cache determina a quantidade de memória que o Serviço de Integração de Dados aloca para cada cache de transformação no início de uma execução de mapeamento.

Se o tamanho do cache for maior do que a memória disponível na máquina, o Serviço de Integração de Dados não poderá alocar memória suficiente e a execução do mapeamento falhará.

Se o tamanho do cache for menor que a quantidade de memória necessária para executar a transformação, o Serviço de Integração de Dados processará parte da transformação na memória e armazenará os dados excedentes nos arquivos de cache. Quando o Serviço de Integração de Dados pagina os arquivos de cache no disco, o tempo de processamento aumenta. Para obter o melhor desempenho, configure o tamanho do cache para que o Serviço de Integração de Dados possa processar a transformação completa na memória.

Por padrão, o Serviço de Integração de Dados calcula automaticamente os requisitos de memória no tempo de execução, com base na quantidade máxima de memória que o serviço pode alocar. Depois de executar um mapeamento no modo de cache automático, você pode ajustar os tamanhos do cache para as transformações. Analise as estatísticas de transformação no log de mapeamento para determinar os tamanhos do cache necessários para obter o melhor desempenho, e configure os tamanhos do cache específicos para as transformações.

## Tipos de Cache

As transformações de Agregador, Associador, Pesquisa e Classificação exigem um cache de índice e um cache de dados. O Serviço de Integração de Dados armazena valores de chave no cache de índice e valores de saída no cache de dados. As transformações de Classificador exigem um único cache. O Serviço de Integração de Dados armazena chaves de classificação e os dados a serem classificados no cache de Classificador.

A seguinte tabela descreve o tipo de informações que o Serviço de Integração de Dados armazena em cada cache:

Objeto de Mapeamento	Tipos e Descrições de Cache
Agregador	<ul style="list-style-type: none"><li>- Índice. Armazena valores de grupo como configurado no grupo por portas.</li><li>- Dados. Armazena cálculos com base no grupo por portas.</li></ul>
Associador	<ul style="list-style-type: none"><li>- Índice. Armazena todas as linhas mestre na condição de união que têm chaves exclusivas.</li><li>- Dados. Armazena linhas de origem mestre.</li></ul>
Pesquisa	<ul style="list-style-type: none"><li>- Índice. Armazena informações de condição de pesquisa.</li><li>- Dados. Armazena dados de pesquisa que não são armazenados no cache de índice.</li></ul>
Classificação	<ul style="list-style-type: none"><li>- Índice. Armazena valores de grupo como configurado no grupo por portas.</li><li>- Dados. Armazena informações de classificação com base no grupo por portas.</li></ul>
Classificador	<ul style="list-style-type: none"><li>- Classificador. Armazena chaves de classificação e dados.</li></ul>

## Arquivos de Cache

Quando você executa um mapeamento, o Serviço de Integração de Dados cria pelo menos um arquivo de cache para cada transformação de Agregador, de Associador, de Pesquisa, de Classificação e de Classificador. Se o Serviço de Integração de Dados não puder executar uma transformação na memória, ele gravará os dados excedentes nos arquivos de cache.

A seguinte tabela descreve os tipos de arquivos de cache que o Serviço de Integração de Dados cria para diferentes objetos de mapeamento:

Objeto de Mapeamento	Arquivo de Cache
Transformações de Agregador, Associador, Pesquisa e Classificação	O Serviço de Integração de Dados cria os seguintes tipos de arquivos de cache: <ul style="list-style-type: none"><li>- Um arquivo de cabeçalho para cada cache de índice e cache de dados</li><li>- Um arquivo de dados para cada cache de índice e cache de dados</li></ul>
Transformação de Classificador	O Serviço de Integração de Dados cria um arquivo de cache de classificador.

Quando você executa um mapeamento, o Serviço de Integração de Dados grava uma mensagem no log de mapeamento que indica o nome do arquivo de cache e o nome da transformação. Quando um mapeamento é concluído, o Serviço de Integração de Dados libera a memória cache e, geralmente, exclui os arquivos de

cache. Você pode encontrar arquivos de cache de índice e de dados no diretório de cache nas seguintes circunstâncias:

- Você configura a transformação de Pesquisa para usar um cache persistente.
- O mapeamento não foi concluído com êxito. A próxima vez que você executar o mapeamento, o Serviço de Integração de Dados excluirá os arquivos de cache existentes e criará novos arquivos.

Como a gravação nos arquivos de cache pode tornar lento o desempenho do mapeamento, configure os tamanhos de cache para executar a transformação na memória.

## Diretório de Arquivo de Cache

Para as transformações de Agregador, de Associador, de Pesquisa e de Classificação, o Serviço de Integração de Dados cria os arquivos de cache no diretório especificado da propriedade do Diretório de Cache. Para transformações de Classificador, o serviço cria os arquivos de cache no diretório especificado da propriedade de Diretório de Trabalho.

Se o processo do Serviço de Integração de Dados não encontrar o diretório, o mapeamento falhará e uma mensagem será gravada no log de mapeamento, indicando que não foi possível criar ou abrir o arquivo de cache. O Serviço de Integração de Dados pode criar vários arquivos de cache. O número de arquivos de cache é limitado pelo espaço em disco disponível no diretório de cache.

## Tamanho do Cache

O tamanho do cache determina a quantidade de memória que o Serviço de Integração de Dados aloca para cada cache de transformação no início de uma execução de mapeamento. Você pode configurar um tamanho do cache de transformação para usar o modo de cache automático ou para usar um valor específico.

### Tamanho do Cache Automático

Por padrão, o tamanho do cache de transformação é definido como Automático. O Serviço de Integração de Dados calcula automaticamente os requisitos de memória em cache no tempo de execução. Você define a quantidade máxima de memória que o serviço pode alocar.

O Serviço de Integração de Dados usa as seguintes diretrizes para alocar automaticamente a memória:

**Aloca mais memória para transformações com mais tempo de processamento.**

O Serviço de Integração de Dados aloca mais memória para transformações que normalmente têm mais tempo de processamento. Por exemplo, o Serviço de Integração de Dados aloca mais memória à transformação do Classificador porque esta costuma levar mais tempo para ser executada.

**Aloca mais memória para o cache de dados do que para o cache de índice.**

As transformações de Agregador, Associador, Pesquisa e Classificação exigem um cache de índice e um cache de dados. Quando o Serviço de Integração de Dados divide a memória alocada para a transformação entre o cache de índice e o de dados, ele aloca mais memória para o cache de dados.

As transformações do Classificador exigem um único cache. O serviço alocará para o cache de Classificador toda a memória que está alocada para a transformação.

## Máximo de Memória para Tamanho do Cache Automático

Você define a quantidade máxima de memória que o Serviço de Integração de Dados pode alocar para caches de transformação na propriedade Memória Máxima por Solicitação dos módulos do Serviço de Integração de Dados na ferramenta Administrator.

Cada módulo executa tipos diferentes de solicitações que têm requisitos diferentes de memória. Por exemplo, as solicitações de mapeamento e de perfil geralmente solicitam mais memória em cache que as solicitações do serviço SQL ou do serviço Web. Você pode configurar a propriedade Memória Máxima por Solicitação para os seguintes módulos do Serviço de Integração de Dados:

- Módulo de Serviço de Mapeamento
- Módulo do Serviço de Criação de Perfil
- Módulo de Serviço SQL
- Módulo do Serviço da Web

**Nota:** As solicitações do Módulo de Serviço de Mapeamento incluem mapeamentos e execução de mapeamentos das tarefas de Mapeamento em um fluxo de trabalho.

Para o Módulo do Serviço de Criação de Perfil, a Memória Máxima por Solicitação define a quantidade máxima de memória que o Serviço de Integração de Dados pode alocar para cada execução de mapeamento para uma única solicitação de perfil.

Para o restante dos módulos, o comportamento da Memória Máxima por Solicitação depende da configuração do Serviço de Integração de Dados. O comportamento depende da propriedade Opções de Inicialização de Trabalho e da propriedade Tamanho Máximo de Memória no Serviço de Integração de Dados.

A seguinte tabela descreve o comportamento de Memória Máxima por Solicitação para os módulos de mapeamento, de serviço SQL e de serviço Web com base na configuração do Serviço de Integração de Dados:

Configuração do Serviço de Integração de Dados	Comportamento de Memória Máxima por Solicitação
Executa trabalhos em processos de sistema locais ou remotos separados, ou o Tamanho Máximo de Memória é 0 (padrão)	<p>Quantidade máxima de memória, em bytes, que o Serviço de Integração de Dados pode alocar para todas as transformações que usam o modo de cache automático em uma única solicitação.</p> <p>O valor que você define para Memória Máxima por Solicitação afeta somente as transformações que usam o modo de cache automático. O Serviço de Integração de Dados alocará a memória separadamente nas transformações nas quais você configurar um tamanho de cache específico. O total de memória usada pela solicitação pode exceder o valor de Máximo de Memória por Solicitação.</p> <p>Por exemplo, a Memória Máxima por Solicitação é definida como 800 MB. Um mapeamento tem três transformações que exigem cache. Configure duas transformações para usar o modo de cache automático e configure a terceira transformação para usar um total de 500 MB para os tamanhos do cache. O Serviço de Integração de Dados aloca um total de 1.300 MB de memória para todos os caches de transformação.</p>
Executa trabalhos no processo do Serviço de Integração de Dados, e o Tamanho Máximo de Memória é maior que 0	<p>A quantidade máxima de memória em bytes que o Serviço de Integração de Dados pode alocar para uma solicitação.</p> <p>O valor que você definir para a propriedade Memória Máxima por Solicitação afeta todas as transformações. O total de memória usada pela solicitação não pode exceder o valor de Máximo de Memória por Solicitação.</p>

Quando você aumenta a quantidade máxima de memória usada para o modo de cache automático, aumenta o tamanho máximo do cache que pode ser usado para todas as solicitações ao módulo. Você pode

aumentar a quantidade máxima de memória para garantir que nenhum arquivo de cache seja paginado no disco. No entanto, como esse valor é usado em todas as solicitações, o Serviço de Integração de Dados pode alocar mais memória que o necessário para algumas solicitações.

## Tamanho do Cache Específico

Você pode configurar um tamanho do cache específico para uma transformação. O Serviço de Integração de Dados aloca a quantidade especificada de memória para o cache de transformação no início da execução do mapeamento. Configure um valor específico em bytes quando ajustar o tamanho do cache.

Na primeira vez que você configurar um tamanho do cache, use o modo de cache automático. Depois de executar o mapeamento, analise as estatísticas de transformação no log de mapeamento para determinar os tamanhos do cache necessários para executar as transformações na memória. Quando você configura o tamanho do cache para usar o valor especificado no log de mapeamento, pode garantir que nenhuma memória alocada seja desperdiçada. No entanto, o tamanho do cache ideal varia com base no tamanho da fonte de dados. Analise os logs de mapeamento depois que o mapeamento subsequente for executado para monitorar as mudanças no tamanho do cache. Se você configurar um tamanho do cache específico para uma transformação reutilizável, verifique se o tamanho do cache é ideal para cada uso da transformação em um mapeamento.

Para definir os tamanhos do cache específicos, configure os valores de tamanho do cache nas propriedades de transformação da Developer tool.

## Aumento do Tamanho do Cache pelo Serviço de Integração de Dados

O Serviço de Integração de Dados cria cada cache de memória baseado no tamanho de cache configurado. Em algumas situações, o Serviço de Integração de Dados pode aumentar o tamanho de cache configurado porque requer mais memória em cache.

O Serviço de Integração de Dados pode aumentar o tamanho configurado do cache por uma das seguintes razões:

**O tamanho do cache configurado é menor do que o tamanho do cache mínimo requerido para processar a operação.**

O Serviço de Integração de Dados requer uma quantidade mínima de memória para inicializar cada mapeamento. Se o tamanho configurado do cache for menor do que o tamanho mínimo requerido, o Serviço de Integração de Dados aumentará o tamanho configurado do cache para que atenda ao requisito mínimo. Se o Serviço de Integração de Dados não puder alocar a memória mínima requerida, o mapeamento falhará.

**O tamanho do cache configurado não é um múltiplo do tamanho da página do cache.**

O Serviço de Integração de Dados armazena dados em cache nas páginas do cache. As páginas em cache devem se ajustar uniformemente dentro do cache. Por exemplo, se você configurar 10 MB (1.048.576 bytes) para o tamanho do cache e o tamanho da página do cache for 10.000 bytes, o Serviço de Integração de Dados aumentará o tamanho configurado do cache para 1.050.000 bytes para torná-lo um múltiplo do tamanho da página de 10.000 bytes.

Quando o Serviço de Integração de Dados aumenta o tamanho de cache configurado, ele continua a executar o mapeamento e grava as seguintes mensagens no log de mapeamento:

```
INFO: MAPPING,      TE_7212,      Increasing [Index Cache] size for transformation
<transformation name> from <configured cache size> to <new cache size>.
```

INFO: MAPPING, TE\_7212, Increasing [Data Cache] size for transformation  
<transformation name> from <configured cache size> to <new cache size>.

## Tamanho do Cache para Caches Particionados

Se você tiver a opção Particionamento, o particionamento do cache criará um cache separado para cada partição que executa uma transformação de Agregador, de Associador, de Classificação, de Pesquisa ou de Classificador. Durante o particionamento do cache, cada partição armazena dados diferentes em um cache separado. Quando o Serviço de Integração de Dados usa o particionamento do cache para essas transformações, o serviço divide o tamanho do cache alocado entre as partições.

Por exemplo, configure o tamanho do cache de transformação para ter 100 MB. O Serviço de Integração de Dados usa quatro partições para executar a transformação. O serviço divide o valor do tamanho do cache para que cada partição use um máximo de 25 MB para o tamanho do cache.

## Otimização do Tamanho do Cache

Para obter o melhor desempenho do mapeamento, configure os tamanhos do cache para que o Serviço de Integração de Dados possa executar a transformação completa na memória.

Para configurar os tamanhos de cache ideais, execute as seguintes tarefas:

1. Defina o nível de rastreamento para inicialização detalhada.
2. Execute o mapeamento no modo de cache automático.
3. Analise o desempenho do armazenamento em cache no log de mapeamento.
4. Configure valores específicos para os tamanhos do cache.

### Etapa 1. Defina o Nível de Rastreamento para Inicialização Detalhada

Na Developer tool, defina o nível de rastreamento como inicialização detalhada para ativar o Serviço de Integração de Dados para gravar as estatísticas de transformação no log de mapeamento. As estatísticas de transformação listam os tamanhos do cache necessários para obter o melhor desempenho. Por padrão, o nível de rastreamento está definido como normal.

Defina o nível de rastreamento para inicialização detalhada de uma das seguintes maneiras:

- Modifique as propriedades avançadas de cada transformação que usa um cache.
- Modifique as propriedades de configuração de mapeamento padrão se você planejar executar o mapeamento pela primeira vez na Developer tool. Para obter mais informações, consulte o *Guia de Ferramentas da Informatica Developer Tool*.
- Modificar as propriedades avançadas de um aplicativo que contém o mapeamento se você planejar executar o mapeamento implantado pela primeira vez da linha de comando. Para obter mais informações, consulte o *Guia de Ferramentas da Informatica Developer Tool*.

## Etapa 2. Executar o Mapeamento no Modo de Cache Automático

Na primeira vez que você executar o mapeamento, use o modo de cache automático para os tamanhos do cache da transformação.

Você pode executar o mapeamento na Developer tool. Ou você pode adicionar o mapeamento a um aplicativo e implantar o aplicativo no Serviço de Integração de Dados para que você possa executar o mapeamento da linha de comando.

## Etapa 3. Analisar o Desempenho do Cache

Depois de executar o mapeamento no modo de cache automático, analise as estatísticas de transformação no log de mapeamento para determinar os tamanhos do cache necessários para obter o melhor desempenho do mapeamento.

Quando uma transformação de Agregador, de Associador, de Pesquisa ou de Classificação é paginada no disco, o log de mapeamento especifica o índice e os tamanhos do cache de dados necessários para executar a transformação na memória. Por exemplo, execute uma transformação de Agregador chamada AGG\_TRANS. O log de mapeamento contém o seguinte texto:

```
CMN_1791, The index cache size that would hold [1098] aggregate groups of input rows for
[AGG_TRANS], in memory, is [286720] bytes
CMN_1790, The data cache size that would hold [1098] aggregate groups of input rows for
[AGG_TRANS], in memory, is [1774368] bytes
```

O log mostra que o cache de índice requer 286.720 bytes e que o cache de dados requer 1.774.368 bytes para executar a transformação na memória sem pagnar o disco.

Quando uma transformação do Classificador é paginada no disco, o log de mapeamento determina que o Serviço de Integração de Dados fez várias transmissões nos dados de origem. O Serviço de Integração de Dados faz várias transmissões nos dados quando ele tem que pagnar no disco para concluir a classificação. A mensagem especifica o número de bytes exigidos para uma única transmissão, que é quando o Serviço de Integração de Dados lê os dados uma vez e executa a classificação na memória sem pagnar no disco.

Por exemplo, execute uma transformação do Classificador chamada SRT\_TRANS. O log de mapeamento contém o seguinte texto:

```
SORT_40427, Sorter Transformation [SRT_TRANS] required 2-pass sort (1-pass temp I/O:
13126221824 bytes). You may try to set the cache size to 14128 MB or higher for 1-pass
in-memory sort.
```

O log mostra que o cache de Classificador requer 14.128 MB para que o Serviço de Integração de Dados faça uma transmissão de dados.

## Etapa 4. Configurar os Tamanhos de Cache Específicos

Para obter o melhor desempenho, configure os tamanhos do cache da transformação para usar os valores especificados no log de mapeamento. Atualize as propriedades de transformação do tamanho do cache de índice e de dados na Developer tool.

1. Na Developer tool, abra a transformação reutilizável ou não reutilizável.
2. Localize as propriedades do tamanho do cache, dependendo dos seguintes tipos de transformação:

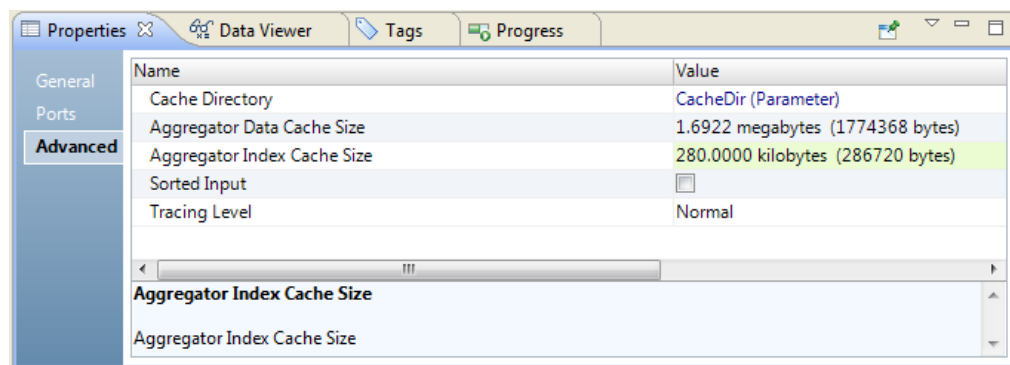
Opção	Descrição
Transformação de Agregador, de Associador, de Classificação ou de Classificador reutilizável	Clique na exibição <b>Avançado</b> .



Opção	Descrição
Transformação de Agregador, de Associador, de Classificação ou de Classificador não reutilizável	Clique na guia <b>Avançado</b> na exibição <b>Propriedades</b> .
Transformação de Pesquisa reutilizável	Clique na exibição <b>Tempo de Execução</b> .
Transformação de Pesquisa não reutilizável	Clique na guia <b>Tempo de Execução</b> na exibição <b>Propriedades</b> .

- Insira os valores em bytes que o log de mapeamento recomendou para os tamanhos do cache de índice e de dados.

A seguinte imagem mostra uma transformação de Agregador não reutilizável que tem valores específicos configurados para os tamanhos do cache de índice e de dados:



- Clique em **Arquivo > Salvar**.

## CAPÍTULO 4

# Transformação de Validador de Endereço

Este capítulo inclui os seguintes tópicos:

- [Transformação de Validador de Endereço - Visão Geral, 79](#)
- [Dados de Referência de Endereço, 79](#)
- [Modos e modelos, 82](#)
- [Grupos de Portas e Seleção de Porta, 82](#)
- [Transformação de Validador de Endereço - Grupos de Portas de Entrada, 82](#)
- [Transformação de Validador de Endereço - Grupos de Portas de Saída, 83](#)
- [Portas de Várias Instâncias, 86](#)
- [Projetos de Validação de Endereço, 87](#)
- [Endereços Formatados e Padrões do Serviço de Correio, 88](#)
- [Conclusão de Endereço Parcial , 89](#)
- [Portas de Status do Validador de Endereço, 90](#)
- [Transformação de Validador de Endereço - Configurações Gerais, 102](#)
- [Propriedades de validação de endereço na janela Preferências, 103](#)
- [Propriedades avançadas de validação de endereço, 106](#)
- [Relatórios de Certificação, 127](#)
- [Configurando uma Transformação de Validador de Endereço, 130](#)
- [Adicionando Portas à Transformação de Validador de Endereço, 130](#)
- [Criando Modelos Definidos pelo Usuário, 131](#)
- [Definindo Modelos do Validador de Endereço, 131](#)
- [Definindo um Relatório de Certificação, 132](#)
- [Transformação de Validador de Endereço em um ambiente não nativo, 132](#)

# Transformação de Validador de Endereço - Visão Geral

A transformação do Validador de Endereço é uma transformação de vários grupos que compara os dados de endereço de entrada com os dados de referência de endereço. Ela determina a precisão dos endereços e corrige os erros. A transformação cria registros que atendem aos padrões de entrega de email para conteúdo e estrutura de dados. A transformação também adiciona informações de status a cada endereço.

A transformação do Validador de Endereço executa as seguintes operações nos dados de endereço:

- A transformação compara os registros de endereço nos dados de origem com as definições de endereço nos dados de referência de endereço.
- Ela gera relatórios de status detalhados sobre a validade de cada endereço de entrada, seu status de entrega e a natureza de qualquer erro ou ambiguidade que ele contenha.
- Ela corrige erros e completa registros de endereço parciais. Para corrigir um endereço, a transformação deve encontrar uma correspondência positiva com um endereço nos dados de referência. A transformação copia os elementos de dados necessários dos dados de referência de endereço para os registros de endereço.
- Ele adiciona informações que não aparecem no endereço padrão mas ajudam na entrega postal, como informações de ponto de entrega e informações de codificação geográfica.
- Ele grava endereços de saída no formato que o projeto de dados e o serviço de correio exigem. Você define o formato quando seleciona as portas de saída na transformação.

## Dados de Referência de Endereço

Um conjunto de dados de referência de endereço descreve os endereços que um serviço de correio nacional reconhece em um país. Antes de realizar a validação de endereço com a transformação do Validador de Endereço, instale os dados de referência de endereço na máquina dos serviços Informatica no domínio. Compre e faça download dos dados de referência de endereço da Informatica.

Instale um arquivo de dados de referência de endereço para cada país identificado pelos dados de endereço da origem. Um país que tem uma grande população poderá exigir vários arquivos. Além disso, você pode instalar arquivos de dados que suplementam ou aprimoram os dados de endereço. O serviço de correio pode usar os dados aprimorados para verificar a precisão dos endereços e para acelerar a entrega de correspondência.

Quando você executa a validação de endereço, a transformação de Validador de Endereço compara cada registro de entrada aos dados de referência de endereço. Se a transformação localizar o endereço de entrada nos dados de referência de endereço, ela poderá atualizar o registro com os dados de endereço corretos e completos. Se você tiver adquirido conjuntos de dados de referência adicionais, a transformação também poderá aprimorar os dados de endereço.

Use a janela **Preferências** na Developer tool para exibir informações sobre os arquivos de dados de referência de endereço na máquina de serviços Informatica no domínio.

## Tipos de Dados de Referência de Endereço

O modo de validação selecionado determina como a transformação compara o endereço de entrada aos dados de referência de endereço.

A transformação do Validador de Endereço pode ler os seguintes tipos de dados de referência de endereço:

### Dados de pesquisa de código de endereço

Instale os dados de pesquisa de código de endereço para recuperar um endereço parcial ou um endereço completo de um valor de código em uma porta de entrada. A completude do endereço depende do nível de suporte ao código de endereço no país ao qual o endereço pertence. Para ler o código de endereço de um endereço de entrada, selecione o as portas específicas do país no grupo de portas Discreto.

Você pode selecionar portas para os seguintes países:

- Áustria. Retorna um endereço para o nível de edifício.
- Alemanha. Retorna um endereço para a localidade, o município ou o nível de rua.
- Japão. Retorna um endereço para o nível de caixa correio exclusivo.
- África do Sul. Retorna um endereço para o nível de rua.
- Coreia do Sul. Retorna um endereço para o nível de caixa correio exclusivo.
- Sérvia. Retorna um endereço para o nível de rua.
- Reino Unido. Retorna um endereço para o nível de caixa correio exclusivo.

A transformação do Validador de Endereço lê dados de pesquisa de código de endereço quando você configura a transformação para execução no modo de pesquisa de código de endereço.

### Dados em lote e interativos

Instale dados em lote e interativos para realizar a validação de endereço em um conjunto de registros de endereço. Use dados em lote e interativos para verificar se os endereços de entrada têm total capacidade de entrega e estão completos com base nos dados postais atuais do serviço de correio nacional.

Quando você configura a transformação para execução no modo em lote, a transformação do Validador de Endereço retorna um único endereço para cada endereço de entrada. Quando você configura a transformação para execução no modo interativo, a transformação do Validador de Endereço retorna um ou mais endereços para cada endereço de entrada.

### Dados CAMEO

Instale dados CAMEO para adicionar dados de segmentação de clientes em registros de endereço residencial. Os dados de segmentação de clientes indicam o nível de renda e as preferências de estilo de vida prováveis dos residentes de cada endereço.

A transformação do Validador de Endereço lê dados CAMEO quando você configura a transformação para execução no modo em lotes ou no modo certificado.

### Dados certificados

Instale dados certificados para verificar se os registros de endereço atendem aos padrões de certificação definidos por um serviço de correio. Um endereço atende a um padrão de certificação se contém elementos de dados que podem identificar uma caixa postal exclusiva, tais como elementos de dados de ponto de entrega. Quando um endereço atende a um padrão de certificação, o serviço de correio cobra uma taxa de entrega reduzida.

Os seguintes países definem padrões de certificação:

- Austrália. Certifica a correspondência de acordo com o padrão AMAS (Sistema de Aprovação de Correspondência de Endereço).
- Canadá. Certifica a correspondência de acordo com o padrão SERP (Programa de Análise e Reconhecimento de Software) padrão.
- França. Certifica a correspondência de acordo com o padrão SNA (Serviço Nacional de Gerenciamento de Endereços).
- Nova Zelândia. Certifica a correspondência de acordo com o padrão SendRight.
- Estados Unidos. Certifica a correspondência de acordo com o padrão CASS (Sistema de Suporte para Precisão de Codificação).

A transformação de Validador de Endereço lê os dados certificados quando você a configura para execução no modo certificado.

#### **Dados de Codificação Geográfica**

Instale dados de codificação geográfica para adicionar dados de codificação geográfica aos registros de endereço. Codificações geográficas são coordenadas de latitude e longitude.

A transformação do Validador de Endereço lê dados de codificação geográfica quando você configura a transformação para execução no modo em lotes ou no modo certificado.

#### **Dados da lista de sugestões**

Instale dados da lista de sugestões para encontrar versões válidas alternativas de um registro de endereço parcial. Use dados da lista de sugestões quando você configurar um mapeamento de validação de endereço para processar registros de endereço um por um, em tempo real. A transformação do Validador de Endereço usa os elementos de dados no endereço parcial para realizar uma verificação de duplicatas nos dados da lista de sugestões. A transformação retorna qualquer endereço válido que inclua as informações no endereço parcial.

A transformação do Validador de Endereço lê dados da lista de sugestões quando você configura a transformação para execução no modo de lista de sugestões.

#### **Dados complementares**

Instale dados complementares para adicionar dados a um registro de endereço que pode ajudar o serviço de correio na entrega de correspondência. Use os dados complementares para adicionar detalhes sobre a área geográfica ou postal contém o endereço. Em alguns países, dados complementares podem fornecer um identificador exclusivo para uma caixa de correio no sistema postal.

A transformação de Validador de Endereço lê os dados complementares quando você a configura para execução no modo em lote ou no modo certificado.

**Nota:** A transformação não lê dados de referência de endereço no modo de reconhecimento de país ou no modo de análise.

## TÓPICOS RELACIONADOS:

- [“ Transformação de Validador de Endereço - Configurações Gerais” na página 102](#)

# Modos e modelos

Ao configurar a transformação de Validador de Endereço, você seleciona o tipo de validação de endereço que essa transformação realiza. A transformação define o tipo de validação como o modo. Selecione o modo na guia **Configurações Gerais** ou como uma propriedade avançada da transformação.

É possível criar um modelo de porta na transformação. Um modelo é um subconjunto de portas de um ou mais grupos de portas. Use um modelo para organizar as portas que você pretende usar em um projeto.

# Grupos de Portas e Seleção de Porta

A transformação de Validador de Endereço contém grupos de portas predefinidos que contêm as portas de entrada e de saída que você pode usar. Ao configurar uma transformação do Validador de Endereço, você procura os grupos e seleciona as portas das quais precisa.

Selecione as portas de entrada que correspondem à estrutura do dados de entrada de endereço. Selecione as portas de saída que contêm os dados de endereço que o projeto requer.

É possível adicionar portas de entrada e de saída diretamente à transformação ou criar um modelo padrão que contenha portas de entrada e de saída. Ao adicionar portas diretamente à transformação, as portas que você seleciona se aplicam somente a essa transformação. Ao adicionar portas ao modelo padrão, as portas que você seleciona se aplicam às transformações de Validador de Endereço futuras que você cria.

Também é possível adicionar à transformação portas de passagem para colunas que você não quer que sejam processadas pela transformação do Validador de Endereço.

# Transformação de Validador de Endereço - Grupos de Portas de Entrada

Para poder conectar dados de endereço a portas de entrada na transformação, procure os grupos de entrada e selecione as portas que correspondem à estrutura e ao conteúdo dos dados de entrada. Procure os grupos de saída e selecione as portas que correspondem aos seus requisitos de dados.

A transformação de Validador de Endereço exibe os grupos de portas em um Modelo Básico e em um Modelo Avançado. Você pode definir a maioria dos endereços usando um grupo de portas no Modelo Básico. Se seus endereços forem altamente complexos, use as portas adicionais disponíveis no Modelo Avançado.

**Nota:** Selecione portas apenas de um grupo de portas de entrada.

A transformação tem os seguintes grupos de portas de entrada:

### Discreto

Use as portas Discreto para ler colunas de dados que contêm informações completas sobre um único elemento de dados, como número da casa, rua ou código postal. Localize o grupo Discreto nos modelos Básico e Avançado.

### Híbrido

Use as portas Híbrido para ler colunas de dados que contêm informações sobre um ou mais elementos de dados. O grupo Híbrido combina portas dos grupos Discreto e de Várias Linhas. Use as portas do grupo Híbrido para criar registros de endereço que você pode enviar para um serviço de correio. As portas do grupo Híbrido estruturam um endereço conforme os padrões do serviço de correio e identificam o tipo de dados de cada linha. Localize o grupo Híbrido nos modelos Básico e Avançado.

### Várias Linhas

Use as portas Várias Linhas para ler colunas de dados que contêm vários elementos de dados. Cada coluna de entrada corresponde a uma linha de um endereço. Para obter melhores resultados, defina os dados de entrada no formato exigido pelo serviço de correio. Selecione as portas Várias Linhas para criar um conjunto de registros de endereço imprimível.

Cada porta Várias Linhas representa uma linha no endereço imprimível, como a seguinte linha de dados de rua:

123 Main Street, Apartment 2

As portas de Várias Linhas não especificam o tipo de dados que aparece em cada linha de endereço. Localize o grupo Várias Linhas nos modelos Básico e Avançado.

### Linha Única

Use as portas Linha Única para ler uma única coluna de dados que contém todos os elementos de endereço para o nível de província e que não inclui um separador entre os elementos. Use a porta Endereço Completo no grupo de portas para enviar os elementos de endereço. O grupo de portas também inclui uma porta País que você pode usar para ler as informações de país referentes ao endereço. Localize o grupo Linha Única nos modelos Básico e Avançado.

## Transformação de Validador de Endereço - Grupos de Portas de Saída

Antes de você poder conectar a transformação de Validador de Endereço a outras transformações ou objetos de dados, determine os tipos de informação necessários e a estrutura que os endereços de saída adotarão.

Procure os grupos de saída e selecione as portas que correspondem aos seus requisitos de dados.

**Nota:** Você pode selecionar portas de vários grupos de saída e pode selecionar portas que têm funcionalidade comum.

A transformação tem os seguintes grupos de saída predefinidos:

### Elementos de Endereço

Grava elementos de dados de rua, como número da casa, número do apartamento e nome de rua para separar as portas. Localize o grupo de Elementos de Endereço nos modelos Básico e Avançado.

**Complementar da AT**

Grava dados em endereços na Áustria que podem ajudar na entrega postal, como dados de código postal em nível de edifício. Localize o grupo Complementar da AT no modelo Básico.

**Complementar de AU**

Grava dados para endereços na Austrália que identificam as regiões geográficas para as quais o Departamento Australiano de Estatísticas atribui os endereços. Localiza o grupo Complementar da AU no modelo Básico.

**Específico da Austrália**

Grava dados em endereços na Austrália que permitem que esses endereços atendam às normas do AMAS (Address Matching Approval System) dos Correios da Austrália. Localize o grupo Específico da Austrália nos modelos Básico e Avançado.

**Complementar da BE**

Grava dados em endereços na Bélgica que podem ajudar na entrega postal. Os dados incluem os códigos de identificação de localidade e região do diretório de estatísticas na Bélgica. Localize o grupo Complementar da BE no modelo Básico.

**Complementar do BR**

Grava dados em endereços no Brasil que podem ajudar na entrega postal, como códigos de identificação do distrito do Instituto Brasileiro de Geografia e Estatística (IBGE). Localize o grupo Complementar do BR no modelo Básico.

**CAMEO**

Gera dados de resumo demográfico e de renda que você pode usar na análise de segmentação de clientes. Localize o grupo CAMEO no modelo Básico.

**Específico do Canadá**

Grava dados em endereços no Canadá que permitem que os endereços atendam às normas do SERP (Software Evaluation and Recognition Program) dos Correios do Canadá. Localize o grupo Específico do Canadá no modelo Básico.

**Complementar da CH**

Grava dados em endereços na Suíça que podem ajudar na entrega postal, como dados de código postal estendidos. Localize o grupo Complementar da CH no modelo Básico.

**Complementar da CZ**

Grava dados para endereços na República Tcheca que podem ajudar a entrega postal, como dados de código postal estendido. Localize o grupo Complementar da CZ no modelo Básico.

**Elementos de Contato**

Contém dados de contato ou pessoa, como nomes, saudações e cargos. Localize o grupo de Elementos de Contato no modelo Avançado.

**País**

Grava o nome ou código do país definido pela Organização Internacional de Normalização (ISO). Localize o grupo de País nos modelos Básico e Avançado.

**Complementar da DE**

Grava dados em endereços na Alemanha que podem ajudar na entrega postal, como dados de código de município e distrito. Localize o grupo Complementar da DE no modelo Básico.



**Complementar da ES**

Grava dados em endereços na Espanha que podem ajudar na entrega postal. Localiza o grupo Complementar da ES no modelo Básico.

**Linha de Endereço Formatada**

Grava endereços formatados para impressão e correspondência. Localize o grupo de Linha de Endereço Formatada nos modelos Básico e Avançado.

**Complementar da FR**

Grava dados em endereços na França que podem ajudar na entrega postal, como códigos de identificação do Instituto Nacional de Estudos de Economia e Estatística (INSEE). Localize o grupo Complementar da FR no modelo Básico.

**Específico da França**

Grava dados em endereços na França que permitem que os endereços atendam aos padrões do Serviço Nacional de Gerenciamento de Endereços (SNA) do La Poste. Localize o grupo Específico da França no modelo Básico.

**Codificação Geográfica**

Gera dados de codificação geográfica de um endereço, como coordenadas de latitude e longitude. Localize o grupo de Codificação Geográfica no modelo Básico.

**Elementos de ID**

Grava o ID de Registro e dados de Chave de Transação. Localize o grupo de Elementos de ID no modelo Avançado.

**IT complementar**

Grava dados em endereços na Itália que podem ajudar na entrega postal. Localiza o grupo Complementar da IT no modelo Básico.

**Complementar do JP**

Grava dados em endereços no Japão que podem ajudar na entrega postal, como códigos Choumei Aza. Localize o grupo de Complementar do JP no modelo Básico.

**Complementar da KR**

Grava dados em endereços na Coreia do Sul que podem ajudar na entrega postal, como identificadores exclusivos que podem especificar versões atuais e não atuais de um determinado endereço. Localiza o grupo complementar da KR no modelo Básico.

**Elementos de Última Linha**

Grava dados que podem ser exibidos na última linha de um endereço doméstico. Localize o grupo de Elementos de Última Linha nos modelos Básico e Avançado.

**Específico da Nova Zelândia**

Grava dados em endereços na Nova Zelândia que permitem que os endereços atendam aos padrões SendRight dos Correios da Nova Zelândia. Localize o grupo Específico da Nova Zelândia no modelo Básico.

**Complementar da PL**

Grava dados em endereços na Polônia que podem ajudar na entrega postal, como dados de Divisão Territorial (TERYT). Localize o grupo Complementar da PL no modelo Básico.

**Resíduo**

Grava elementos de dados que a transformação não pode analisar para outras portas. Localize o grupo de Resíduo nos modelos Básico e Avançado.

**Complementar da RS**

Grava dados em endereços na Sérvia que podem ajudar na entrega postal, como dados de sufixo de código postal. Localize o grupo de Complementar da RS no modelo Básico.

**Complementar da RU**

Grava dados em endereços na Rússia que podem ajudar na entrega postal, como o identificador do Sistema de Endereçamento de Informações Federais do endereço. Localize o grupo Complementar da RU no modelo Básico.

**Informações de Status**

Gera dados detalhados sobre a qualidade de cada endereço de entrada e saída. Localize o grupo de Informações de Status no modelo Básico.

**Complementar do Reino Unido**

Grava dados em endereços no Reino Unido que podem ajudar na entrega postal, como dados de pontos de entrega e dados de levantamento cartográfico. Localize o grupo de Complementar do Reino Unido no modelo Básico.

**Específico dos EUA**

Grava dados em endereços nos Estados Unidos que permitem que os endereços atendam aos padrões do Sistema de Suporte de Precisão de Codificação (CASS) do Serviço Postal dos Estados Unidos. Localize o grupo Específico dos EUA no modelo Básico.

**Complementar dos EUA**

Grava dados geográficos e demográficos, como Padrão de Processamento de Informações Federais (FIPS) para endereços nos Estados Unidos. Localize o grupo de Complementar dos EUA no modelo Básico.

**XML**

Grava dados do registro de endereço em uma estrutura XML definida pela biblioteca de softwares do Validador de Endereço. Localize o grupo XML no modelo Avançado.

**Complementar da ZA**

Grava dados em endereços na África do Sul que podem ajudar na entrega postal, como os dados do Banco de Dados Nacional de Endereços. Localize o grupo Complementar da ZA no modelo Básico.

## Portas de Várias Instâncias

Vários tipos de dados de endereço podem ocorrer mais de uma vez em um endereço. Você pode selecionar várias instâncias de uma porta quando o endereço contiver vários casos de um elemento de dados.

Uma porta de várias instâncias pode conter até seis instâncias. Vários endereços usam uma instância de uma porta para cada elemento de dados que contém. Alguns endereços usam uma segunda instância de uma porta. Um conjunto pequeno de endereços usa mais de uma instância de porta.

Geralmente, a primeira instância de uma porta é o nome principal ou a área mais ampla que a porta identifica. Você deve verificar o relacionamento entre as instâncias de porta de todas as portas que selecionar.

### Exemplo de Portas de Rua Completa

Um registro de endereço do Reino Unido poderá conter dois nomes de rua se uma delas for parte de um plano de rua mais amplo.

A seguinte tabela contém um endereço que usa duas portas Rua Completa:

Porta	Dados
Número da Rua Completo 1	1A
Rua Completa 1	THE PHYGTLE
Rua Completa 2	SOUTH STREET
Nome da Localidade 1	NORFOLK
Código Postal 1	NR25 7QE

No exemplo, os dados de rua em Rua Completa 1 são dependentes dos dados de rua em Rua Completa 2. Os dados de Número de Rua Completo 1 se referem aos dados de Rua Completa 1.

**Nota:** Embora a Rua Completa 1 especifique a localização da caixa de correio, a Rua Completa 2 pode ser a rua maior.

### Exemplo de Portas de Contato

Um registro de endereço poderá conter vários contatos quando um deles fizer parte da mesma residência.

A seguinte tabela contém um endereço que usa duas portas Nome de Contato:

Porta	Dados
Nome do Contato 1	MR. JOHN DOE
Nome do Contato 2	MS. JANE DOE
Linha de Endereço Formatada 1	2 MCGRATH PLACE EAST
Linha de Endereço Formatada 2	ST. JOHN'S NL A1B 3V4
Linha de Endereço Formatada 3	CANADA

No exemplo, a organização pode decidir a prioridade a ser aplicada ao Nome de Contato 1 ou ao Nome de Contato 2. A transformação do Validador de Endereço não prioriza os dados de contato.

Se você formatar endereços para a saída de impressão, poderá usar várias instâncias das portas Linha de Endereço Formatada. É possível selecionar até 12 portas de Linha de Endereço Formatada.

## Projetos de Validação de Endereço

Você pode usar a transformação do Validador de Endereço em vários tipos de projeto. Crie um modelo de endereço com portas diferentes para cada tipo de projeto.

Você pode definir um projeto de validação de endereço com um ou mais dos seguintes objetivos:

### **Criar endereços formatados em conformidade com os padrões do serviço de correio**

É possível preparar um grande conjunto de registros de endereço para uma campanha de mala direta. Se você criar os endereços no formato preferido do serviço de correio, os custos serão significativamente reduzidos. Ao preparar endereços para correspondência, selecione portas de saída que gravam cada linha do endereço formatado em uma única porta. Você pode selecionar uma porta diferente para o nome de contato, para as linhas de endereço de rua, e para a localidade e linhas de CEP.

### **Organizar endereços por indicadores de renda e estilo de vida**

Você pode adicionar dados de segmentação de cliente aos registros de endereço residencial. Os dados de segmentação de cliente indicam as preferências de nível de renda e estilo de vida provável dos residentes de cada endereço. Selecione portas do grupo de saída CAMEO para adicionar dados de segmentação de clientes aos registros de endereço. Você pode usar dados de segmentação de clientes em campanhas de mala direta direcionadas a vários mercados consumidores.

### **Criar endereços certificados pelo serviço de correio**

Quando você preparar um conjunto de registro para o Correio da Austrália, Correio do Canadá ou Serviço Postal dos Estados Unidos (USPS), poderá adicionar dados que confirmam a capacidade de entrega de cada endereço.

A transformação do Validador de Endereço pode gerar relatórios para verificar se os registros de endereço estão completos e corretos conforme os padrões de dados de cada serviço de correio.

### **Criar endereços que atendem aos requisitos regulatórios**

Você pode verificar se os registros de endereço mantidos pela sua organização estão corretos conforme as normas da indústria e do governo. Selecione as portas de saída que gravam cada elemento de dados de endereço em um campo separado. Além disso, selecione as portas de status de validação de endereço que fornecem informações detalhadas sobre a precisão e à totalidade dos dados de saída.

### **Completar endereços parciais**

Você pode inserir um endereço parcial e recuperar o endereço completo válido que corresponde ao endereço parcial nos dados de referência. Para completar endereços parciais, configure a transformação para ser executada no modo de lista de sugestões ou no modo interativo. Você pode inserir o endereço de entrada como uma única linha na porta Endereço Completo.

### **Melhorar a qualidade de dados de endereços**

Paralelamente a outros projetos de dados, você pode melhorar a estrutura e a qualidade geral dos dados do conjunto de dados de endereço. Por exemplo, o conjunto de dados pode conter mais colunas do que você precisa, ou pode conter o mesmo tipo de dados em diversas colunas. Você pode reduzir o número de colunas do conjunto de dados e pode simplificar as colunas usadas para diferentes tipos de dados.

## **Endereços Formatados e Padrões do Serviço de Correio**

Ao preparar registros de endereço para uma campanha de mala direta, crie uma estrutura de endereço imprimível que corresponda aos padrões de formatação do serviço de correio.

Por exemplo, o USPS mantém o seguinte formato de endereço para endereços domésticos nos Estados Unidos:

Line 1

Person/Contact Data

JOHN DOE

Line 2	Street Number, Street, Sub-Building	123 MAIN ST NW STE 12
Line 3	Locality, State, ZIP Code	ANYTOWN NY 12345

É possível definir um formato de endereço imprimível que grava cada linha de endereço em uma única porta. Você pode usar portas que reconhecem os tipos de dados em cada linha, ou você pode usar portas que preenchem a estrutura do endereço, independentemente dos dados de cada linha.

A tabela a seguir mostra as maneiras diferentes que você pode formatar um endereço dos Estados Unidos para impressão:

Para Este Endereço	Use Estas Portas	Ou Use Estas Portas
JOHN DOE	Linha de Destinatário 1	Linha de Endereço Formatada 1
123 MAIN ST NW STE 12	Linha de Endereço de Entrega 1	Linha de Endereço Formatada 2
ANYTOWN NY 12345	Última Linha Específica do País 1	Linha de Endereço Formatada 3

Use as portas de linha de Endereço Formatada quando o conjunto de dados contiver tipos diferentes de endereço, como endereços comerciais e residenciais. Um endereço comercial pode precisar de três linhas de endereço para dados de contato e organização. A transformação do Validador de Endereço garante que cada endereço comercial ou residencial seja formatado corretamente usando as portas de Linha de Endereço Formatada somente quando necessário. No entanto, as portas de Linha de Endereço Formatada não identificam o tipo de dados.

Use as portas de Linha de Destinatário, Linha de Endereço de Entrega e Última Linha Específica do País quando todos os endereços seguem um formato. As portas de Linha de Destinatário, Linha de Endereço de Entrega e Última Linha Específica do País separam os elementos de dados do endereço por tipo de informação e tornam o conjunto de dados fácil de entender.

**Nota:** Você pode selecionar outras portas para processar esse endereço. Esse exemplo se concentra nas portas que formatam os endereços para impressão e entrega.

### Dados demográficos e geográficos

Ao criar um conjunto de registro para uma campanha de mala direta, você pode adicionar vários tipos de dados que não aparecem no endereço. Use esses dados para revisar a distribuição demográfica e geográfica dos itens de correspondência.

Por exemplo, você pode identificar o Distrito Congressional ao qual um endereço dos Estados Unidos pertence. Você também poderá gerar coordenadas de latitude e longitude se o país destino incluir as coordenadas nos dados de referência do sistema de correspondência.

## Conclusão de Endereço Parcial

Quando você usa o modo de lista de sugestões ou o modo interativo, pode inserir um endereço incompleto e recuperar os endereços válidos e completos dos dados de referência.

Selecione o modo de lista de sugestões quando estiver em dúvida quanto a um endereço e desejar exibir uma lista de possíveis endereços válidos. Selecione o modo interativo quando tiver certeza de um endereço

e desejar validar a forma completa. Em cada caso, a transformação do Validador de Endereço pesquisa os dados de referência de endereço e retorna todos os endereços que contêm os dados de entrada.

Considere as seguintes regras e diretrizes quando você configurar a transformação para execução no modo de lista de sugestões ou no modo interativo:

- Você pode definir um endereço de entrada em várias portas ou inserir todos os elementos de endereço na porta de entrada Endereço Completo.
- Quando você configurar a transformação no modo de lista de sugestões, selecione portas do grupo de entrada Discreto. Como alternativa, selecione uma porta Endereço Completo e, opcionalmente, selecione uma porta Nome de País do grupo Várias Linhas.
- O modo de lista de sugestões e o modo interativo podem retornar vários endereços para cada endereço de entrada. A propriedade Contagem Máxima de Resultados especifica um limite superior para o número de endereços retornados. Se o número de endereços correspondentes for maior do que o valor de Contagem Máxima de Resultados, a porta Overflow de Contagem retornará o número de endereços adicionais.
- A Verificação de Endereço da Informatica refere-se ao modo de lista de sugestões como modo de conclusão rápida.

## Portas de Status do Validador de Endereço

A transformação de Validador de Endereço grava informações de status nos elementos de endereço que ele lê e grava em portas de entrada e saída. Use as portas de Informações de Status para exibir as informações de status.

Você pode selecionar as seguintes portas de status:

### **Código de Resolução de Endereço**

Descreve os elementos de endereço não válidos no endereço. Selecione a porta do grupo de portas de Informações de Status no Modelo Básico.

### **Tipo de Endereço**

Indica o tipo de endereço em casos em que o serviço de correio reconhece mais de uma forma de um endereço. Selecione a porta do grupo de portas de Informações de Status no Modelo Básico.

### **Status da Entrada do Elemento**

Descreve os níveis de semelhança entre os elementos de endereço de entrada e os dados de referência. Selecione a porta do grupo de portas de Informações de Status no Modelo Básico.

### **Relevância do Elemento**

Identifica os elementos de endereço que o serviço de correio exige para identificar uma caixa de correio para o endereço. Selecione a porta do grupo de portas de Informações de Status no Modelo Básico.

### **Status do Resultado do Elemento**

Descreve qualquer atualização que a validação de endereço faz para o endereço de entrada. Selecione a porta do grupo de portas de Informações de Status no Modelo Básico.

### **Status do Resultado do Elemento Estendido**

Indica a presença de dados adicionais para um endereço nos dados de referência. A porta pode conter informações detalhadas sobre as atualizações que a validação de endereço faz para um endereço. Selecione a porta do grupo de portas de Informações de Status no Modelo Básico.

### Status da Codificação Geográfica

Descreve o tipo de dados de codificação geográfica que a validação de endereço retorna para um endereço. Selecione a porta no grupo de portas Codificação Geográfica no Modelo Básico.

### Pontuação de Capacidade de Envio por Correio

Indica a probabilidade de que o serviço de correio possa entregar um item de correio para o endereço. Selecione a porta do grupo de portas de Informações de Status no Modelo Básico.

### Código de Correspondência

Descreve os resultados da validação de endereços e operações de correção de endereço em um endereço de entrada. Selecione a porta do grupo de portas de Informações de Status no Modelo Básico.

### Porcentagem do Resultado

Representa o grau de similaridade entre um endereço de entrada e o endereço de saída como um valor de porcentagem. Selecione a porta do grupo de portas de Informações de Status no Modelo Básico.

## Definições de Código de Status do Elemento

As portas de Status de Entrada do Elemento, Relevância do Elemento, Status de Resultado do Elemento e Status do Elemento Estendido fornecem informações de status sobre a validade dos elementos de dados de entrada e saída. Selecione as portas de elemento para revisar os resultados de uma operação de validação de endereço.

Os códigos contêm as seguintes informações:

- Os códigos de Status de Entrada do Elemento representam a qualidade da correspondência encontrada entre os dados do endereço de entrada e os dados de referência.
- Os códigos de Relevância do Elemento identificam os elementos do endereço que são necessários para entrega de endereço no país de destino.
- Os códigos de Status de Resultado do Elemento descrevem qualquer alteração feita nos dados de entrada durante o processamento.
- Os códigos de Status de Resultado do Elemento Estendido indicam que os dados de referência de endereço contêm informações adicionais sobre o elemento do endereço.

Cada porta retorna um código de 20 caracteres em que cada caractere se refere a outro elemento de dados de endereço. Ao ler os códigos de saída nas portas de elemento, é necessário saber o elemento ao qual cada caractere se refere. Os 20 caracteres consistem em 10 pares. Os dois códigos em cada par representam um tipo de informação de endereço. Por exemplo, a primeira posição do código de retorno representa as informações básicas de código postal.

**Nota:** A porta de Código de Resolução de Endereço retorna uma string de 20 caracteres baseada nos mesmos elementos de endereço das portas de Status do Elemento.

A tabela seguinte descreve os elementos de endereço que os valores de cada posição identificam:

Posição	Elemento de Endereço	Descrição	Exemplo de Elemento de Endereço
1	Nível de código postal 0	Informações básicas de código postal, como um CEP de cinco dígitos.	O CEP de cinco dígitos 10118
2	Nível de código postal 1	Informações adicionais de código postal, como os quatro dígitos finais de um Código de CEP+4.	0110, no Código CEP+4 10118-0110

Posição	Elemento de Endereço	Descrição	Exemplo de Elemento de Endereço
3	Nível de localidade 0	Localização primária, como cidade ou município.	Londres, na Inglaterra
4	Nível de localidade 1	Localidade dependente, subúrbio, vilarejo.	Islington, em Londres
5	Nível da província 0	Região primária dentro de um país, como um nome de estado dos Estados Unidos, um nome de província no Canadá ou um cantão da Suíça.	Estado de Nova York
6	Nível da província 1	Nome de condado nos Estados Unidos.	Condado de Queens no Estado de Nova York
7	Nível de rua 0	Informação de rua primária.	Rua South Great George
8	Nível de rua 1	Informações de rua dependente.	George's Arcade, na Rua South Great George
9	Nível de número 0	Número de edifício ou casa relacionado à chave primária.	460, na Rua South Great George
10	Nível de número 1	Número de edifício ou casa relacionado à rua dependente.	81, no George's Arcade
11	Nível de serviço de entrega 0	Descritor e número de case postale ou caixa postal.	Caixa Postal 111
12	Nível de serviço de entrega 1	Código do correio responsável pela entrega.	MAIN STN
13	Nível de edifício 0	Nome ou número de edifício. Não identifica um número de casa.	Alice Tully Hall
14	Nível de edifício 1	Nome ou número de edifício adicional.	Teatro Starr, em Alice Tully Hall
15	Nome de subedifício 0	Nome ou número de apartamento, suíte ou andar.	80, em 350 5th Avenue, Andar 80
16	Nome de subedifício 1	Informações de apartamento, suíte ou andar, quando emparelhadas com informações de nível de Subedifício 0	80-18, onde 18 é o número da suíte e 80 é o número do andar
17	Nível da organização 0	Nome da empresa.	AddressDoctor® GmbH
18	Nível da organização 1	Informações adicionais corporativas, como uma empresa principal.	Informatica Corporation



Posição	Elemento de Endereço	Descrição	Exemplo de Elemento de Endereço
19	Nível de país 0	Nome do país.	Estados Unidos da América
20	Nível de país 1	Território.	Ilhas Virgens dos Estados Unidos

Quando um nome de porta tem um sufixo de número, o nível 0 refere-se aos dados na porta número 1 e o nível 1 refere-se aos dados nas portas de número 2 a 6.

As informações do nível 0 podem preceder ou vir depois das informações do nível 1 em um endereço de impresso. Por exemplo, os códigos Postais de nível 1 aparecem depois dos códigos Postais de nível 0 e a Localidade de nível 1 precede a Localidade de nível 0.

## Valores da Porta de Saída do Código de Resolução de Endereço

O Código de Resolução de Endereço é uma string de 20 caracteres em que cada caractere na string representa um elemento de endereço de entrada diferente. O valor de um caractere descreve qualquer elemento de endereço inválido na posição correspondente no endereço.

A tabela a seguir descreve os valores da porta de Código de Resolução de Endereço:

Código	Descrição
2	O elemento de endereço é necessário para entrega, mas não está presente no endereço de entrada. Os dados de referência de endereço não contêm o elemento de endereço ausente. Uma saída de 2 indica que o endereço não é válido para entrega sem o elemento de endereço.
3	O elemento de endereço é um número de casa ou rua que está fora do intervalo válido para o endereço. Por exemplo, o elemento de endereço contém um número de casa que não existe na rua especificada. O modo de lista de sugestões retorna endereços alternativos.
4	A validação de endereço não pode verificar ou corrigir o elemento de endereço porque o endereço de entrada contém mais de uma instância do elemento.
5	O elemento de endereço é ambíguo no endereço atual, e os dados de referência de endereço contêm alternativas. A validação de endereço copia o elemento de entrada para o endereço de saída. Por exemplo, o elemento de endereço é um código postal válido que não corresponde a uma localidade válida no endereço.
6	O elemento de endereço contradiz outro elemento no endereço. A validação de endereço não consegue determinar o elemento correto do endereço. O endereço de saída copia o endereço de entrada.
7	O elemento de endereço não pode ser corrigido antes que várias alterações sejam feitas no endereço. A validação de endereço pode corrigir o endereço, mas o número de alterações indica que o endereço não é confiável.
8	Os dados não estão em conformidade com as regras de validação do serviço de correio.

## Valores de Porta de Saída do Status da Entrada do Elemento

O Status de Entrada do Elemento é uma string de 20 caracteres em que cada caractere representa um elemento de endereço de entrada diferente. O valor de cada caractere representa o tipo de processamento executado no elemento de endereço.

Localize a porta no grupo de portas de Informações de Status.

A seguinte tabela descreve os códigos que o Status da Entrada do Elemento pode retornar em cada posição na string de saída no modo de lote, certificado ou lista de sugestões:

Código	Descrição
0	O endereço de entrada não contém dados na posição atual.
1	Os dados de referência não contém os dados na posição atual.
2	Não é possível verificar os dados porque os dados de referência estão ausentes.
3	Os dados na posição atual estão incorretos. O banco de dados de referência sugere que o número ou o valor do serviço de entrega está fora do intervalo esperado pelos dados de referência. Nos modos em lote e certificado, a transformação transmite os dados de entrada na posição atual de forma incorreta como saída.
4	Os dados na posição atual correspondem aos dados de referência, mas contém erros.
5	Os dados na posição atual correspondem aos dados de referência, mas a transformação corrigiu ou padronizou os dados.
6	Os dados na posição atual correspondem aos dados de referência e não contém erros.

A seguinte tabela descreve os códigos que o Status da Entrada do Elemento pode retornar em cada posição na string de saída no modo de análise:

Código	Descrição
0	O endereço de entrada não contém dados na posição atual.
1	A transformação moveu o elemento na posição atual para outra posição no endereço de saída.
2	O elemento na posição atual é correspondente ao valor de dados de referência, mas a transformação normalizou o elemento no endereço de saída.
3	Os dados na posição atual estão corretos.

## Valores de Porta de Saída da Relevância do Elemento

Os valores de Relevância do Elemento indicam se um elemento de endereço é necessário para entrega postal. Localize a porta no grupo de portas de Informações de Status.

O valor de Relevância do Elemento é uma cadeia de 20 caracteres na qual cada caractere representa um tipo diferente de dados de endereço. Depois de executar o mapeamento de validação de endereço, analise a saída da porta para identificar os elementos de endereço necessários para cada endereço. Use os resultados

para verificar se você selecionou as portas de saída certas para os dados de endereço. Se você não selecionar uma porta de saída para um elemento de dados de endereço relevante, a saída para o endereço não será válida.

A seguinte tabela descreve os códigos que a Relevância do Elemento pode retornar em cada posição da cadeia de saída:

Código	Descrição
0	Não é relevante para a entrega no endereço.
1	Relevante para a entrega no endereço. O serviço de correio nacional não pode entregar no endereço sem os dados nessa posição da cadeia de saída.

**Nota:** Os valores de Relevância de Elemento estão disponíveis para endereços com um valor de Código de Correspondência de Cx ou Vx no modo do lote ou Cx, Vx, I3 ou I4 no modo interativo. Outros códigos de avaliação, como Status de Entrada do Elemento, Status de Resultado do Elemento, Status de Resultado do Elemento Estendido e Código de Resolução de Endereço retornam valores, independentemente do valor de Código de Correspondência.

## Valores de Porta de Saída do Status do Resultado do Elemento

O Status de Resultado do Elemento é uma string de 20 caracteres em que cada caractere representa um elemento de endereço de entrada diferente. O valor para cada caractere descreve qualquer atualização que o processo de validação faz para o elemento de endereço.

Localize a porta no grupo de portas de Informações de Status.

A tabela a seguir descreve os valores da porta Status de Resultado do Elemento:

Código	Descrição
0	O endereço de saída não contém dados na posição atual.
1	A transformação não pode localizar os dados da posição atual nos dados de referência. A transformação copia os dados de entrada para os dados de saída.
2	Os dados na posição atual não são verificados, mas são padronizados.
3	Os dados na posição atual são verificados, mas não correspondem aos dados de referência. Os dados de referência sugerem que os dados de número não estão no intervalo válido. A transformação copia os dados de entrada para a porta de saída. É aplicável no modo em lote.
4	A transformação copia os dados de entrada para os dados de saída porque os dados de referência estão ausentes.
5	Os dados na posição atual foram validados, mas não foram alterados porque há várias correspondências nos dados de referência. É aplicável no modo em lote.
6	A validação de dados excluiu o valor de entrada na posição atual.

Código	Descrição
7	Os dados na posição atual foram validados, mas os dados de entrada continham um erro de ortografia. A validação corrigiu o erro com um valor dos dados de referência.
8	Os dados na posição atual foram validados e atualizados com um valor dos dados de referência. Um valor de 8 também pode significar que o banco de dados de referência contém dados adicionais para o elemento de entrada. Por exemplo, a validação poderá adicionar um número de edifício ou de subedifício se encontrar uma correspondência perfeita para o nome da rua ou do edifício.
9	Os dados na posição atual foram validados, mas não foram alterados e o status de entrega não está claro. Por exemplo, o valor de DPV está errado.
C	Os dados na posição atual foram validados e verificados, mas os dados de nome estão desatualizados. A validação alterou os dados de nome.
Dia	Os dados na posição atual foram validados e verificados, mas foram alterados de um exônimo para um nome oficial.
E	Os dados na posição atual foram validados e verificados. No entanto, a validação de endereço padronizou o idioma ou a distinção entre maiúsculas e minúsculas dos caracteres. A validação de endereço poderá alterar o idioma se o valor corresponder totalmente a um idioma alternativo. Por exemplo, a validação de endereço pode alterar "Brussels" para "Bruxelles" em um endereço da Bélgica.
F	Os dados na posição atual foram validados e verificados, mas não foram alterados devido a uma correspondência perfeita com os dados de referência.

As posições 19 e 20 na string de saída relacionam-se aos dados do país.

A seguinte tabela descreve os valores que a validação pode retornar para as posições 19 e 20:

Código	Descrição
0	O endereço de saída não contém dados na posição atual.
1	A validação de endereço não reconhece os dados do país.
4	A validação de endereço identifica o país no valor de País Padrão na transformação do Validador de Endereço.
5	A validação de endereço não pode determinar o país porque os dados de referência contém várias correspondências.
6	A validação de endereço identifica o país em um script.
7	A validação de endereço identifica o país no formato de endereço.
8	A validação de endereço identifica o país nos dados da cidade principal.
9	A validação de endereço identifica o país nos dados da província.
C	A validação de endereço identifica o país nos dados do território.

Código	Descrição
Dia	A validação de endereço identifica o país no nome de país, mas o nome contém erros.
E	A validação de endereço identifica o país nos dados de endereço, por exemplo com um código ISO ou um nome de país.
F	A validação de endereço identifica o país no valor Forçar País definido na transformação do Validador de Endereço.

## Valores de Porta de Saída do Status do Resultado do Elemento Estendido

O Status do Resultado do Elemento Estendido é uma cadeia de 20 caracteres em que cada caractere representa um elemento de endereço de entrada diferente. Os códigos de saída da porta complementam os dados de status na porta de Status de Entrada do Elemento e de Status de Resultado do Elemento. Essa porta de saída também pode indicar a presença de informações adicionais sobre um elemento de endereço nos dados de referência.

Localize a porta no grupo de portas de Informações de Status.

A seguinte tabela descreve os valores de porta do Status do Resultado do Elemento Estendido:

Código	Descrição
1	Os dados de referência de endereço contêm informações adicionais sobre o elemento de endereço. A validação de endereço não requer essas informações adicionais.
2	A validação de endereço atualizou o elemento de endereço para resolver um erro de dados ou de formato. A validação de endereço não verificou o elemento de endereço.
3	A validação de endereço atualizou o elemento de endereço para resolver um erro de dados ou de formato. A validação de endereço verificou os dados numéricos no elemento de endereço.
4	A validação de endereço moveu o elemento de endereço para outro campo de forma a resolver um erro de formato.
5	Os dados de referência de endereço contêm uma versão alternativa do elemento de endereço, como um nome de localidade preferencial.
6	A validação de endereço não verificou todas as partes do elemento de endereço. O elemento inclui dados que a validação de endereço não pode validar.
7	A validação de endereço encontrou um elemento de endereço válido na posição errada em um endereço. A validação de endereço moveu o elemento de endereço para a posição correta.
8	A validação de endereço encontrou um elemento de endereço válido no campo de dados errado. A validação de endereço moveu o elemento de endereço para o campo correto.
9	A validação de endereço gerou o elemento de saída de acordo com as regras de validação do serviço de correio.

Código	Descrição
A	A validação de endereço encontrou elementos de endereço de diferentes tipos de endereço que são elegíveis para a posição atual. A validação de endereço selecionou o elemento de endereço de saída que está em conformidade com as regras da operadora de correio no país de destino.
B	A validação de endereço não pode determinar a relevância do elemento. A validação de endereço retorna o valor padrão para o país especificado pelo endereço.
C	Modo de lista de sugestões. A validação de endereço pode retornar sugestões de endereços adicionais para o elemento de endereço. Para retornar sugestões adicionais, atualize a propriedade Contagem Máxima de Resultados para a transformação do Validador de Endereço.
Dia	A validação de endereço interpolou os dados numéricos no elemento de endereço.
E	A validação de endereço não pode retornar o elemento de endereço no idioma preferencial. A validação de endereço retorna o elemento no idioma padrão.
S	Modo de pesquisa de código de endereço. O endereço de entrada está desatualizado.

## Valores de Porta de Saída da Pontuação de Capacidade de Envio por Correio

O valor da Pontuação de Capacidade de Envio por Correio inclui uma estimativa da capacidade de entrega do endereço de saída. Use a pontuação de capacidade de envio por correio como um indicador geral da capacidade de entrega do endereço. Localize a porta no grupo de portas de Informações de Status.

A transformação de Validador de Endereço considera vários fatores ao calcular a pontuação de capacidade de envio por correio. A transformação baseia os cálculos principalmente no valor do Código de Correspondência e no valor do Status do Resultado do Elemento para o endereço. Outros fatores que influenciam a pontuação de capacidade de envio por correio incluem a relevância postal dos valores de endereço e a granularidade dos dados de referência do país.

O valor da porta de Pontuação de Capacidade de Envio por Correio fornece uma estimativa da capacidade de entrega do endereço. A pontuação não é um indicador preciso ou definitivo da capacidade de entrega do endereço.

A seguinte tabela descreve os códigos de saída da Pontuação de Capacidade de Envio por Correio:

Valor	Resumo	Descrição
5	Totalmente confiante	Indica que a validação de endereço controlou e verificou todos os elementos relevantes do endereço de entrada.
4	Quase certeza	Indica um dos seguintes cenários: <ul style="list-style-type: none"> <li>- Um ou mais elementos de endereço relevantes não podem ser verificados devido à ausência de dados de referência. Outros elementos de endereço são verificados.</li> <li>- A validação de endereço corrigiu um ou mais elementos relevantes com um alto grau de confiança. Isso ocorre quando a validação de endereço encontra uma única correspondência entre o endereço de entrada e os dados de referência, e o grau de variação é muito baixo.</li> </ul>

Valor	Resumo	Descrição
3	Deve estar bem	Indica que a validação de endereço corrigiu um ou mais elementos relevantes no endereço de entrada. A validação de endereço encontrou uma única correspondência entre o endereço de entrada e os dados de referência, e o grau de variação é aceitável.
2	Boa possibilidade	Indica que a validação de endereço não pode corrigir ou verificar o endereço por um dos seguintes motivos: <ul style="list-style-type: none"> <li>- A validação de endereço não pode identificar uma correspondência de candidato nos dados de referência com confiança suficiente.</li> <li>- A validação de endereço encontrou várias correspondências de candidatos com níveis de confiança semelhantes.</li> </ul> O serviço de correio pode ser capaz de fazer a entrega para o endereço.
1	Com risco	Indica que a validação de endereço pode encontrar correspondências parciais de dados de referência apenas para o endereço de entrada.
0	Entrega impossível	Indica que a validação de endereço não pode encontrar uma correspondência para o endereço nos dados de referência. Estão faltando muitos elementos no endereço de entrada, ou a validação de endereço não pode verificar a maioria dos elementos no endereço.

## Valores de Porta de Saída do Código de Correspondência

O valor do Código de Correspondência resume os resultados da comparação do endereço de entrada com os dados de referência. O código também resume qualquer correção que a transformação fez no endereço. Localize a porta no grupo de portas de Informações de Status.

A tabela a seguir descreve os valores de porta de saída do Código de Correspondência:

Código	Descrição
A1	A pesquisa de código de endereço encontrou um endereço parcial ou um endereço completo para o código de entrada.
A0	A pesquisa de código de endereço não encontrou endereços para o código de entrada.
C4	Corrigido. Todos os elementos postais relevantes são verificados.
C3	Corrigido. Alguns elementos não podem ser verificados.
C2	Corrigido, mas o status de entrega não está claro devido à ausência de dados de referência.
C1	Corrigido, mas o status de entrega não está claro porque a padronização do usuário introduziu erros.
I4	Os dados não puderam ser corrigidos completamente, mas há uma única correspondência com um endereço nos dados de referência.
I3	Os dados não puderam ser corrigidos completamente, e há várias correspondências com endereços nos dados de referência.

Código	Descrição
I2	Os dados não podem ser corrigidos. O modo em lote retorna endereços sugeridos parcialmente.
I1	Os dados não podem ser corrigidos. O modo em lote não pode sugerir um endereço.
N7	Erro de validação. A validação não ocorreu porque a validação de uma única linha não está desbloqueada.
N6	Erro de validação. A validação não ocorreu porque a validação de uma única linha não é compatível com o país de destino.
N5	Erro de validação. A validação não ocorreu porque o banco de dados de referência está desatualizado.
N4	Erro de validação. A validação não ocorreu porque o banco de dados de referência está corrompido ou mal formatado.
N3	Erro de validação. A validação não ocorreu porque não foi possível desbloquear os dados do país.
N2	Erro de validação. A validação não ocorreu porque o banco de dados de referência exigido não está disponível.
N1	Erro de validação. A validação não ocorreu porque o país não foi reconhecido ou não é compatível.
Q3	Modo de Lista de Sugestões. A validação de endereço pode recuperar um ou mais endereços completos dos dados de referência de endereço que correspondem ao endereço de entrada.
Q2	Modo de Lista de Sugestões. A validação de endereço pode combinar os elementos do endereço de entrada e os elementos dos dados de referência de endereço para criar um endereço completo.
Q1	Modo de Lista de Sugestões. A validação de endereço não pode sugerir um endereço completo. Para gerar uma sugestão de endereço completo, adicione os dados ao endereço de entrada.
Q0	Modo de Lista de Sugestões. Não há dados de entrada suficientes para gerar uma sugestão.
RB	País reconhecido com base na abreviação. Reconhece os códigos de país ISO de dois e três caracteres. Também pode reconhecer abreviações comuns, como "GER" para Alemanha.
RA	País reconhecido com base na configuração Forçar País da transformação.
R9	País reconhecido com base na configuração País Padrão da transformação.
R8	País reconhecido pelo nome de país.
R7	País reconhecido pelo nome de país, mas a transformação identificou erros nos dados do país.
R6	País reconhecido pelos dados de território.



Código	Descrição
R5	País reconhecido pelos dados de província.
R4	País reconhecido pelos dados da cidade principal.
R3	País reconhecido pelo formato de endereço.
R2	País reconhecido por um script.
R1	País não reconhecido porque várias correspondências estão disponíveis.
R0	País não reconhecido.
S4	Modo de análise. O endereço foi analisado perfeitamente.
S3	Modo de análise. O endereço foi analisado com vários resultados.
S1	Modo de análise. Ocorreu um erro de análise devido a um formato de entrada incompatível.
V4	Verificado. Os dados de entrada estão corretos. A validação de endereço verificou todos os elementos postais relevantes. As entradas são perfeitamente correspondentes.
V3	Verificado. Os dados de entrada estão corretos, mas alguns ou todos os elementos foram padronizados, ou a entrada contém nomes ou exônimos desatualizados.
V2	Verificado. Os dados de entrada estão corretos, mas alguns elementos não podem ser verificados devido a dados de referência incompletos.
V1	Verificado. Os dados de entrada estão corretos, mas a padronização do usuário afetou negativamente a capacidade de entrega. Por exemplo, o tamanho do código postal é muito pequeno.

## Valores de Porta de Saída do Status da Codificação Geográfica

A tabela a seguir descreve os valores de porta de saída do Status da Codificação Geográfica. Localizar essa porta no grupo de portas de Codificação Geográfica.

Selecione essa porta se você tiver instalado dados de referência de codificação geográfica para o país de um endereço de entrada.

Valor	Descrição
EGC0	Não é possível anexar codificações geográficas ao endereço de entrada porque as codificações geográficas não estão disponíveis para o endereço.
EGC1-3	Reservado para uso futuro.
EGC4	As codificações geográficas são parcialmente precisas no nível de CEP.
EGC5	As codificações geográficas são precisas no nível de CEP.
EGC6	As codificações geográficas são precisas no nível de localidade.

Valor	Descrição
EGC7	As codificações geográficas são precisas no nível de rua.
EGC8	As codificações geográficas são precisas no nível de número da casa. As codificações geográficas estimam a localização do número da casa e incluem um deslocamento para o lado da rua que contém a caixa de correio.
EGC9	As codificações geográficas são precisas em relação ao ponto de chegada ou ao endereço específico.
EGCA	As codificações geográficas são precisas em relação ao centro do setor do território.
EGCC	O banco de dados de codificação geográfica está corrompido.
EGCN	Não é possível localizar o banco de dados de codificação geográfica .
EGCU	O banco de dados de codificação geográfica não está desbloqueado.

**Nota:** A Informatica não emite mais dados de referência para centroide de setor e codificações geográficas residenciais.

## Transformação de Validador de Endereço - Configurações Gerais

Defina as configurações gerais para configurar os parâmetros necessários para a validação de endereço.

Você pode configurar as seguintes propriedades na exibição **Configurações Gerais**:

### País Padrão

Especifica o conjunto de dados de referência de endereço que a transformação usará se não puder identificar um país de destino no endereço de entrada. Selecione Nenhum se os seus dados incluem informações de país.

Você também pode definir o país padrão como uma propriedade avançada na transformação.

### Forçar País

Propriedade opcional. Substitui o nome ou a abreviação do país no endereço de entrada pelo nome ou pela abreviação do país padrão. Se o endereço de entrada não identificar um país, a transformação anexará os dados do país padrão ao endereço.

### Separador de Linha

Especifica o símbolo delimitador que separa os campos de dados em um endereço de uma única linha.

Você também pode especificar um separador de linha como uma propriedade avançada na transformação.

### Estilo de Formatação de Maiúsculas e Minúsculas

Define o estilo de formatação de maiúsculas e minúsculas dos caracteres para dados de saída. Selecione a opção Misto para seguir o padrão de dados de referência de endereço para as letras maiúsculas iniciais. Selecione a opção Preservado para gravar o endereço no estilo de formatação de maiúsculas e minúsculas usado pelos dados de referência de endereço.

Você também pode definir o estilo de letras maiúsculas e minúsculas como uma propriedade avançada na transformação.

#### Modo

Determina o tipo de validação que a transformação executa.

Selecione uma das seguintes opções:

Tipo de Modo	Descrição
Pesquisa de código de endereço	Retorna um endereço parcial ou completo dos dados de referência quando você fornece um código de endereço como uma entrada. Vários países suportam códigos de endereço que representam a localidade, a rua, o edifício ou a caixa de correio exclusiva de um endereço.
Lote	Executa a validação de endereço nos registros em um conjunto de dados. A validação em lote concentra-se na totalidade do endereço e na capacidade de entrega. O modo em lote não retorna sugestões para endereços de baixa qualidade. Lote é o modo padrão.
Certificado	Executa a validação de endereço nos registros em um conjunto de dados para os padrões de certificação do país especificado. Os padrões de certificação exigem que cada endereço identifique uma caixa de correio exclusiva. Você pode realizar a validação de endereço certificado em endereços da Austrália, França, Nova Zelândia, Reino Unido e Estados Unidos.
Reconhecimento de país	Determina um país de destino do endereço postal. A transformação não realiza a validação de endereço no modo de reconhecimento de país.
Interativo	Completa um endereço válido incompleto. Quando um endereço de entrada incompleto corresponde a mais de um endereço nos dados de referência, a transformação retorna todos os endereços válidos até o limite que a Contagem Máxima de Resultados especifica.
Analisar	Analisa dados em campos de endereço. A transformação não realiza a validação de endereço no modo de análise.
Lista de sugestões	Retorna uma lista de endereços válidos dos dados de referência quando um endereço de entrada contém informações fragmentadas. Quando um fragmento do endereço corresponde a mais de um endereço nos dados de referência, a transformação retorna todos os endereços válidos até o limite que a Contagem Máxima de Resultados especifica.

Você também pode definir o modo como uma propriedade avançada na transformação.

## Propriedades de validação de endereço na janela Preferências

Você pode exibir as propriedades do mecanismo de validação de endereço e os arquivos de dados de referência de endereço que o mecanismo lê na Developer tool. A Developer tool expõe as propriedades do mecanismo que o Serviço de Integração de Dados usa para executar mapeamentos de validação de

endereço. A Developer tool lista as propriedades do Serviço do Gerenciamento de Conteúdo que rege as operações de validação de endereço.

Use a janela **Preferências** na Developer tool para revisar as propriedades. Selecione a opção **Status de Conteúdo** na janela **Preferências** para identificar o Serviço do Gerenciamento de Conteúdo usado pelo Serviço de Integração de Dados atual. Para exibir as propriedades, selecione o Serviço do Gerenciamento de Conteúdo local.

Você pode exibir as seguintes propriedades:

#### Dados de Validação de Endereço

As propriedades de dados de validação de endereço listam os tipos de dados de referência que o Serviço do Gerenciamento de Conteúdo atual pode fornecer ao Serviço de Integração de Dados. As propriedades também indicam os países aos quais os dados de referência se aplicam.

#### Mecanismo de Validação de Endereço

As propriedades do mecanismo de validação de endereço incluem a versão atual do mecanismo, o mecanismo no qual os componentes de certificação foram atualizados mais recentemente e o método de pré-carregamento de dados.

#### Licença de Validação de Endereço

As propriedades de licença de validação de endereço incluem informações de licença para os dados de referência que o Serviço do Gerenciamento de Conteúdo atual pode fornecer ao Serviço de Integração de Dados.

## Propriedades de dados de validação de endereço

As propriedades de dados de validação de endereço listam os tipos de dados de referência que o Serviço do Gerenciamento de Conteúdo atual pode fornecer ao Serviço de Integração de Dados. As propriedades também incluem os países aos quais os dados de referência se aplicam.

A tabela a seguir descreve as propriedades de dados exibidas quando você seleciona o Serviço do Gerenciamento de Conteúdo na exibição **Status de Conteúdo**:

Propriedade	Descrição
ISO de País	O país ao qual se aplica o arquivo de dados de referência de endereço. A propriedade mostra o código ISO de três caracteres para o país.
Data de Expiração	A data de expiração do arquivo atual. A Informatica lança um arquivo mais recente na data de expiração. Você pode usar o arquivo de dados de referência de endereço atual após a data de expiração, mas os dados no arquivo podem não ser mais precisos.
Tipo de País	O tipo de validação de endereço que você pode realizar com os dados. Selecione o tipo de processamento na opção <b>Modo</b> na guia <b>Configurações Gerais</b> . Se o modo que você selecionar não corresponder a um arquivo de dados de endereço no domínio, o mapeamento de validação de endereço falhará.
Data de Expiração de Desbloqueio	A data de expiração da licença. Não será possível usar qualquer versão do arquivo após a data de expiração de desbloqueio. As propriedades Data de Expiração de Desbloqueio e Data de Validade na exibição Propriedades de Licença de Validação de Endereço representam as mesmas informações.
Data de Início de Desbloqueio	A data em que a licença entra em vigor para o modo que a propriedade Tipo de País identifica e o país que a propriedade ISO de País identifica. Você não pode usar qualquer versão do arquivo antes da data de início de desbloqueio.

## Propriedades de licença de validação de endereço

As propriedades de licença de validação de endereço incluem informações de licença para os dados de referência que o Serviço do Gerenciamento de Conteúdo atual pode fornecer ao Serviço de Integração de Dados.

A tabela a seguir descreve as propriedades de licença que são exibidas quando você seleciona o Serviço do Gerenciamento de Conteúdo na exibição **Status de Conteúdo**:

Propriedade	Descrição
Código de Desbloqueio	O código de licença que desbloqueia os dados de referência para o modo que a propriedade Tipo de Código identifica. A Developer tool exibe os quatro primeiros caracteres do código e mascara os outros caracteres.
Tipo de Código	O modo de validação de endereço que você pode executar com os dados que a licença especifica. A Informatica emite um único código de licença para cada modo. O código de licença pode ser aplicado a um ou mais países.  Selecione o tipo de processamento na opção <b>Modo</b> na guia <b>Configurações Gerais</b> . Se o modo que você selecionar não corresponder a um arquivo de dados de endereço no domínio, o mapeamento de validação de endereço falhará.
Lista de Países	Os países para os quais o código de desbloqueio desbloqueia os dados de referência. A propriedade Lista de Países contém um ou mais códigos ISO de três caracteres para cada país.
Status	O status do código de licença. A propriedade retorna OK quando o arquivo de licença é válido.
Data de Validade	A data de expiração da licença.  As propriedades Data de Validade e Data de Expiração de Desbloqueio na exibição Propriedades de Dados de Validação de Endereço representam as mesmas informações.

## Propriedades do mecanismo de validação de endereço

As propriedades do mecanismo de validação de endereço incluem a versão atual do mecanismo, o mecanismo no qual os componentes de certificação foram atualizados mais recentemente e o método de pré-carregamento de dados.

A tabela a seguir descreve as propriedades do mecanismo que são exibidas quando você seleciona o Serviço do Gerenciamento de Conteúdo na exibição **Status de Conteúdo**:

Propriedade	Valor
Versão do Mecanismo	A versão do mecanismo de validação de endereço que o Serviço de Integração de Dados executa.
Versão do CASS	A versão do mecanismo de validação de endereço na qual a Informatica atualizou mais recentemente os componentes de certificação CASS. Use a propriedade para identificar a versão do mecanismo em um relatório de certificação CASS.  A propriedade também inclui o ciclo de certificação CASS compatível com o mecanismo. Por exemplo, o mecanismo pode ser compatível com o ciclo de certificação N.
Versão do AMAS	A versão do mecanismo de validação de endereço na qual a Informatica atualizou mais recentemente os componentes de certificação AMAS. Use a propriedade para identificar a versão do mecanismo em um relatório de certificação AMAS.

Propriedade	Valor
Versão do SendRight	A versão do mecanismo de validação de endereço na qual a Informatica atualizou mais recentemente os componentes de certificação SendRight. Use a propriedade para identificar a versão do mecanismo em um relatório de certificação SendRight.
Versão do SERP	A versão do mecanismo de validação de endereço na qual a Informatica atualizou mais recentemente os componentes de certificação SERP. Use a propriedade para identificar a versão do mecanismo em um relatório de certificação SERP.
Versão do SNA	A versão do mecanismo de validação de endereço na qual a Informatica atualizou mais recentemente os componentes de certificação SNA. Use a propriedade para identificar a versão do mecanismo em um relatório de certificação SNA.
Método de Pré-Carregamento	O método que o Serviço de Integração de Dados usa para carregar previamente o banco de dados de referência na memória. As propriedades do Serviço do Gerenciamento de Conteúdo especificam os países para os quais o Serviço de Integração de Dados carrega previamente os dados de referência. Os valores possíveis são MAP e LOAD. O valor padrão é MAP. Os métodos MAP e LOAD alocam um bloco de memória e, na sequência, leem os dados de referência no bloco. No entanto, o método MAP pode compartilhar dados de referência entre vários processos.
Tamanho do Cache	O tamanho do cache de dados que o Serviço de Integração de Dados usa para os dados de referência que o serviço não carrega previamente. Os valores possíveis são NONE, SMALL e LARGE. O valor padrão é LARGE.
Uso Máximo de Memória	O número de megabytes de memória que o mecanismo de validação de endereço pode alocar. O valor padrão é 4096.
Contagem Máxima de Objetos do Endereço	O número máximo de instâncias de validação de endereço que o Serviço de Integração de Dados pode executar ao mesmo tempo. O valor padrão é 3.
Contagem Máxima de Segmentos	O número máximo de segmentos que a validação de endereço pode usar. O valor padrão é 2.
Contagem Máxima de Resultados	O número máximo de endereços que a validação de endereço pode retornar quando você executa um mapeamento no modo de lista de sugestões. O valor padrão é 20. O limite máximo na propriedade é 100.
Data Atual	A data atual. A Developer tool retorna os valores de propriedade que se aplicam na data atual.
Gravar BOM XML	Indica se o Serviço de Integração de Dados grava uma marca de ordem de byte no arquivo GetConfig.xml. Os valores possíveis são ALWAYS, IF_NECESSARY e NEVER. O valor padrão é IF_NECESSARY.
Codificação XML	Identifica a codificação XML que o mecanismo de validação de endereço usa para ler e gravar dados.

## Propriedades avançadas de validação de endereço

Configure as propriedades avançadas para determinar como o Serviço de Integração de Dados processa dados para a transformação de Validador de Endereço.

## Alias de Localidade

Determina se a validação de endereço substitui um alias de localidade válido pelo nome oficial da localidade.

Um alias de localidade é um nome de localidade alternativo que o USPS reconhece como um elemento em um endereço que pode receber entregas. Você pode usar a propriedade quando configurar a transformação do Validador de Endereço para validar os registros de endereço dos Estados Unidos no modo Certificado.

A seguinte tabela descreve as opções de Alias de Localidade:

Opção	Descrição
Desativado	Desativa a propriedade Alias de Localidade.
Oficial	Substitui qualquer nome de localidade alternativo ou alias de localidade pelo nome oficial da localidade. Opção padrão.
Preservar	Preserva um nome de localidade alternativo ou um alias de localidade válido. Se o nome da localidade de entrada não for válido, a validação de endereço o substituirá pelo nome oficial.

## Rua de Alias

Determina se a validação de endereço substitui um alias de rua pelo nome oficial da rua.

Um alias de rua é um nome alternativo de rua que o USPS reconhece como um elemento em um endereço que pode receber entregas. Você pode usar a propriedade quando configurar a transformação do Validador de Endereço para validar os registros de endereço dos Estados Unidos no modo Certificado.

A seguinte tabela descreve as opções de Rua de Alias:

Opção	Descrição
Desativado	Não aplica a propriedade.
Oficial	Substitui qualquer nome de rua ou de alias alternativo pelo nome de rua oficial. Opção padrão.
Preservar	Preserva um nome de rua ou alias de rua alternativo. Se o nome de rua de entrada não for válido, a validação de endereço substituirá o nome pelo nome oficial.

## Estilo de Formatação de Maiúsculas e Minúsculas

Especifica o estilo de maiúsculas e minúsculas de caracteres que a transformação aplica aos dados de endereços de saída.

A tabela seguinte descreve as opções de Estilo de Formatação de Maiúsculas e Minúsculas:

Opção	Descrição
Atribuir Parâmetro	Usa um parâmetro que você especifica para definir o estilo de maiúsculas e minúsculas.
Minúsculas	Grava o endereço de saída em letras minúsculas.

Opção	Descrição
Misto	Usa o estilo de maiúsculas e minúsculas em uso no país de destino quando possível.
Preservado	Aplica o estilo de maiúsculas e minúsculas utilizado pelos dados de referência de endereço. Opção padrão.
Sem Alteração	Não aplica um estilo de letras maiúsculas e minúsculas ao endereço. <b>Nota:</b> A opção <b>Sem Alteração</b> não garante que o endereço de saída corresponderá à formatação de maiúsculas e minúsculas do endereço de entrada. Se a validação de endereço substituir um elemento de endereço por um elemento dos dados de referência, o elemento seguirá a formatação de maiúsculas e minúsculas usada pelos dados de referência.
Maiúsculas	Grava o endereço de saída em letras maiúsculas.

Você também pode configurar o estilo de maiúsculas e minúsculas na guia **Configurações Gerais**.

### Utilização de Parâmetros

Você pode usar um dos seguintes parâmetros para especificar o estilo de maiúsculas e minúsculas:

- LOWER. Grava o endereço de saída em letras minúsculas.
- MIXED. Usa o estilo de maiúsculas e minúsculas em uso no país de destino quando possível.
- NATIVE. Aplica o estilo de maiúsculas e minúsculas utilizado pelos dados de referência de endereço. Opção padrão. Corresponde a opção **Preservado**.
- NOCHANGE. Não aplica um estilo de letras maiúsculas e minúsculas ao endereço.
- UPPER. Grava o endereço de saída em letras maiúsculas.

Insira o valor do parâmetro em letras maiúsculas.

## País de Origem

Identifica o país no qual os registros de endereço são enviados.

Selecione um país da lista. A propriedade é vazia por padrão.

## Tipo de País

Determina o formato do nome ou abreviação do país nos dados de saída do Endereço Completo ou da porta de Linha de Endereço Formatada. A transformação grava o nome ou abreviação do país no formato padrão de país selecionado.

A tabela seguinte descreve as opções Tipo de País:

Opção	País
ISO 2	Códigos de país de dois caracteres ISO
ISO 3	Códigos de país de três caracteres ISO
ISO #	Códigos de país de três dígitos ISO



Opção	País
Abreviação	(Reservado para uso futuro)
CN	Canadá
DA	(Reservado para uso futuro)
DE	Alemanha
EN	Grã-Bretanha (padrão)
ES	Espanha
FI	Finlândia
FR	França
GR	Grécia
IT	Itália
JP	Japão
HU	Hungria
KR	Coreia, República da
NL	Holanda
PL	Polônia
PT	Portugal
RU	Rússia
SA	Arábia Saudita
SE	Suécia

## País Padrão

Especifica o conjunto de dados de referência de endereço que a transformação usa quando um registro de endereço não identifica um país de destino.

Selecione um país da lista. Use a opção padrão se os registros de endereço incluírem informações de país. O padrão é Nenhum.

Você também pode configurar o país padrão na guia **Configurações Gerais**.

### Utilização de Parâmetro

Você pode usar um parâmetro para especificar o país padrão. Quando você criar o parâmetro, insira o código ISO 3166-1 alpha-3 para o país como o valor de parâmetro. Quando você inserir um valor de parâmetro, use caracteres em letras maiúsculas. Por exemplo, se todos os registros de endereço incluírem informações de país, insira NONE.

## Prioridade de Endereço Duplo

Determina o tipo de endereço a validar. Defina a propriedade quando os registros de endereço de entrada contiverem mais de um tipo de dados de endereço válido.

Por exemplo, use a propriedade quando um registro de endereço contiver os elementos de caixa postal e os elementos de rua. A validação de endereço lê os elementos de dados que contêm o tipo de dados de endereço que você especificar. A validação de endereço ignora quaisquer dados incompatíveis no endereço.

A tabela a seguir descreve as opções na propriedade Prioridade de Endereço Duplo:

Opção	Descrição
Serviço de entrega	Valida elementos de dados do serviço de entrega em um endereço, tais como elementos de caixa postal.
Administrador postal	Valida os elementos de endereço exigidos pelo serviço de correio local. Opção padrão.
Rua	Valida elementos de dados de rua em um endereço, como elementos de número de edifício e elementos de nome de rua.

## Abreviação de Elemento

Determina se a transformação retorna a forma abreviada de um elemento do endereço. Você pode definir a transformação para retornar o formato abreviado se os dados de referência de endereço tiverem abreviações.

Por exemplo, Serviço Postal dos Estados Unidos (USPS) mantém formatos curtos e longos de muitos nomes de ruas e de localidades. A forma curta de HUNTSVILLE BROWNSFERRY RD é HSV BROWNS FRY RD. Você pode selecionar a propriedade Abreviação de Elemento quando os valores de rua ou de localidade excedem o tamanho de campo máximo especificado pelo USPS.

A opção fica desmarcada por padrão. Defina a propriedade como ON para retornar os valores de endereço abreviados. A propriedade retorna o nome de localidade e o código de localidade abreviados quando você usa a transformação no modo em lotes. A propriedade retorna o nome de rua, o nome de localidade e o código de localidade abreviados quando você usa a transformação no modo certificado.

## Instâncias de Execução

Especifica o número de segmentos que o Serviço de Integração de Dados tenta criar para a transformação atual em tempo de execução. O Serviço de Integração de Dados considerará o valor Instâncias de Execução se você substituir a propriedade de tempo de execução Paralelismo Máximo no mapeamento que contém a transformação. O valor padrão de Instâncias de Execução é 1.

O Serviço de Integração de Dados considera vários fatores para determinar o número de segmentos a serem atribuídos à transformação. Os principais fatores são o valor de Instâncias de Execução e os valores no mapeamento e nos serviços de aplicativo associados no domínio.

O Serviço de Integração de Dados lê os seguintes valores ao calcular o número de segmentos a serem usados para a transformação:

- O valor máximo de *Paralelismo Máximo* no Serviço de Integração de Dados. O padrão é 1.
- Qualquer valor de *Paralelismo Máximo* que você define no nível de mapeamento. O padrão é Automático.
- O valor de *Instâncias de Execução* na transformação. O padrão é 1.

Se você substituir o valor de Paralelismo Máximo no nível de mapeamento, o Serviço de Integração de Dados tentará usar o valor mais baixo nas propriedades para determinar o número de segmentos.

Se você usar o valor de padrão de Paralelismo Máximo no nível de mapeamento, o Serviço de Integração de Dados ignorará o valor de Instâncias de Execução.

O Serviço de Integração de Dados também considera a propriedade *Contagem Máxima de Objetos de Endereço* no Serviço do Gerenciamento de Conteúdo ao calcular o número de segmentos que devem ser criados. A propriedade *Contagem Máxima de Objetos de Endereço* determina o número máximo instâncias de validação de endereço que podem ser executadas simultaneamente em um mapeamento. O valor da propriedade *Contagem Máxima de Objetos de Endereço* deve ser maior que ou igual ao valor de *Paralelismo Máximo* no Serviço de Integração de Dados.

## Regras e Diretrizes para a Propriedade Instâncias de Execução

Considere as seguintes regras e diretrizes quando você definir o número de instâncias de execução:

- Vários usuários podem executar mapeamentos simultâneos em um Serviço de Integração de Dados. Para calcular o número correto de segmentos, divida o número de unidades centrais de processamento que o serviço pode acessar pelo número de mapeamentos simultâneos.
- No PowerCenter, o arquivo de configuração *AD50.cfg* especifica o número máximo instâncias de validação de endereço que podem ser executadas simultaneamente em um mapeamento.
- Quando você usa o valor padrão de Instâncias de Execução e os valores padrão de Paralelismo Máximo, as operações de transformação não são particionáveis.
- Ao definir o valor de Instâncias de Execução como maior que 1, você modifica a transformação de Validador de Endereço de uma transformação passiva para uma transformação ativa.

## Expansão de Intervalo Flexível

Impõe um limite prático ao número de endereços que a transformação do Validador de Endereço retorna quando você define a propriedade Intervalos para Expandir. É possível definir a propriedade Intervalos para Expandir e a propriedade Expansão de Intervalo Flexível ao configurar a transformação para execução no modo de lista de sugestões.

A propriedade Intervalos para Expandir determina como a transformação retorna sugestões de endereço quando um endereço de entrada não contém dados de número de casa. Se o endereço de entrada não incluir dados contextuais, como um código postal completo, a propriedade Intervalos para Expandir pode gerar um grande número de endereços muito semelhantes. A propriedade Expansão de Intervalo Flexível restringe o número de endereços que a propriedade Intervalos para Expandir gera para um único endereço. Defina a propriedade Expansão de Intervalo Flexível como Ligada ao definir a propriedade Intervalo para Expandir como Tudo.

A tabela a seguir descreve as opções na propriedade Expansão de Intervalo Flexível:

Opção	Descrição
Ligada	A validação de endereço limita o número de endereços que a propriedade Intervalos para Expandir adiciona à lista de sugestões. Opção padrão.
Desligada	A validação de endereço não limita o número de endereços que a propriedade Intervalos para Expandir adiciona à lista de sugestões.

**Nota:** A transformação do Validador de Endereço aplica a propriedade Expansão de Intervalo Flexível de uma maneira diferente para cada endereço que ela retorna para a lista de sugestões. A transformação não impõe um limite fixo ao número de endereços expandidos na lista. A transformação também considera a

configuração da propriedade Contagem Máxima de Resultados ao calcular o número de endereços expandidos a incluir na lista.

## Tipo de Dados de Codificação Geográfica

Determina como a transformação do Validador de Endereço calcula dados de codificação geográfica para um endereço. Codificações geográficas são coordenadas de latitude e longitude.

Os resultados de codificação geográfica que a transformação retorna dependem dos dados de referência de codificação geográfica que você instala. Para obter informações sobre dados de referência de codificação geográfica, entre em contato com a Informatica.

Selecione uma das seguintes opções de codificação geográfica:

### Ponto de chegada

Retorna as coordenadas de latitude e longitude da entrada de um prédio ou de um setor do território. Opção padrão.

Você pode selecionar a opção de ponto de chegada de endereços nos seguintes países:

Austrália, Áustria, Canadá, Croácia, Dinamarca, Estônia, Finlândia, França, Alemanha, Hungria, Itália, Letônia, Liechtenstein, Lituânia, Luxemburgo, México, Mônaco, Holanda, Noruega, Polônia, Eslováquia, Eslovênia, Suécia, Suíça e Estados Unidos.

Se você especificar as codificações geográficas do ponto de chegada e a transformação do Validador de Endereço não puder retornar as codificações geográficas de um endereço, a transformação retornará as codificações geográficas interpoladas.

### Padrão

Retorna as coordenadas de latitude e longitude estimadas da entrada do prédio ou o setor do território. Uma codificação geográfica estimada também é chamada de codificação geográfica interpolada.

A transformação do Validador de Endereço usa as codificações geográficas mais próximas disponíveis nos dados de referência para estimar as codificações geográficas do endereço.

**Nota:** A Informatica não emite mais dados de referência para centroide de setor ou codificações geográficas residenciais.

### Utilização de Parâmetros

É possível usar um parâmetro para especificar o tipo de codificação geográfica. Insira `ARRIVAL_POINT` ou `NONE`. Para retornar as codificações geográficas padrão, insira `NONE`.

Insira o valor do parâmetro em letras maiúsculas.

## Comprimento Máximo de Campo Global

Determina o número máximo de caracteres em qualquer linha do endereço. Se a transformação do Validador de Endereço gravar uma linha de endereço de saída que contenha mais caracteres do que você especificar, a transformação abreviará a elementos de endereço na linha.

Use a propriedade para controlar o comprimento da linha no endereço. Por exemplo, os padrões SNA exigem que um endereço não contenha mais de 38 caracteres em qualquer linha. Se você gerar endereços para o padrão SNA, defina o Comprimento Máximo de Campo Global como 38.

O padrão é 1024.

## Utilização de Parâmetros

É possível usar um parâmetro para especificar o número máximo de endereços. Para definir o valor do parâmetro, insira um número inteiro de 0 a 1024.

## Descritor Preferencial Global

Determina o formato do descritor de edifício, de subedifício e de rua que a transformação do Validador de Endereço grava nos dados de saída. Selecione um descritor quando os dados de referência de endereço do país de destino contiverem um intervalo de descritores de um ou mais elementos de dados.

A seguinte tabela descreve as opções na propriedade:

Opção	Descrição
Banco de Dados	Retorna o descritor que o banco de dados de referência especifica para o elemento no endereço. Se o banco de dados não especificar um descritor para o endereço, a transformação copiará o valor de entrada no endereço de saída. O valor padrão é banco de dados.
Longo	Retorna a forma completa do descritor, por exemplo <i>Rua</i> .
Preservar Entrada	Copia o descritor do endereço de entrada para o endereço de saída. Se o descritor de entrada não for uma versão válida do descritor, a transformação retornará um descritor válido equivalente do banco de dados de referência.
Curta	Retorna uma forma abreviada do descritor, por exemplo <i>R</i> .

## Tipo de Formato de Entrada

Descreve o tipo mais comum de informações contidas em dados de entrada sem campos. Use a propriedade Tipo de Formato de Entrada ao conectar dados de entrada às portas de Linha de Endereço Formatada ou Endereço Completo. Selecione a opção que melhor descreva as informações dos dados de origem de mapeamento.

Selecione uma das seguintes opções:

- Tudo
- Endereço
- Organização
- Contato
- Organização/Contato  
O endereço inclui informações da organização e informações de contato.
- Organização/Departamento  
O endereço inclui informações da organização e informações do departamento.

O padrão é Tudo.

## Formato de Entrada com País

Especifica se a entrada contém dados de países. Selecione a propriedade se você conectar dados às portas de entrada Endereço Completo e Linha de Endereço Formatada, e se os dados contiverem informações do país.

A opção fica desmarcada por padrão.

## Separador de Linha

Especifica o símbolo delimitador que indica quebras de linha em um endereço formatado.

Selecione uma das seguintes opções:

- Atribua um parâmetro para identificar o separador de linha
- Retorno de carro (CR)
- Vírgula
- Avanço de Linha (LF)
- Nenhum
- Ponto e vírgula
- Guia
- Nova Linha do Windows (CRLF)

O padrão é ponto e vírgula.

Você também pode configurar o separador de linha na guia **Configurações Gerais**.

### Utilização de Parâmetros

É possível usar um parâmetro para especificar o separador de linha. O valor do parâmetro diferencia maiúsculas de minúsculas. Insira o valor do parâmetro em caracteres maiúsculos.

Insira um dos seguintes valores:

- CR
- CRLF
- COMMA
- LF
- PIPE
- SEMICOLON
- SPACE
- TAB

## Alternativas de Correspondência

Determina se a validação de endereço reconhece nomes de locais alternativos, como sinônimos ou nomes históricos, em um endereço de entrada. A propriedade se aplica a dados de rua, localidade e província.

**Nota:** A propriedade Alternativas de Correspondência não preserva os nomes alternativos em um endereço validado.

A tabela seguinte descreve as opções de Alternativas de Correspondência:

Opção	Descrição
Tudo	Reconhece todos os nomes de rua e de localidade alternativos conhecidos. Opção padrão.
Somente arquivos	Reconhece apenas nomes históricos. Por exemplo, a validação de endereço valida "Constantinopla" como uma versão histórica de "Istambul".
Nenhum	Não reconhece nomes de rua ou de localidade alternativos.
Sinônimos somente	Reconhece apenas sinônimos e exônimos. Por exemplo, a validação de endereço valida "Londres" como um exônimo de "London".

## Arquivo Morto Estendido Correspondente

Determina se a validação de endereço retorna um código de ponto de entrega exclusivo para um endereço japonês desatualizado.

Os arquivos de dados de referência de endereço para o Japão incluem endereços desatualizados ou retirados junto com os endereços atuais das caixas de correio correspondentes. Quando você seleciona a propriedade Arquivo Morto Estendido Correspondente, a validação de endereço retorna o código de ponto de entrega para a versão atual de cada endereço. A validação de endereço também grava um valor na porta Status do Resultado do Elemento Estendido para indicar que o endereço de entrada está desatualizado.

Para recuperar o endereço atual dos dados de referência de endereço, insira o código do endereço como um elemento de entrada.

A seguinte tabela descreve as opções de Arquivo Morto Estendido Correspondente:

Opção	Descrição
Desativado	Não aplica a propriedade.
Ativado	Retorna o código de endereço para a versão atual de um endereço japonês desatualizado.

A propriedade Arquivo Morto Estendido Correspondente usa dados complementares e dados de pesquisa de códigos de endereço para o Japão. Para aplicar a propriedade à validação de endereço, configure a transformação para que ela seja executada no modo de pesquisa de código.

## Escopo de Correspondência

Determina a quantidade de correspondência de dados que a transformação encontra em relação aos dados de referência de endereço durante a validação de endereço.

A tabela seguinte descreve as opções do Escopo de Correspondência:

Opção	Descrição
Tudo	Valida todas as portas selecionadas. Opção padrão.
Ponto de Entrega	Valida dados de endereço de edifícios e subedifícios, além dos dados validados pela opção Rua.
Localidade	Valida dados de província, localidade e código postal.
Rua	Valida dados de endereço de rua, além dos dados validados pela opção Localidade.

## Contagem Máxima de Resultados

Determina o número máximo de endereços que a validação de endereço pode retornar no modo de lista de sugestões.

Você pode definir um número máximo no intervalo de 1 a 100. O padrão é 20.

**Nota:** O modo de lista de sugestões realiza uma verificação de endereço nos dados de referência de endereço e retorna uma lista de endereços que são correspondências possíveis para o endereço de entrada. Quando você verifica um modo de lista de sugestões, a validação de endereço retorna as melhores correspondências primeiro.

### Utilização de Parâmetros

É possível usar um parâmetro para especificar o número máximo de endereços. Para definir o valor do parâmetro, insira um número inteiro de 0 a 100.

## Modo

Determina o tipo de análise de endereço executado pela transformação. Também é possível configurar o modo na guia **Configurações Gerais** da transformação.

### TÓPICOS RELACIONADOS:

- [“ Transformação de Validador de Endereço - Configurações Gerais” na página 102](#)



## Nível de Otimização

Determina como a transformação corresponde os dados de endereço de entrada e dados de referência de endereço. A propriedade define o tipo de correspondência que a transformação deve encontrar entre os dados de entrada e os dados de referência antes de atualizar o registro de endereço.

A seguinte tabela descreve as opções de Nível de Otimização:

Opção	Descrição
Restrita	A transformação analisa números de edifícios ou residências a partir de informações de rua antes de realizar a validação. Caso contrário, a transformação validará os elementos de endereço de entrada estritamente de acordo com a estrutura da porta de entrada. A opção restrita executa a validação de endereço mais rápida, mas pode retornar resultados menos precisos que as outras opções.
Padrão	A transformação analisa vários tipos de informações de endereço a partir de dados de entrada antes de realizar a validação. Quando você selecionar a opção padrão, a transformação atualizará um endereço se conseguir corresponder vários valores de entrada aos dados de referência. É o padrão.
Ampla	A transformação usa as configurações de análise padrão e realiza operações de análise adicionais nos dados de entrada. Quando você selecionar a opção Ampla, a transformação atualizará um endereço se conseguir corresponder pelo menos um valor de entrada aos dados de referência. A opção Ampla aumenta o tempo de execução do mapeamento.

### Utilização de Parâmetros

É possível usar um parâmetro para especificar o nível de otimização. Insira NARROW, STANDARD ou WIDE. Insira o valor do parâmetro em letras maiúsculas.

## Tipo de Formato de Saída

Descreve o tipo mais comum de informações que a transformação grava na porta de saída Endereço Completo e Linha de Endereço Formatada. Selecione a opção que melhor descreve os dados que você espera na porta de saída.

Selecione uma das seguintes opções:

- Tudo
- Endereço
- Organização
- Contato
- Organização/Contato  
O endereço inclui informações da organização e informações de contato.
- Organização/Departamento  
O endereço inclui informações da organização e informações do departamento.

O padrão é Tudo.

## Formato de Saída com País

Determina se a transformação grava dados de identificação de país nas portas de saída de Endereço Completo ou de Linha de Endereço Formatada.

A opção fica desmarcada por padrão.

## Idioma Preferencial

Determina os idiomas em que a transformação de Validador de Endereço retorna elementos de endereço quando os conjuntos de dados de referência contêm dados em mais de um idioma. Você pode definir um idioma preferencial para endereços na Bélgica, no Canadá, na China, na Finlândia, em Hong Kong, na Irlanda, em Israel, em Macau, na Suíça e em Taiwan.

A transformação de Validador de Endereço pode retornar dados de endereços nos seguintes idiomas:

- O idioma padrão para o endereço nos dados de referência de endereço. O idioma padrão é a principal língua falada na região à qual cada endereço pertence.
- Qualquer outro idioma ao qual os dados de referência de endereço oferecem suporte para um endereço. Por exemplo, os dados de referência da Bélgica contêm elementos de endereço em flamengo, francês e alemão.

Os dados de referência de endereço podem conter dados para um elemento de endereço único ou para um endereço completo em vários idiomas. Por exemplo, a validação de endereço pode retornar todos os elementos de endereço da Irlanda no idioma inglês e pode retornar informações de rua, localidade e província no idioma irlandês. Além disso, os dados de referência podem especificar idiomas padrão diferentes para endereços em diversas partes de um país. Por exemplo, nos dados de referência da Suíça, o idioma padrão varia de região para região entre francês, alemão e italiano.

A seguinte tabela resume as opções que você pode selecionar na propriedade Idioma Preferencial:

Opção	Descrição
Banco de dados	Retorna cada endereço no idioma que os dados de referência de endereço especifica. Os dados de referência de endereço podem especificar diferentes idiomas para endereços em diferentes regiões em um país. Banco de dados é a opção padrão.
Alternativa 1, Alternativa 2, Alternativa 3	Retorna elementos de endereço em um idioma alternativo com base nos dados de referência. Os idiomas alternativos dependem do país ao qual o endereço pertence.
Inglês	Retorna elementos de endereço em inglês quando os dados de referência contêm os dados em inglês. Retorna os outros elementos de endereço no idioma padrão da região ao qual o endereço pertence.
Preservar Entrada	Retorna as informações de endereço no idioma de entrada. A validação de endereço preserva o idioma se os dados de referência contiverem as informações de endereço no idioma de entrada. Se a validação de endereço detectar mais de um idioma suportado no endereço de entrada, ele retornará o endereço no idioma do banco de dados. Se a Verificação de Endereço não puder retornar um elemento no idioma de entrada, ela retornará o elemento no idioma do banco de dados.

**Nota:** Um conjunto de dados de referência de endereço pode conter alguns elementos de endereço em um idioma não padrão, mas não outros. Se a transformação não puder localizar um elemento no idioma que a propriedade especificar, a transformação retornará o elemento no idioma padrão.

Quando você definir uma opção de idioma preferencial, verifique se o conjunto de caracteres especificado pela propriedade Script Preferencial é compatível com os dados de endereço de saída que você espera.

### Suporte multilíngue para endereços na Bélgica

A tabela a seguir descreve os idiomas que você pode especificar para endereços na Bélgica:

Opção	Descrição
Banco de Dados	Valor padrão. Retorna endereços no idioma principal da região à qual o endereço pertence. O idioma pode ser flamengo, francês ou alemão.
Inglês	Retorna as informações de província, localidade e rua em inglês se os dados de referência de endereço contiverem as informações em inglês. Retorna os outros elementos de endereço no idioma principal da região à qual o endereço pertence.
Alternativa 1	Retorna as informações de província, localidade e rua em flamengo.
Alternativa 2	Retorna as informações de província, localidade e rua em francês.
Alternativa 3	Retorna as informações de província, localidade e rua em alemão.
Preservar Entrada	Retorna as informações de endereço no idioma de entrada. A validação de endereço preserva o idioma se os dados de referência contiverem as informações de endereço no idioma de entrada. Se a validação de endereço detectar mais de um idioma suportado no endereço de entrada, ele retornará o endereço no idioma do banco de dados. Se a Verificação de Endereço não puder retornar um elemento no idioma de entrada, ela retornará o elemento no idioma do banco de dados.

### Suporte multilíngue para endereços no Canadá

A tabela a seguir descreve os idiomas que você pode especificar para endereços no Canadá:

Opção	Descrição
Banco de Dados	Valor padrão. Retorna endereços em inglês para todas as províncias, exceto Quebec. Retorna endereços de Quebec em francês.
Inglês	Retorna todos os endereços em inglês.
Alternativa 1	Retorna todos os endereços em inglês.
Alternativa 2	Retorna endereços de Quebec em francês. Em províncias diferentes de Quebec, a transformação retorna os descritores de rua, as informações de trajetos e os nomes das províncias em francês, e retorna outros elementos de endereço em inglês.
Preservar Entrada	Retorna as informações de endereço no idioma de entrada. A validação de endereço preserva o idioma se os dados de referência contiverem as informações de endereço no idioma de entrada. Se a validação de endereço detectar mais de um idioma suportado no endereço de entrada, ele retornará o endereço no idioma do banco de dados. Se a Verificação de Endereço não puder retornar um elemento no idioma de entrada, ela retornará o elemento no idioma do banco de dados.

## Suporte a vários idiomas para endereços na China

A tabela a seguir descreve os idiomas que você pode especificar para endereços na China:

Opção	Descrição
Banco de Dados	Valor padrão. Retorna todas as informações de endereço em chinês.
Inglês	Retorna as versões no idioma inglês do descritor de rua e dos valores de trajetos de ruas. Retorna todas as outras informações de endereço no idioma chinês. Os elementos de endereço em inglês omitem elementos de transliteração, como "shi".
Alternativa 1	Retorna todas as informações de endereço no idioma do banco de dados.
Alternativa 2	Retorna todas as informações de endereço no idioma do banco de dados.
Alternativa 3	Retorna todas as informações de endereço no idioma do banco de dados.
Preservar Entrada	Retorna as informações de endereço no idioma de entrada. A validação de endereço preserva o idioma se os dados de referência contiverem as informações de endereço no idioma de entrada. Se a validação de endereço detectar mais de um idioma suportado no endereço de entrada, ele retornará o endereço no idioma do banco de dados. Se a Verificação de Endereço não puder retornar um elemento no idioma de entrada, ela retornará o elemento no idioma do banco de dados.

Considere as seguintes regras e diretrizes quando você selecionar o idioma preferencial:

- Para retornar o endereço no idioma chinês, selecione Banco de Dados, Alternativa 1, Alternativa 2 ou Alternativa 3.  
Para retornar o endereço em um conjunto de caracteres chinês, defina a propriedade Script Preferencial como Banco de Dados.
- Para retornar o descritor de rua e as informações de trajetos de ruas no idioma inglês, selecione Inglês.  
Para retornar o endereço em um conjunto de caracteres Latino ou ASCII, defina a propriedade Script Preferencial como o valor LATIN ou ASCII.
- Se você selecionar um valor LATIN ou ASCII como o script preferencial, e Banco de Dados como o idioma preferencial, a validação de endereço retornará os dados de endereço no Pinyin.

## Suporte multilíngue para endereços na Finlândia

A tabela a seguir descreve os idiomas que você pode especificar para endereços na Finlândia:

Opção	Descrição
Banco de Dados	Valor padrão. Retorna todas as informações de endereço em finlandês.
Alternativa 1	Retorna todas as informações de endereço no idioma do banco de dados.

Opção	Descrição
Alternativa 2	Retorna as informações de rua, localidade e província em sueco. Retorna todas as outras informações em finlandês.
Preservar Entrada	Retorna as informações de endereço no idioma de entrada. A validação de endereço preserva o idioma se os dados de referência contiverem as informações de endereço no idioma de entrada. Se a validação de endereço detectar mais de um idioma suportado no endereço de entrada, ele retornará o endereço no idioma do banco de dados. Se a Verificação de Endereço não puder retornar um elemento no idioma de entrada, ela retornará o elemento no idioma do banco de dados.

## Suporte multilíngue para endereços em Hong Kong

A tabela a seguir descreve os idiomas que você pode especificar para endereços em Hong Kong:

Opção	Descrição
Banco de Dados	Valor padrão. Retorna todas as informações de endereço em chinês.
Inglês	Retorna todas as informações de endereço em inglês.
Alternativa 1	Retorna todas as informações de endereço no idioma do banco de dados.
Alternativa 2	Retorna todas as informações de endereço em inglês.
Alternativa 3	Retorna todas as informações de endereço no idioma do banco de dados.
Preservar Entrada	Retorna as informações de endereço no idioma de entrada. A validação de endereço preserva o idioma se os dados de referência contiverem as informações de endereço no idioma de entrada. Se a validação de endereço detectar mais de um idioma suportado no endereço de entrada, ele retornará o endereço no idioma do banco de dados. Se a Verificação de Endereço não puder retornar um elemento no idioma de entrada, ela retornará o elemento no idioma do banco de dados.

Considere as seguintes regras e diretrizes quando você selecionar o idioma preferencial para Hong Kong:

- Para retornar o endereço em um conjunto de caracteres chinês, defina a propriedade Script Preferencial como Banco de Dados.
- Para retornar o endereço em um conjunto de caracteres Latino ou ASCII, defina a propriedade Script Preferencial como o valor LATIN ou ASCII.
- O idioma dos dados de entrada pode afetar a operação da opção Preservar Entrada em um endereço de Hong Kong. A validação de endereço identifica o idioma de entrada como inglês quando os dados de entrada usam caracteres ASCII de 7 bits e inclui um descritor de idioma inglês.

## Suporte a vários idiomas para endereços na Irlanda

A tabela a seguir descreve os idiomas que você pode especificar para endereços na Irlanda:

Opção	Descrição
Banco de Dados	Valor padrão. Retorna todas as informações de endereço em inglês.
Inglês	Retorna todas as informações de endereço em inglês.
Alternativa 1	Retorna todas as informações de endereço em inglês.
Alternativa 2	Retorna as informações de rua, localidade e província em irlandês. Retorna todas as outras informações de endereço em inglês.
Preservar Entrada	Retorna as informações de endereço no idioma de entrada. A validação de endereço preserva o idioma se os dados de referência contiverem as informações de endereço no idioma de entrada. Se a validação de endereço detectar mais de um idioma suportado no endereço de entrada, ele retornará o endereço no idioma do banco de dados. Se a Verificação de Endereço não puder retornar um elemento no idioma de entrada, ela retornará o elemento no idioma do banco de dados.

## Suporte a vários idiomas para endereços em Israel

A tabela a seguir descreve os idiomas que você pode especificar para endereços em Israel:

Opção	Descrição
Banco de Dados	Valor padrão. Retorna todas as informações de endereço em hebraico.
Inglês	Retorna todas as informações de endereço em inglês.
Alternativa 1	Retorna todas as informações de endereço em hebraico.
Alternativa 2	Retorna todas as informações de endereço em inglês.
Preservar Entrada	Retorna as informações de endereço no idioma de entrada. A validação de endereço preserva o idioma se os dados de referência contiverem as informações de endereço no idioma de entrada. Se a validação de endereço detectar mais de um idioma suportado no endereço de entrada, ele retornará o endereço no idioma do banco de dados. Se a Verificação de Endereço não puder retornar um elemento no idioma de entrada, ela retornará o elemento no idioma do banco de dados.

Considere as seguintes regras e diretrizes quando você selecionar o idioma preferencial:

- Para retornar os endereços em um conjunto de caracteres hebraicos, defina a propriedade Script Preferencial como Banco de Dados.
- Para retornar os endereços em um conjunto de caracteres Latino ou ASCII, defina a propriedade Script Preferencial como o valor LATIN ou ASCII.
- Se você selecionar um conjunto de caracteres latinos como o script preferencial e selecionar Hebraico como o idioma preferencial, a validação do endereço transliterará o endereço hebraico para caracteres latinos. Para obter melhores resultados em um conjunto de caracteres latinos, selecione Inglês como o idioma preferido.

## Suporte multilíngue para endereços em Macau

A tabela a seguir descreve os idiomas que você pode especificar para endereços em Macau:

Opção	Descrição
Banco de Dados	Valor padrão. Retorna todas as informações de endereço em chinês.
Alternativa 1	Retorna todas as informações de endereço no idioma do banco de dados.
Alternativa 2	Retorna todas as informações de endereço em português.
Preservar Entrada	Retorna as informações de endereço no idioma de entrada. A validação de endereço preserva o idioma se os dados de referência contiverem as informações de endereço no idioma de entrada. Se a validação de endereço detectar mais de um idioma suportado no endereço de entrada, ele retornará o endereço no idioma do banco de dados. Se a Verificação de Endereço não puder retornar um elemento no idioma de entrada, ela retornará o elemento no idioma do banco de dados.

- Para retornar o endereço em um conjunto de caracteres chinês, defina a propriedade Script Preferencial como Banco de Dados.
- Para retornar o endereço em um conjunto de caracteres Latino ou ASCII, defina a propriedade Script Preferencial como o valor LATIN ou ASCII.
- O idioma dos dados de entrada pode afetar a operação da opção Preservar Entrada em um endereço de Macau. A validação de endereço identifica o idioma de entrada como português quando os dados de entrada usam caracteres ASCII de 7 bits e inclui um descritor de idioma português.

## Suporte multilíngue para Suíça

A tabela a seguir descreve os idiomas que você pode especificar para endereços na Suíça:

Opção	Descrição
Banco de Dados	Valor padrão. Retorna endereços no idioma principal da região à qual o endereço pertence. Por exemplo, a validação de endereço retorna um endereço de Zurique em alemão e um endereço de Genebra em francês.
Inglês	Retorna as informações de localidade e província em inglês se o banco de dados de endereço de referência contém as informações em inglês. Retorna os outros elementos de endereço no idioma principal da região à qual o endereço pertence. A validação de endereço retorna as informações de localidade em inglês para algumas localidades, por exemplo, Genebra e Zurique.
Alternativa 1	Retorna as informações de província e localidade em alemão.
Alternativa 2	Retorna as informações de província e localidade em francês.

Opção	Descrição
Alternativa 3	Retorna as informações de província e localidade em italiano.
Preservar Entrada	Retorna as informações de endereço no idioma de entrada. A validação de endereço preserva o idioma se os dados de referência contiverem as informações de endereço no idioma de entrada. Se a validação de endereço detectar mais de um idioma suportado no endereço de entrada, ele retornará o endereço no idioma do banco de dados. Se a Verificação de Endereço não puder retornar um elemento no idioma de entrada, ela retornará o elemento no idioma do banco de dados.

**Nota:** A validação de endereço também retorna informações de rua para endereços em Bienna no idioma alternativo configurado.

### Suporte multilíngue para Taiwan

A tabela a seguir descreve os idiomas que você pode especificar para endereços em Taiwan:

Opção	Descrição
Banco de Dados	Valor padrão. Retorna todas as informações de endereço em chinês.
Inglês	Retorna todas as informações de endereço em inglês.
Preservar Entrada	Retorna as informações de endereço no idioma de entrada. A validação de endereço preserva o idioma se os dados de referência contiverem as informações de endereço no idioma de entrada. Se a validação de endereço detectar mais de um idioma suportado no endereço de entrada, ele retornará o endereço no idioma do banco de dados. Se a Verificação de Endereço não puder retornar um elemento no idioma de entrada, ela retornará o elemento no idioma do banco de dados.

Considere as seguintes regras e diretrizes quando você selecionar o idioma preferencial:

- Para retornar o endereço em um conjunto de caracteres chinês, defina o parâmetro Script Preferencial como Banco de Dados.
- Para retornar o endereço em um conjunto de caracteres Latino ou ASCII, defina o parâmetro Script Preferencial como o valor LATIN ou ASCII.
- O idioma dos dados de entrada pode afetar a operação da opção Preservar Entrada em um endereço de Taiwan. A validação de endereço identifica o idioma de entrada como inglês quando os dados de entrada usam caracteres ASCII de 7 bits e inclui um descritor de idioma inglês.

## Script Preferencial

Determina o conjunto de caracteres que a transformação do Validador de Endereço usa para dados de saída.

A tabela seguinte descreve as opções da propriedade:

Opção	Descrição
ASCII (Simplificado)	Retorna um endereço em caracteres ASCII.
ASCII (Estendido)	Retorna um endereço em caracteres ASCII e expande qualquer caractere especial em um endereço. Por exemplo, Õ é transliterado para OE.



Opção	Descrição
Banco de Dados	Retorna um endereço no conjunto de caracteres que os dados de referência de endereço usam para o idioma padrão. O valor padrão é Banco de Dados.
Latino	Retorna um endereço no conjunto de caracteres Latino.
Latino (Alternativo)	Retorna um endereço em um conjunto de caracteres Latino alternativo. Por exemplo, especifique Latino para retornar um endereço da Coreia do Sul na transliteração Romanização Revisada. Especifique Latino (Alternativo) para retornar um endereço da Coreia do Sul na transliteração mais antiga ISO/TR 11941.
Administrador Postal	Retorna um endereço no script que o serviço postal local ao endereço prefere.
Administrador Postal (Alt)	Retorna um endereço em um script que o serviço postal local ao endereço aprova como um script alternativo.
Preservar Entrada	Retorna dados de endereço no conjunto de caracteres que o endereço de entrada usa.

A transformação pode processar uma fonte de dados que contém dados em vários idiomas e conjuntos de caracteres. A transformação converte todos os dados de entrada para o conjunto de caracteres Unicode UCS-2 e processa os dados no formato UCS-2. Depois de processar os dados, a transformação converte os dados de cada registro de endereço no conjunto de caracteres que você especifica na propriedade. O processo é chamado de transliteração.

A transliteração pode usar representações numéricas de cada caractere em um conjunto de caracteres quando converte os caracteres para o processamento. Ela também poderá converter caracteres foneticamente quando não houver representação numérica equivalente de um caractere. Se a transformação do Validador de Endereço não puder mapear um caractere em UCS-2, ela converterá o caractere para um espaço.

**Nota:** Se você atualizar o idioma ou script preferencial na transformação, verifique se o idioma e o código de caractere selecionado são compatíveis.

## Intervalos para Expandir

Determina como a transformação do Validador de Endereço retorna endereços sugeridos para um endereço que não especifica um número de casa. Use a propriedade ao executar a transformação no modo de lista de sugestões.

A transformação do Validador de Endereço lê um endereço parcial ou incompleto no modo de lista de sugestões. A transformação compara o endereço a os dados de referência de endereço e retorna todos os endereços semelhantes para o usuário final. Se o endereço de entrada não contém um número de casa, a transformação pode retornar uma ou mais sugestões de número da casa para a rua. A propriedade Intervalos para Expandir determina como a transformação retorna os endereços.

A transformação pode retornar o intervalo de números de casa válidos em um único endereço, ou pode retornar um endereço separado para cada número da casa válido. A transformação também pode retornar um endereço para cada número no intervalo do número de casa mais baixo na rua ao mais alto.

A tabela seguinte descreve as opções da propriedade:

Opção	Descrição
Tudo	A validação de endereço retorna um endereço sugerido para cada número de casa no intervalo de possíveis números de residências na rua.
Nenhum	A validação de endereço retorna um único endereço que identifica o menor e o maior número de casa no intervalo válido para a rua.
Somente com itens válidos	A validação de endereço retorna um endereço sugerido para cada número de casa reconhecido pelos dados de referência de endereço como um endereço com capacidade de entrega.

**Nota:** O modo de lista de sugestões pode usar outros elementos do endereço para especificar o intervalo válido de números de rua. Por exemplo, um código de CEP pode identificar o bloco da cidade que contém a caixa postal do endereço. A transformação do Validador de Endereço pode usar o código de CEP para identificar o maior e o menor número de residência no bloco.

Se a transformação não pode determinar um intervalo de números de residências dentro de limites práticos, o número de endereços sugeridos pode aumentar para um tamanho inutilizável. Para restringir o número de endereços que a propriedade Intervalos para Expandir gera, defina a propriedade Expansão de Intervalo Flexível como Ligada.

## Padronizar Endereços Inválidos

Determina se o processo de validação de endereço padroniza os valores de dados em um endereço sem capacidade de entrega. A propriedade se aplica a registros de endereço que retornam um status de Código de Correspondência no intervalo de I1 a I4.

Ao padronizar os dados, você aumenta a probabilidade de que um processo de dados downstream retorne resultados precisos. Por exemplo, um mapeamento de análise duplicada pode retornar uma pontuação de correspondência mais alta para dois registros de endereço que apresentem elementos de endereço comum no mesmo formato.

A validação de endereço pode padronizar os seguintes elementos de endereço:

- Elementos de sufixo de rua, como estrada e alameda.
- Elementos pré-direcionais e pós-direcionais, como norte, sul, leste e oeste.
- Elementos de serviço de entrega, como caixa postal.
- Elementos sub-edifício, como apartamento, andar e suite.
- Nomes de estados ou províncias. A padronização retorna as formas abreviadas dos nomes.

A tabela seguinte descreve as opções da propriedade:

Opção	Descrição
Desativado	A validação de endereço não corrige erros de dados. Opção Padrão.
Ligado	A validação de endereço corrige erros de dados.

### Utilização de Parâmetros

Você pode atribuir um parâmetro para especificar a diretiva de padronização para erros de dados. Insira OFF ou ON como o valor do parâmetro. Insira o valor em letras maiúsculas.

## Nível de Rastreamento

Define a quantidade de detalhes incluída no log.

Você pode configurar níveis de rastreamento para logs.

Configure a seguinte propriedade na guia **Avançado**:

### **Nível de Rastreamento**

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.

## Relatórios de Certificação

Você pode gerar um relatório que descreve o status dos endereços enviados para validação no modo certificado. O relatório verifica se o endereço definido atende aos padrões de certificação definidos pelo serviço de correio.

A transformação do Validador de Endereço gera relatórios para os seguintes padrões:

### **Address Machine Approval System (AMAS)**

O Correio da Austrália define o padrão de certificação AMAS.

### **Coding Accuracy Support System (CASS)**

O USPS define o padrão de certificação CASS.

### **SendRight**

O Correio da Nova Zelândia define o padrão de certificação SendRight.

### **Programa de Reconhecimento e Avaliação de Software (SERP)**

O Correio do Canadá define o padrão de certificação SERP.

**Nota:** É possível usar a transformação do Validador de Endereço para certificar registros de endereço franceses para o padrão de certificação SNA (Serviço de Gerenciamento de Endereços Nacionais). No entanto, você não gera um relatório para certificação de endereço SNA na transformação de Validador de Endereço.

Ao fornecer os itens de correio que atendem aos padrões de certificação para o serviço de correio, você envia o relatório de certificação com o endereço definido. O relatório inclui dados sobre a organização. Você pode inserir os dados ao configurar a transformação de Validador de Endereço. A transformação adiciona os dados da organização ao arquivo de relatório.

## Campos de Relatório AMAS

Ao configurar um relatório AMAS, você fornece informações sobre a organização que envia o conjunto de registros de endereço certificados para o Correio da Austrália. Salve ou imprima o relatório e inclua-o com os registros de endereço que você enviar para o Correio da Austrália.

Use a exibição **Relatórios** para inserir as informações.

A seguinte tabela descreve as informações que você insere:

Campo	Descrição
Nome do Arquivo de Relatório	Nome e localização do relatório que a validação de endereço cria. Por padrão, o Validador de Endereço cria o relatório no diretório <code>bin</code> da máquina do Serviço de Integração de Dados. Para gravar o arquivo de relatório em outro local na máquina do Serviço de Integração de Dados, digite o caminho e o nome do arquivo. Você pode inserir um caminho totalmente qualificado ou um caminho relativo. O caminho relativo usa o diretório <code>bin</code> como um diretório raiz. O diretório especificado deverá existir antes de você executar o mapeamento de validação de endereço.
Nome da Lista de Endereços	Nome do conjunto de registros de endereço que você envia para o Correio da Austrália.
Nome do Processador da Lista	Nome da organização que envia o conjunto de registros de endereço.
Nome do Gerente/ Proprietário da Lista	Nome do gerente ou proprietário dos dados de endereço na organização.
Número do Telefone	Número de telefone para contato da organização que envia o conjunto de registros de endereço.
Endereço	Endereço da organização que envia o conjunto de registros de endereço.

#### TÓPICOS RELACIONADOS:

- [“Definindo um Relatório de Certificação” na página 132](#)

## Campos de Relatório CASS

Ao configurar um relatório CASS, você fornece informações sobre a organização que envia o conjunto de registros de endereço certificados ao USPS. Salve ou imprima o relatório e inclua-o com os registros de endereço que você enviar para o USPS.

Use a exibição **Relatórios** para inserir as informações.

A seguinte tabela descreve as informações que você insere:

Campo	Descrição
Nome do Arquivo de Relatório	O nome e a localização do relatório criado pela validação de endereço. Por padrão, o Validador de Endereço cria o relatório no diretório <code>bin</code> da máquina do Serviço de Integração de Dados. Para gravar o arquivo de relatório em outra localização na máquina do Serviço de Integração de Dados, insira o caminho e o nome do arquivo. Você pode inserir um caminho totalmente qualificado ou relativo. O caminho relativo usa o diretório <code>bin</code> como um diretório raiz. O diretório especificado deverá existir antes que você execute o mapeamento de validação de endereço.
Nome/ID da Lista	Nome ou número de identificação da lista de endereços que você envia ao prestador de serviços postais.

Campo	Descrição
Nome do Processador da Lista	Nome da organização que executa a validação de endereço.
Name/Address	Nome e endereço postal da organização que executa a validação de endereço.

**Nota:** A transformação de Validador de Endereço lê metadados sobre os arquivos de dados de referência CASS a partir do arquivo de dados de referência em lote e interativos para os Estados Unidos. Para gerar um relatório CASS, a transformação deve ler os arquivos de dados de referência CASS atuais e o arquivo de dados de referência em lote e interativos atual.

## TÓPICOS RELACIONADOS:

- [“Definindo um Relatório de Certificação” na página 132](#)

## Relatório SendRight

Ao configurar um relatório SendRight, você fornece informações sobre a organização que envia o conjunto de registros de endereço certificados ao Correio da Nova Zelândia. Salve ou imprima o relatório e inclua-o com os registros de endereço que você enviar para o Correio da Nova Zelândia.

Use a exibição **Relatórios** para inserir as informações.

A seguinte tabela descreve as informações que você insere:

Campo	Descrição
Nome do Cliente	O nome da organização que envia o conjunto de registros de endereço.
Número NZP do Cliente:	O número da conta do Correio da Nova Zelândia da organização que envia o conjunto de registros de endereço. Se um serviço postal envia os registros em nome da organização, insira o número de identificação do transporte do serviço postal (TPID).
Banco de Dados de Clientes	O nome do arquivo que contém o conjunto de registros de endereço. A transformação do Validador de Endereço cria o relatório na localização que você especificar no Serviço do Gerenciamento de Conteúdo. Use a ferramenta Administrator para definir a localização.
Endereço do Cliente	O endereço da organização que envia o conjunto de registros de endereço.

## TÓPICOS RELACIONADOS:

- [“Definindo um Relatório de Certificação” na página 132](#)

## Campos de Relatório SERP

Ao configurar um relatório SERP, você fornece informações sobre a organização que envia o conjunto de registros de endereço certificados ao Correio do Canadá. Salve ou imprima o relatório e inclua-o com os registros de endereço que você enviar para o Correio do Canadá.

Use a exibição **Relatórios** para inserir as informações.

A seguinte tabela descreve as informações que você insere:

Campo	Descrição
Nome do Arquivo de Relatório	O nome e a localização do relatório criado pela validação de endereço. Por padrão, o Validador de Endereço cria o relatório no diretório <code>bin</code> da máquina do Serviço de Integração de Dados. Para gravar o arquivo de relatório em outra localização na máquina do Serviço de Integração de Dados, insira o caminho e o nome do arquivo. Você pode inserir um caminho totalmente qualificado ou relativo. O caminho relativo usa o diretório <code>bin</code> como um diretório raiz. O diretório especificado deverá existir antes que você execute o mapeamento de validação de endereço.
Número CPC do Cliente	O número de cliente emitido pelo Canada Post Corporation para a organização que realiza a validação de endereço.
Nome/Endereço do Cliente	O nome e o endereço da organização que realiza a validação de endereço.

#### TÓPICOS RELACIONADOS:

- [“Definindo um Relatório de Certificação” na página 132](#)

## Configurando uma Transformação de Validador de Endereço

Use uma transformação de Validador de Endereço para validar e melhorar a qualidade dos seus dados de endereço postal.

A transformação de Validador de Endereço lê dados de referência de endereço. Verifique se a ferramenta Developer pode acessar os arquivos de dados de referência de endereço de que você precisa.

1. Abra a transformação.
2. Clique na exibição **Configurações Gerais** e configure as propriedades gerais.
3. Clique na exibição **Modelos** para adicionar portas de entrada e saída.
4. Clique na exibição **Relatórios** para gerar relatórios para certificação de endereço do serviço postal.
5. Clique na exibição **Avançado** para configurar as propriedades avançadas de validação de endereço.
6. Conecte as portas de entrada e saída.

**Nota:** Conecte ao grupo de portas de entrada de **Passagem** as portas de entrada que você não quer que sejam validadas pela Transformação de Validador de Endereço.

## Adicionando Portas à Transformação de Validador de Endereço

Use a exibição **Modelos** para adicionar portas à transformação de Validador de Endereço.

1. Clique na exibição **Modelos**.

2. Expanda um modelo.
  - Escolha o **Modelo Básico** para adicionar campos de endereço comuns.
  - Escolha o **Modelo Avançado** para adicionar campos de endereço especializados.
3. Expanda o grupo de portas de entrada que corresponde ao formato dos seus dados de entrada. Os grupos de portas de entrada são **Distinto, Várias Linhas e Híbrido**.
4. Selecione as portas de entrada.

**Sugestão:** Pressione a tecla **CTRL** para selecionar várias portas.
5. Clique com o botão direito do mouse nas portas e selecione **Adicionar porta à transformação**.
6. Expanda o grupo de portas de saída que contém os campos de que você precisa.
7. Clique com o botão direito do mouse nas portas e selecione **Adicionar porta à transformação**.
8. Para adicionar portas de passagem a colunas que você não quer validar, clique na guia **Portas**, selecione o grupo de portas de entrada de **Passagem** e clique em **Novo**.

## Criando Modelos Definidos pelo Usuário

Crie modelos para agrupar as portas de endereço que você planeja reutilizar.

Você cria modelos personalizados selecionando portas dos modelos Básico e Avançado. Você pode selecionar os modelos personalizados ao criar transformações de Validador de Endereço subsequentes.

**Nota:** Os modelos não são objetos de repositório. Os modelos residem no computador que você usa para criá-los.

1. Selecione a exibição **Modelos**.
2. Clique em **Novo**.
3. Digite um nome para o modelo.
4. Expanda o **Modelo Básico** ou o **Modelo Avançado** e selecione as portas necessárias.
5. Clique em **OK**.

## Definindo Modelos do Validador de Endereço

Os modelos de Validador de Endereço definem as portas de entrada e saída padrão para as transformações de Validador de Endereço.

As transformações de Validador de Endereço não contêm portas de entrada e saída padrão. No entanto, você pode definir um modelo para especificar as portas de entrada e saída usadas pelas transformações de Validador de Endereço.

**Nota:** Os modelos não são objetos de repositório. Os modelos residem no computador que você usa para criá-los.

Para definir um modelo de Validador de Endereço, execute as seguintes etapas:

1. Selecione a exibição **Modelos**.
2. Expanda o **Modelo Básico** ou o **Modelo Avançado** e selecione as portas necessárias.

3. Selecione **Criar modelo AV padrão usando portas selecionadas**.
4. Para redefinir o modelo e remover todas as portas, selecione **Limpar modelo AV padrão**.

## Definindo um Relatório de Certificação

Ao definir um relatório de certificação na transformação de Validador de Endereço, você configura as opções nas exibições **Configurações Gerais** e **Relatórios**.

1. Na exibição **Configurações Gerais**, defina a opção **Modo** para *Certificado*.
2. Na exibição **Relatórios**, selecione o tipo de relatório a ser gerado. É possível selecionar os seguintes tipos de relatório:

Opção	Descrição
Relatório AMAS	Contém informações que o Correio da Austrália exige para processar o conjunto de registros.
Relatório CASS	Contém informações que o USPS exige para processar o conjunto de registros.
Relatório SendRight	Contém informações que o Correio da Nova Zelândia exige para processar o conjunto de registros.
Relatório SERP	Contém informações que o Correio do Canadá exige para processar o conjunto de registros.

3. Para cada tipo de relatório selecionado, insira os detalhes do relatório.

Envie o arquivo do relatório para prestador de serviços postais com a lista de registros de endereço que você validou com a transformação de Validador de Endereço.

### TÓPICOS RELACIONADOS:

- [“Campos de Relatório AMAS” na página 127](#)
- [“Campos de Relatório CASS” na página 128](#)
- [“Relatório SendRight” na página 129](#)
- [“Campos de Relatório SERP” na página 129](#)

## Transformação de Validador de Endereço em um ambiente não nativo

O processamento da transformação de Validador de Endereço em um ambiente não nativo depende do mecanismo que executa a transformação.

Considere o suporte para os seguintes mecanismos de tempo de execução não nativos:

- Mecanismo Blaze. Suportado com restrições.
- Mecanismo Spark. Com suporte com restrições em mapeamentos de fluxo e em lote.



- Mecanismo Databricks Spark. Sem suporte.

## Transformação de Validador de Endereço no Mecanismo Blaze

A transformação de Validador de Endereço é suportada com as seguintes restrições:

- A transformação de Validador de Endereço não pode gerar um relatório de certificação.
- A transformação de Validador de Endereço falha na validação quando é configurada para ser executada no modo Interativo ou no modo Lista de Sugestões.

## Transformação de Validador de Endereço no Mecanismo Spark

A transformação de Validador de Endereço é suportada com as seguintes restrições:

- A transformação de Validador de Endereço não pode gerar um relatório de certificação.
- A transformação de Validador de Endereço falha na validação quando é configurada para ser executada no modo Interativo ou no modo Lista de Sugestões.

## Transformação de Validador de Endereço em um Mapeamento de Streaming

Os mapeamentos de fluxo têm as mesmas regras de processamento que os mapeamentos em lote no mecanismo Spark.

## CAPÍTULO 5

# Transformação de Agregador

Este capítulo inclui os seguintes tópicos:

- [Transformação de Agregador - Visão Geral, 134](#)
- [Transformações de Agregador em Mapeamentos Dinâmicos, 135](#)
- [Desenvolvendo uma Transformação de Agregador, 135](#)
- [Portas da Transformação de Agregador, 135](#)
- [Expressões Agregadas, 136](#)
- [Agrupar por Portas, 138](#)
- [Caches de Agregador, 141](#)
- [Entrada Classificada para uma Transformação de Agregador, 141](#)
- [Propriedades Avançadas da Transformação de Agregador, 143](#)
- [Criando uma Transformação de Agregador Reutilizável, 144](#)
- [Criando uma Transformação de Agregador Não Reutilizável, 144](#)
- [Dicas para Transformações de Agregador, 145](#)
- [Solucionando Problemas de Transformações de Agregador, 145](#)
- [Transformação de agregador em um ambiente não nativo, 146](#)

## Transformação de Agregador - Visão Geral

Configure uma transformação de Agregador para realizar cálculos agregados, como médias e somas, em grupos de dados. É possível usar uma transformação de Agregador para remover linhas duplicadas. A transformação de Agregador é uma transformação ativa.

A transformação de Agregador é diferente da transformação de Expressão, pois você pode configurar a transformação de Agregador para realizar cálculos em grupos de dados. Uma transformação de Expressão retorna resultados para cada linha.

Por exemplo, é possível calcular o salário médio para funcionários em cada departamento de uma organização. Configure um grupo por número de departamento. Configure uma expressão para calcular o salário médio e retornar o resultado para cada número de departamento exclusivo.

Use a linguagem de transformação para criar expressões agregadas.

O Serviço de Integração de Dados realiza cálculos agregados ao ler dados e armazena esses dados em um cache agregado. É possível classificar os dados de entrada para melhorar o desempenho. O Serviço de Integração de Dados não criará o cache se você classificar os dados de entrada.

# Transformações de Agregador em Mapeamentos Dinâmicos

É possível usar uma transformação de Agregador em um mapeamento dinâmico. Você pode configurar portas dinâmicas na transformação e fazer referência às portas geradas.

As seguintes tarefas podem ser realizadas para configurar uma transformação de Agregador em um mapeamento dinâmico:

## **Fazer referência a uma porta dinâmica ou a uma porta gerada como uma coluna Agrupar por.**

É possível incluir portas dinâmicas ou portas geradas como colunas Agrupar por. Se você especificar uma porta dinâmica como coluna Agrupar por, estará especificando que todas as portas geradas nessa porta dinâmica são colunas Agrupar por. Por isso, não será possível especificar uma porta gerada como coluna Agrupar por se você tiver especificado a porta dinâmica pai como uma coluna Agrupar por. Se você fizer referência a uma porta gerada e esta não existir em tempo de execução, o mapeamento falhará. É possível parametrizar as colunas Agrupar por para indicar quais delas devem ser agrupadas em tempo de execução.

## **Fazer referência a uma porta gerada na expressão agregada.**

É possível incluir uma porta gerada na expressão agregada. Se essa porta não existir em tempo de execução, o mapeamento falhará. Não é possível fazer referência a uma porta dinâmica na expressão agregada.

## **Criar a expressão agregada em uma porta de saída.**

Não é possível criar uma expressão agregada em uma porta gerada ou em uma porta dinâmica. A expressão agregada não pode ser uma expressão dinâmica.

## **Parametrizar valores na expressão agregada.**

É possível incluir parâmetros na expressão agregada. No entanto, você não pode usar um parâmetro de expressão ou um parâmetro de porta na expressão agregada.

# Desenvolvendo uma Transformação de Agregador

Ao criar uma transformação de Agregador, defina a expressão a ser executada em cada linha. Defina uma lista de portas Agrupar por para retornar resultados.

Para criar uma transformação de Agregador, realize as seguintes etapas:

1. Defina a transformação e crie as portas.
2. Configure uma expressão agregada em uma porta de saída ou variável.
3. Defina um grupo de portas para retornar resultados agregados.

# Portas da Transformação de Agregador

Uma transformação de Agregador tem vários tipos de portas que você pode usar para configurar grupos e agregar expressões.

A exibição Portas de uma transformação de Agregador tem os seguintes campos:

**Nome**

Nome da porta.

**Tipo**

O tipo de dados da porta.

**Precisão**

Comprimento do campo.

**Escala**

O número de posições à direita do decimal para dados numéricos.

**Entrada**

Indica que os dados são de uma transformação em upstream.

**Saída**

Indica que a porta retorna o valor de uma expressão. A expressão pode incluir expressões não agregadas e cláusulas condicionais. Você pode criar várias portas de saída de agregação.

**Passagem**

Indica que a porta é uma porta de entrada/saída que retorna os dados inalterados.

**Variável**

Indica que a porta pode armazenar valores ou cálculos para uso em uma expressão. Portas variáveis não podem ser portas de entrada ou de saída. Elas transmitem dados apenas no interior da transformação.

**Expressão**

A expressão para agregar linhas ou grupos de linhas.

**Valor padrão**

O valor padrão para uma porta com valores inválidos ou nulos.

**Regras de Entrada**

Um conjunto de regras que filtram as portas a serem incluídas ou excluídas na transformação com base em nomes de porta ou no tipo de dados. Configure regras de entrada ao definir portas dinâmicas.

## Expressões Agregadas

Configure expressões agregadas em portas variáveis ou nas portas de saída de uma transformação de Agregador. Não será possível inserir uma expressão se a porta for de entrada, de passagem ou dinâmica.

Uma expressão agregada pode incluir cláusulas condicionais e funções que não são agregadas. A função agregada também pode incluir uma função agregada aninhada dentro de outra função agregada, como:

```
MAX( COUNT( ITEM ) )
```

O resultado da expressão agregada varia com base nas portas de agrupamento da transformação. Por exemplo, a seguinte expressão agregada encontra a quantidade total de itens vendidos:

```
SUM( QUANTITY )
```

No entanto, se você usar a mesma expressão e agrupar pela porta ITEM, o Serviço de Integração de Dados retornará a quantidade total por item.

Você pode criar uma expressão agregada em qualquer porta de saída e usar várias portas de agregação em uma transformação.

## Funções de Agregação

Configure funções agregadas em uma transformação de Agregador. É possível aninhar uma função de agregação dentro de outra função de agregação.

A linguagem de transformação inclui as seguintes funções de agregação:

- AVG
- COUNT
- FIRST
- LAST
- MAX
- MEDIAN
- MIN
- PERCENTILE
- STDDEV
- SUM
- VARIANCE

Se você usar uma porta em uma expressão na transformação de Agregador, mas não usar a porta em uma função agregada, o Serviço de Integração de Dados usará a última linha na porta para processar a expressão.

Por exemplo, você cria uma transformação de Agregador que contém as portas `COMMISSIONS` e `SALARY`. A porta `SALARY` é uma porta de agrupamento.

Você pode usar a seguinte expressão em uma porta de saída:

```
SUM(COMMISSIONS)
```

O Serviço de Integração de Dados processa a função Agregador e retorna a soma dos valores na porta `COMMISSIONS` na porta de saída.

Você pode modificar a expressão para a seguinte expressão:

```
SUM(COMMISSIONS) + COMMISSIONS
```

Para processar a expressão, o Serviço de Integração de Dados retorna a soma dos valores na porta `COMMISSIONS` e adiciona o valor da última linha na porta `COMMISSIONS` para o valor de retorno na porta de saída.

Para uma porta de saída diferente, você pode usar a seguinte expressão:

```
SUM(COMMISSIONS) + SALARY
```

Para processar a expressão, o Serviço de Integração de Dados retorna a soma dos valores na porta `COMMISSIONS` e adiciona o valor da última linha na porta `SALARY` para o valor de retorno na porta de saída. Observe que os valores em cada linha da porta `SALARY` são os mesmos porque a porta `SALARY` é uma porta de agrupamento.

## Funções de Agregação Aninhadas

É possível incluir várias funções de nível único ou várias funções aninhadas em diferentes portas de saída em uma transformação de Agregador.

Não é possível incluir funções aninhadas e de nível único em uma transformação de Agregador. Portanto, se uma transformação de Agregador contiver uma função de nível único em qualquer porta de saída, não será possível usar uma função aninhada em nenhuma outra porta dessa transformação. Quando você inclui funções de nível único e aninhadas na mesma transformação de Agregador, a ferramenta Developer marca o mapeamento ou maplet como inválido. Se precisar criar funções aninhadas e de nível único, crie transformações de Agregador separadas.

## Cláusulas Condicionais em Expressões Agregadas

Use cláusulas condicionais na expressão agregada para reduzir o número de linhas usadas na agregação. A cláusula condicional pode ser qualquer cláusula que avalie como TRUE ou FALSE.

Por exemplo, use a seguinte expressão para calcular o total de comissões dos funcionários que excederam sua cota trimestral:

```
SUM( COMMISSION, COMMISSION > QUOTA )
```

## Agrupar por Portas

É possível definir grupos de linhas a serem agregadas em vez de executar uma agregação entre todos os dados de entrada. Por exemplo, você pode calcular as vendas totais da empresa ou pode localizar as vendas totais agrupadas por região.

Para definir um grupo para a expressão agregada, selecione as portas apropriadas de entrada, entrada/saída, saída e variáveis na transformação de Agregador. Você pode selecionar vários grupos por portas, a fim de criar um novo grupo para cada combinação exclusiva. Em seguida, o Serviço de Integração de Dados executa a agregação definida para cada grupo.

Quando você agrupa valores, o Serviço de Integração de Dados produz uma linha para cada grupo. Se você não agrupa os valores, o Serviço de Integração de Dados retorna uma linha para todas as linhas de entrada. O Serviço de Integração de Dados retorna a última linha de cada grupo com o resultado da agregação. Você pode especificar o retorno de uma linha específica. Por exemplo, se você usar a função agregadora FIRST, o Serviço de Integração de Dados retornará a primeira linha.

Quando várias portas de agrupamento são selecionadas na transformação de Agregador, o Serviço de Integração de Dados usa a ordem das portas para determinar a ordem do agrupamento. A ordem dos grupos pode afetar os resultados. Ordene as portas de agrupamento para garantir o agrupamento apropriado. Você pode alterar a ordem das portas depois de selecionar as portas no grupo /

Por exemplo, é possível criar uma porta de saída denominada Price\_Out. A expressão para Price\_Out é SUM (Qty \* Price). Você define Store\_ID e Item como as portas de agrupamento. A transformação retorna o preço total para cada item por loja.

As linhas de entrada podem conter os seguintes dados:

Store_ID	Item	Qty	Price
101	battery	3	2.99
101	battery	1	3.19
101	battery	2	2.59

Store_ID	Item	Qty	Price
101	AAA	2	2.45
201	battery	1	1.99
201	battery	4	1.59
301	battery	1	2.45

O Serviço de Integração de Dados executa o cálculo de agregação nos seguintes grupos exclusivos:

Store_Id	Item
101	battery
101	AAA
201	battery
301	battery

O Serviço de Integração de Dados retorna os valores de Store\_ID, Item, Qty e Price da última linha com a soma de (Price \* Qty) para cada item por loja:

Store_ID	Item	Qty	Price	Price_Out
101	battery	2	2.59	17.34
101	AAA	2	2.45	4.90
201	battery	4	1.59	8.35
301	battery	1	2.45	2.45

## Configurar Portas de Agrupamento

Defina as portas de agrupamento na guia **Agrupar por** da exibição **Propriedades** da transformação.

A seguinte imagem mostra a guia Agrupar por:

A guia Agrupar por contém as seguintes opções:

### Especificar por

Selecione **Valor** ou **Parâmetro**. Selecione **Valor** para usar nomes de portas. Escolha **Parâmetro** para usar um parâmetro de lista de portas.

### Adicionar

Aceita um nome de porta que você digita manualmente. Você deve digitar um nome de porta válido antes de clicar em **Adicionar**.

### Escolher

Clique em **Escolher** para selecionar portas a serem adicionadas ao grupo. A Developer tool fornece uma lista de portas de transformação para escolher.

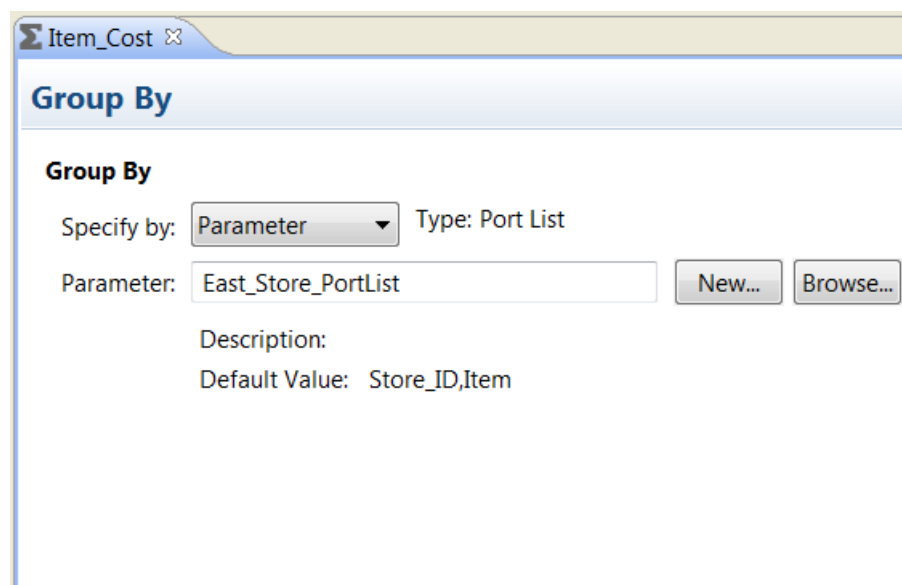
### Mover para Cima e Mover para Baixo

Você pode alterar a ordem das portas no grupo. Selecione o nome da porta e, em seguida, clique em um dos botões Mover para movê-lo para cima ou para baixo na ordem de classificação.

## Parâmetros de Agrupamento

É possível configurar um parâmetro de lista de portas que contém uma ou mais portas para inclusão no grupo. Crie um parâmetro de lista de portas selecionando portas em uma lista das portas na transformação.

A seguinte imagem mostra a guia **Agrupar por** quando você usa um parâmetro para identificar as portas do grupo:



É possível procurar um parâmetro de lista de portas ou clicar em **Novo** para criar esse parâmetro. Se você optar por criar um parâmetro de lista de portas, poderá selecionar as portas em uma lista das portas da transformação.

## Valores Padrão de Grupo por Portas

O Serviço de Integração de Dados não cria um grupo quando uma porta de agrupamento contém valores nulos. É possível definir um valor padrão para cada porta no grupo de forma a substituir qualquer valor de entrada nulo. Em seguida, o Serviço de Integração de Dados pode incluir as linhas nos totais de agregação.



## Expressões Não Agregadas

Use expressões não agregadas no grupo por portas para modificar ou substituir grupos.

Por exemplo, se você deseja substituir 'pilha AAA' antes de agrupar, pode criar um grupo por porta de saída, chamado CORRECTED\_ITEM, usando a seguinte expressão:

```
IIF( ITEM = 'AAA battery', battery, ITEM )
```

## Caches de Agregador

Quando você executa um mapeamento que usa uma transformação de Agregador, o Serviço de Integração de Dados cria um cache de índice e um cache de dados na memória para processar a transformação. Se o Serviço de Integração de Dados exigir mais espaço que o disponível no cache de memória, ele armazenará os dados excedentes nos arquivos de cache.

O Serviço de Integração de Dados cria os seguintes caches para a transformação de Agregador:

- Cache de índice que armazena valores de grupo como configurado no grupo por portas.
- Cache de dados que armazena cálculos com base no grupo por portas.

O Serviço de Integração de Dados não usa a memória em cache para executar uma transformação de Agregador com portas classificadas. Não é necessário configurar memória cache para transformações de Agregador que usam portas classificadas.

## Entrada Classificada para uma Transformação de Agregador

Você pode aumentar o desempenho da transformação de Agregador com a opção de entrada classificada.

Quando você usa a entrada classificada, o Serviço de Integração de Dados assume que todos os dados estejam classificados por grupo e realiza cálculos de agregação enquanto lê as linhas de um grupo. Quando necessário, o Serviço de Integração de Dados armazena as informações do grupo na memória. Para usar a opção Entrada Classificada, você deverá passar dados classificados à transformação do Agregador. Se você usar a entrada classificada, a transformação de Agregador fornecerá uma saída classificada.

Quando você não usa a entrada classificada, o Serviço de Integração de Dados realiza cálculos de agregação conforme lê. Como os dados não são classificados, o Serviço de Integração de Dados armazena os dados de cada grupo até ler toda a origem, para garantir que todos os cálculos de agregação sejam precisos.

Por exemplo, uma transformação de Agregador tem o grupo por portas STORE\_ID e ITEM, com a opção de entrada classificada selecionada. Quando você passar os seguintes dados por meio do Agregador, o Serviço de Integração de Dados executará uma agregação para as três linhas no grupo 101/pilha quando encontrar o grupo 201/pilha:

STORE_ID	ITEM	QTY	PRICE
101	'battery'	3	2.99
101	'battery'	1	3.19

STORE_ID	ITEM	QTY	PRICE
101	'battery'	2	2.59
201	'battery'	4	1.59
201	'battery'	1	1.99

Se você usar a entrada classificada e não classificar anteriormente os dados corretamente, o Serviço de Integração de Dados falhará na execução do mapeamento.

## Condições de Entrada Classificada

Determinadas condições podem impedir que você use a entrada classificada.

Você não poderá utilizar a entrada classificada se alguma das seguintes condições for verdadeira:

- A expressão de agregação contém funções de agregação aninhadas.
- Os dados de origem são controlados por dados.

Quando uma dessas condições for verdadeira, o Serviço de Integração de Dados processará a transformação como se você não estivesse usando a entrada classificada.

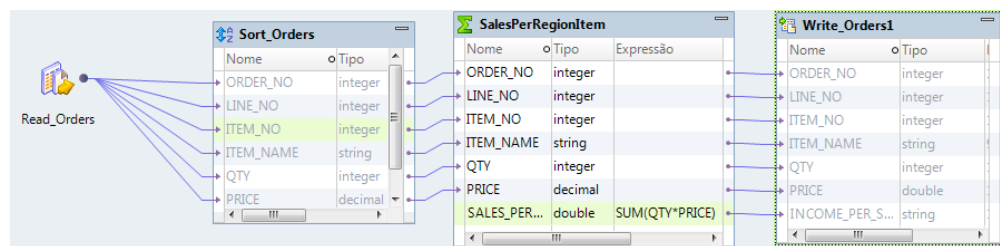
## Classificando Dados em uma Transformação de Agregador

Para usar a entrada classificada, passe dados classificados por meio da transformação de Agregador.

Você deve classificar dados pelo grupo por portas do Agregador na ordem em que eles aparecem na transformação de Agregador.

Para entrada relacional e de arquivo simples, use a transformação de Ordenador para classificar os dados no mapeamento antes de passá-los para a transformação de Agregador. Você poderá colocar a transformação de Ordenador em qualquer lugar do mapeamento antes do Agregador se nenhuma transformação alterar a ordem dos dados classificados. O grupo por colunas na transformação do Agregador deve estar na mesma ordem em que aparece na transformação Classificador.

O mapeamento a seguir mostra uma transformação Classificador, configurada para classificar os dados de origem ascendente por ITEM\_NO:



A transformação classifica os dados como a seguir:

ITEM_NO	ITEM_NAME	QTY	PRICE
345	Soup	4	2.95
345	Soup	1	2.95
345	Soup	2	3.25

ITEM_NO	ITEM_NAME	QTY	PRICE
546	Cereal	1	4.49
546	Cereal	2	5.25

Com a entrada classificada, a transformação do Agregador retorna estes resultados:

ITEM_NAME	QTY	PRICE	INCOME_PER_ITEM
Cereal	2	5.25	14.99
Soup	2	3.25	21.25

## Propriedades Avançadas da Transformação de Agregador

Configure propriedades que ajudam a determinar como o Serviço de Integração de Dados processa dados para a transformação de Agregador.

Configure as seguintes propriedades avançadas para uma transformação de Agregador:

### Diretório de Cache

O diretório onde o Serviço de Integração de Dados cria os arquivos de cache de índice e de dados. Verifique se o diretório existe e se contém espaço em disco suficiente para os arquivos de cache.

Insira vários diretórios separados por ponto-e-vírgula para melhorar o desempenho durante o particionamento do cache. O particionamento do cache cria um cache separado para cada partição que processa a transformação.

O padrão é o parâmetro do sistema CacheDir. Você pode configurar outro parâmetro do sistema ou parâmetro definido pelo usuário para essa propriedade.

### Tamanho do Cache de Dados

Quantidade de memória que o Serviço de Integração de Dados aloca para o cache de dados para a transformação no início da execução do mapeamento. Selecione Automático para que o Serviço de Integração de Dados calcule automaticamente os requisitos de memória no tempo de execução. Insira um valor específico em bytes quando você ajustar o tamanho do cache. O padrão é Automático.

### Tamanho do Cache de Índice

Quantidade de memória que o Serviço de Integração de Dados aloca para o cache de índice para a transformação no início da execução do mapeamento. Selecione Automático para que o Serviço de Integração de Dados calcule automaticamente os requisitos de memória no tempo de execução. Insira um valor específico em bytes quando você ajustar o tamanho do cache. O padrão é Automático.

### Entrada Classificada

Indica que os dados de entrada são pré-classificados por grupos. Selecione essa opção somente se o mapeamento transferir dados classificados para a transformação de Agregador.

### Nível de Rastreamento

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.

### TÓPICOS RELACIONADOS:

- [“Tamanho do Cache” na página 72](#)

## Criando uma Transformação de Agregador Reutilizável

Crie uma transformação de Agregador reutilizável para usar em vários mapeamentos ou mapplets.

1. Selecione um projeto ou pasta na exibição **Object Explorer**.
2. Clique em **Arquivo > Novo > Transformação**.  
A caixa de diálogo **Novo** é exibida.
3. Selecione a transformação de Agregador.
4. Clique em **Avançar**.
5. Insira um nome para a transformação.
6. Clique em **Concluir**.  
A transformação aparece no editor.
7. Clique no botão **Novo** para adicionar uma porta à transformação.
8. Edite a porta para definir o nome, o tipo de dados e a precisão.
9. Determine o tipo de cada porta: entrada, saída, passagem ou variável.
10. Clique no campo Expressão para configurar expressões agregadas para portas de saída. Você pode selecionar as portas e os parâmetros para definir a expressão agregada.
11. Clique na exibição **Avançado** e edite as propriedades da transformação.

## Criando uma Transformação de Agregador Não Reutilizável

Crie uma transformação de Agregador não reutilizável em um mapeamento ou mapplet.

1. Em um mapeamento ou mapplet, arraste uma transformação de Agregador da paleta de Transformação para o editor.  
A transformação aparece no editor.
2. Na exibição **Propriedades**, edite o nome da transformação e a descrição.
3. Na guia **Portas**, clique no botão **Novo** para adicionar portas à transformação.
4. Edite as portas para definir o nome, o tipo de dados e a precisão.
5. Determine o tipo de cada porta: entrada, saída, passagem ou variável.

6. Configure expressões agregadas para as portas de saída.
7. Na exibição **Avançado**, edite as propriedades da transformação.

## Dicas para Transformações de Agregador

Você pode seguir dicas para usar as transformações de Agregador com mais eficácia.

### **Utilize a entrada classificada para diminuir o uso de caches de agregação.**

A entrada classificada reduz a quantidade de dados armazenados no cache durante a execução do mapeamento e melhora o desempenho. Utilize esta opção com a transformação de Ordenador para passar dados classificados para a transformação de Agregador.

### **Limite as portas de entrada/saída ou de saída conectadas.**

Limite o número de portas de entrada/saída ou de saída conectadas para reduzir o volume de dados que a transformação de Agregador armazena no cache de dados.

### **Filtre os dados antes de agregá-los.**

Se utilizar uma transformação de Filtro no mapeamento, coloque-a antes da transformação de Agregador, a fim de reduzir uma agregação desnecessária.

### **Somente uma transformação de Agregador classificada fornece saídas classificadas.**

Se você usar a entrada não classificada e quiser produzir uma saída classificada, use uma transformação de Classificador após a transformação de Agregador.

## Solucionando Problemas de Transformações de Agregador

Você pode solucionar problemas de transformações de Agregador.

### **Selecionei a entrada classificada, mas o mapeamento leva o mesmo tempo que antes.**

Você não poderá utilizar a entrada classificada se alguma das seguintes condições for verdadeira:

- A expressão de agregação contém funções de agregação aninhadas.
- Os dados de origem são controlados por dados.

Quando uma dessas condições for verdadeira, o Serviço de Integração de Dados processará a transformação como se você não estivesse usando a entrada classificada.

### **Um mapeamento com uma transformação de Agregador provoca baixo desempenho.**

O Serviço de Integração de Dados pode estar paginando no disco. É possível melhorar o desempenho aumentando os tamanhos dos caches de dados e de índice nas propriedades da transformação.

# Transformação de agregador em um ambiente não nativo

O processamento da transformação de Agregador em um ambiente não nativo depende do mecanismo que executa a transformação.

Considere o suporte para os seguintes mecanismos de tempo de execução não nativos:

- Mecanismo Blaze. Suportado com restrições.
- Mecanismo Spark. Com suporte com restrições em mapeamentos de fluxo e em lote.
- Mecanismo Databricks Spark. Suportado com restrições.

## Transformação de agregador no mecanismo Blaze

Algumas regras de processamento para o mecanismo Blaze são diferentes das regras de processamento do Serviço de Integração de Dados.

### Validação de Mapeamento

A validação do mapeamento falha nas seguintes situações:

- A transformação contém portas de variável com estado.
- A transformação contém funções não suportadas em uma expressão.

### Funções de agregação

Se você usar uma porta em uma expressão na transformação de Agregador, mas não usar a porta em uma função agregada, o mecanismo de tempo de execução poderá usar qualquer linha na porta para processar a expressão.

A linha que o mecanismo de tempo de execução usa pode não ser a última linha na porta. O processamento é distribuído e, portanto, o mecanismo de tempo de execução pode não conseguir determinar a última linha na porta.

### Otimização do cache de dados

O cache de dados da transformação de Agregador é otimizado para usar o comprimento variável para armazenar tipos de dados binary e string que passam pela transformação de Agregador. A otimização está ativada para tamanhos de registros de até 8 MB. Se o tamanho do registro for maior que 8 MB, a otimização de tamanho variável será desativada.

Quando o comprimento de variável é usado para armazenar dados que passam pela transformação de Agregador no cache de dados, a transformação de Agregador é otimizada para usar entrada classificada e uma transformação de Classificador de passagem é inserida antes da transformação de Agregador no mapeamento de tempo de execução.

Para visualizar a transformação de Classificador, visualize o mapeamento otimizado ou visualize o plano de execução no ambiente de validação do Blaze.

Durante a otimização do cache de dados, o cache de dados e o cache de índice da transformação de Agregador são definidos como Automático. O cache de classificação da transformação de Classificador é configurado para o mesmo tamanho que o cache de dados para a transformação de Agregador. Para configurar o cache de classificação, você deve configurar o tamanho do cache de dados para a transformação de Agregador.

## Transformação de agregador no mecanismo Spark

Algumas regras de processamento para o mecanismo Spark são diferentes das regras de processamento do Serviço de Integração de Dados.

### Validação de Mapeamento

A validação do mapeamento falha nas seguintes situações:

- A transformação contém portas de variável com estado.
- A transformação contém funções não suportadas em uma expressão.

### Funções de agregação

Se você usar uma porta em uma expressão na transformação de Agregador, mas não usar a porta em uma função agregada, o mecanismo de tempo de execução poderá usar qualquer linha na porta para processar a expressão.

A linha que o mecanismo de tempo de execução usa pode não ser a última linha na porta. O processamento é distribuído e, portanto, o mecanismo de tempo de execução pode não conseguir determinar a última linha na porta.

### Otimização do cache de dados

Você não pode otimizar o cache de dados para a transformação para armazenar dados usando o tamanho da variável.

## Transformação de Agregador em um mapeamento de fluxo

Os mapeamentos de fluxo têm regras de processamento adicionais que não se aplicam a mapeamentos em lote.

### Validação de Mapeamento

A validação do mapeamento falha nas seguintes situações:

- Um pipeline de fluxo contém mais de uma transformação de Agregador.
- Um pipeline de fluxo contém uma transformação de Agregador e uma transformação de Classificação.
- Uma transformação de Agregador está upstream de uma transformação de Pesquisa.
- Uma transformação de Agregador está no mesmo pipeline de fluxo que uma transformação de Pesquisa passiva configurada com uma condição de pesquisa de desigualdade.

## Transformação de agregador no mecanismo Databricks Spark

Algumas regras de processamento para o mecanismo Databricks Spark são diferentes das regras de processamento do Serviço de Integração de Dados.

### Validação de Mapeamento

A validação do mapeamento falha nas seguintes situações:

- A transformação contém portas de variável com estado.
- A transformação contém funções não suportadas em uma expressão.

## Funções de agregação

Se você usar uma porta em uma expressão na transformação de Agregador, mas não usar a porta em uma função agregada, o mecanismo de tempo de execução poderá usar qualquer linha na porta para processar a expressão.

A linha que o mecanismo de tempo de execução usa pode não ser a última linha na porta. O processamento é distribuído e, portanto, o mecanismo de tempo de execução pode não conseguir determinar a última linha na porta.

## Otimização do cache de dados

Você não pode otimizar o cache de dados para a transformação para armazenar dados usando o tamanho da variável.



## CAPÍTULO 6

# Transformação de Associação

Este capítulo inclui os seguintes tópicos:

- [Transformação de Associação - Visão Geral, 149](#)
- [Alocação de Memória, 150](#)
- [Propriedades Avançadas da Transformação de Associação, 151](#)

## Transformação de Associação - Visão Geral

A transformação de Associação processa dados de saída de uma transformação de Correspondência. Ela cria vínculos entre registros duplicados que estão atribuídos a diferentes clusters de correspondência, de forma que esses registros possam ser associados juntos em operações de consolidação de dados e de gerenciamento de dados mestres.

A transformação de Associação gera um valor **AssociationID** para cada linha em um grupo de registros associados e grava os valores de ID em uma porta de saída.

A transformação de Consolidação lê a saída da transformação de Associação. Use uma transformação de Consolidação para criar um registro mestre com base nos registros com valores comuns de ID de associação.

A transformação de Associação aceita valores de dados numéricos e de string nas portas de entrada. Se você adicionar uma porta de entrada de outro tipo de dados, a transformação converterá os valores de dados da porta em strings.

A porta de saída AssociationID grava dados de inteiros. A transformação pode gravar dados de string em uma porta AssociationID se a transformação foi configurada em uma versão anterior do Informatica Data Quality.

### Exemplo: Associando Saídas de Transformação de Correspondência

A tabela a seguir contém três registros que poderiam identificar o mesmo indivíduo:

ID	Nome	Endereço	Cidade	Estado	Código Postal	SSN
1	David Jones	100 Admiral Ave.	Nova York	NY	10547	987-65-4321
2	Dennis Jones	1000 Alberta Ave.	New Jersey	NY	-	987-65-4321
3	D. Jones	Admiral Ave.	Nova York	NY	10547-1521	-

Uma operação de análise de duplicata definida em uma transformação de Correspondência não identifica os três registros como duplicatas uns dos outros pelos seguintes motivos:

- Se você definir uma pesquisa de duplicata por dados de nome e endereço, os registros 1 e 3 serão identificados como duplicatas, mas o registro 2 será omitido.
- Se você definir uma pesquisa de duplicata por dados de nome e número de Seguridade Social, os registros 1 e 2 serão identificados como duplicatas, mas o registro 3 será omitido.
- Se você definir uma pesquisa de duplicata por todos os três atributos (nome, endereço e número de Seguridade Social), a transformação de Correspondência poderá não identificar nenhum dos registros como correspondências.

A transformação de Associação vincula dados de diferentes clusters de correspondência, para que os registros que compartilham um ID de cluster recebam um valor de AssociationID comum. Neste exemplo, todos os três registros recebem o mesmo AssociationID, conforme mostrado na seguinte tabela:

ID	Nome	Endereço	Cidade	Estado	Código Postal	SSN	ID de Cluster de Nome e Endereço	ID de Cluster de Nome e SSN	ID da Associação
1	David Jones	100 Admiral Ave.	Nova York	NY	10547	987-65-4320	1	1	1
2	Dennis Jones	1000 Alberta Ave.	New Jersey	NY	-	987-65-4320	2	1	1
3	D. Jones	Alberta Ave.	Nova York	NY	10547-1521	-	1	2	1

Você pode consolidar os dados de registro duplicados na transformação de Consolidação.

## Alocação de Memória

É possível definir a quantidade mínima de memória cache utilizada pela transformação de Associação. A configuração padrão é 400.000 bytes.

Defina o valor na propriedade **Tamanho de Arquivo de Cache** na guia **Avançado**.

O valor padrão representa a quantidade mínima de memória usada pela transformação. A transformação de Associação tentará obter um múltiplo do valor padrão, dependendo do número de portas no qual você associar. A transformação usa esta fórmula para obter memória cache:

$(\text{Número de portas de associação} + 1) \times \text{memória cache padrão}$

Por exemplo, se você configurar sete portas de associação, a transformação tentará alocar 3,2 milhões de bytes, ou 3,05 MB, para a memória cache.

Se você alterar a configuração padrão, a transformação não tentará obter memória adicional.

**Nota:** Se você inserir um valor de memória cache menor que 65536, a transformação de Associação lerá o valor em megabytes.

# Propriedades Avançadas da Transformação de Associação

A transformação de Associação contém propriedades avançadas que determinam o comportamento da memória cache e o nível de rastreamento.

Você pode configurar as seguintes propriedades avançadas:

## **Diretório de Arquivo de Cache**

Especifica o diretório no qual o Serviço de Integração de Dados grava os dados temporários da transformação atual. O Serviço de Integração de Dados grava os arquivos temporários no diretório quando o volume de dados de entrada é maior do que a memória disponível do sistema. O Serviço de Integração de Dados exclui os arquivos temporários depois que o mapeamento é executado.

Você pode inserir um caminho de diretório na propriedade ou usar um parâmetro para identificar o diretório. Especifique um caminho local na máquina do Serviço de Integração de Dados. O Serviço de Integração de Dados deve ser capaz de gravar no diretório. O valor padrão é o parâmetro do sistema CacheDir.

## **Tamanho do Arquivo de Cache**

Determina a quantidade de memória do sistema que o Serviço de Integração de Dados usa para classificar os dados de entrada na transformação. O valor padrão é 400.000 bytes. Você pode usar um parâmetro para especificar o tamanho do arquivo de cache.

Antes de classificar os dados, o Serviço de Integração de Dados aloca a quantidade de memória especificada por você. Se a operação de classificação gerar uma quantidade maior de dados, o Serviço de Integração de Dados gravará o excesso de dados no diretório do arquivo de cache. Se a operação de classificação exigir mais memória do que a memória do sistema e o armazenamento de arquivos podem fornecer, o mapeamento falhará.

**Nota:** Se você inserir um valor igual ou maior que 65536, a transformação lerá o valor em bytes. Se você inserir um valor menor, a transformação lerá o valor em megabytes.

## **Nível de Rastreamento**

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.

## CAPÍTULO 7

# Transformação de Exceção de Registro Inválido

Este capítulo inclui os seguintes tópicos:

- [Visão Geral da Transformação de Exceção de Registro Inválido, 152](#)
- [Tipos de Registro de Saída de Exceção de Registro Inválido, 153](#)
- [Fluxo do Processo de Gerenciamento de Exceção de Registro Inválido, 154](#)
- [Mapeamentos de Exceção de Registro Inválido, 154](#)
- [Portas de Exceção de Registro Inválido , 157](#)
- [Exibição Configuração de Exceção de Registro Inválido, 158](#)
- [Atribuição de Problema de Exceção de Registro Inválido , 160](#)
- [Propriedades Avançadas da Transformação de Exceção, 161](#)
- [Configurando uma Transformação de Exceção de Registros Inválidos, 162](#)
- [Exemplo de Mapeamento de Exceção de Registro Inválido, 163](#)

## Visão Geral da Transformação de Exceção de Registro Inválido

A transformação de Exceção de Registro Inválido é uma transformação ativa que lê a saída de um processo de qualidade de dados e identifica registros que exigem revisão manual. A transformação de Exceção de Registro Inválido é uma transformação de vários grupos.

Configure uma transformação de Exceção de Registro Inválido para analisar a saída de um processo que identifica problemas de qualidade de dados em registros. Um registro com um problema de qualidade de dados que precisa de mais análise é uma exceção.

A transformação de Exceção de Registro Inválido recebe entrada de outra transformação ou de um objeto de dados em outro mapeamento. A entrada da transformação de Registro Inválido deve conter uma ou mais portas de problema de qualidade que recebe descrições de texto de problemas de qualidade de dados. A entrada para a transformação de Exceção de Registro Inválido também pode conter uma pontuação numérica de registro que a transformação pode usar para determinar a qualidade dos dados de cada registro. Defina um limite de pontuação superior e inferior na transformação de Exceção para classificar os registros como de qualidade válida e inválida com base na pontuação do registro. A transformação de Exceção de Registro Inválido grava exceções e texto de problema de qualidade associado a uma tabela de Registros Inválidos.

Por exemplo, uma organização precisa validar endereços de clientes antes de enviar correspondências para os clientes. Um desenvolvedor cria um mapeamento que valida a cidade, o estado e o código postal do cliente em tabelas de referência com uma transformação de Rotulador. A transformação de Rotulador valida os campos e adiciona uma pontuação de registro para cada linha com base nos resultados. A transformação de Rotulador também adiciona o texto que descreve os problemas de qualidade para cada registro que tem um erro. A transformação de Rotulador adiciona texto de problema de qualidade, como `cidade não válida` ou `Código de CEP em branco`, para cada exceção. A transformação de Exceção de Registro Inválido grava registros de cliente que precisam de revisão manual para a tabela de Registros Inválidos. Os analistas de dados analisam e corrigem os registros inválidos na ferramenta Analyst.

## Tipos de Registro de Saída de Exceção de Registro Inválido

A Exceção de Registro Inválido examina as pontuações do registro de entrada para determinar a qualidade do registro. Ele retorna os registros para diferentes grupos de saída

A transformação de Exceção identifica os seguintes tipos de registros com base em cada registro de pontuação:

### Bons registros

Registros com pontuações maior ou igual ao limite superior. Bons registros são válidos e não exigem análise. Por exemplo, se configurar o limite superior como 90, qualquer registro com limite 90 ou superior não precisará de revisão.

### Registros inválidos

Os registros com pontuações inferiores ao limite superior e pontuações superiores ou iguais ao limite inferior. Registros inválidos são exceções que você precisa examinar na ferramenta Analyst. Por exemplo, quando o limite inferior é 40, qualquer registro com uma pontuação de 40 até 90 precisa de revisão manual.

### Registros rejeitados

Registros com pontuações menores que o limite inferior. Registros rejeitados não são válidos. Por padrão, a transformação de Exceção descarta registros rejeitados do fluxo de dados. Por exemplo, qualquer registro com uma pontuação de 40 ou menos é um registro rejeitado.

**Nota:** Se os campos de problemas de qualidade são NULL, o registro não é uma exceção. Quando qualquer problema de qualidade contém texto ou contém uma string vazia, o registro é uma exceção. Verifique se uma posta de problema de qualidade contém valores nulos quando um campo não tiver erros. Se as portas de problema de qualidade contém espaços em branco em vez de valores nulos, a transformação de Exceção sinaliza a cada registro como uma exceção. Quando um usuário precisar corrigir os problemas na ferramenta Analyst, o usuário não poderá filtrar exceções por problema de qualidade de dados.

Quando um registro tem uma pontuação menor que zero ou maior que 100, a linha não é válida. O Serviço de Integração de Dados registra uma mensagem de erro de que a linha não é válida e ignorará o processamento do registro.

Se você não conectar um registro de pontuação como entrada à transformação de Exceção, a transformação gravará todos os registros que contiverem problemas de qualidade na tabela de registros inválidos.

Quando você incluir a transformação de Exceção de Registro Inválido em uma tarefa de Mapeamento, poderá configurar uma tarefa Humana no mesmo fluxo de trabalho para incluir uma revisão manual de exceções. A tarefa Humana começa quando uma tarefa de Mapeamento é encerrada no fluxo de trabalho. A tarefa Humana exige que usuários acessem a ferramenta Analyst para resolver problemas de qualidade. Um

usuário pode atualizar os dados e alterar o status de qualidade de cada registro na tabela de registros inválidos.

## Fluxo do Processo de Gerenciamento de Exceção de Registro Inválido

A transformação de Exceção recebe pontuações do registro de transformações de qualidade de dados e cria as tabelas que contêm registros com níveis diferentes de qualidade de dados. Você deve configurar as transformações de qualidade de dados para encontrar problemas de qualidade e fornecer uma pontuação do registro para cada linha.

Você pode configurar as transformações de qualidade de dados em um único mapeamento ou pode criar mapeamentos para diferentes estágios do processo de qualidade de dados.

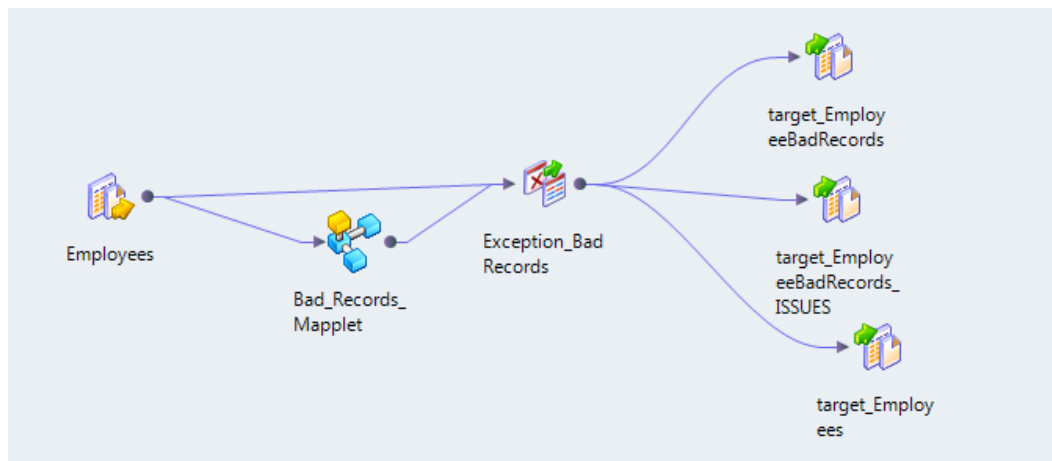
Conclua as seguintes tarefas de gerenciamento de exceção de registro inválido:

1. Na ferramenta Developer, defina as transformações que geram valores de pontuação para dados de origem com base nos problemas de qualidade de dados que você definir. Defina as transformações que retornam texto para descrever a qualidade dos dados de origem. Por exemplo, você pode configurar uma transformação de Rotulador para verificar os dados de origem com relação às tabelas de referência e depois gravar um rótulo descritivo para cada comparação. Você pode definir uma regra IF/THEN em uma transformação de Decisão para examinar um campo de dados. Você pode definir várias transformações e mapplets que executam diferentes operações de qualidade de dados.
2. Configure uma transformação de Exceção para analisar as pontuações do registro que ele recebe das operações de qualidade de dados. Configure a transformação para gravar registros para tabelas de banco de dados com base em valores de pontuação nos registros. Você pode criar tabelas separadas para registros válidos, registros inválidos, problemas de qualidade e registros rejeitados.
3. Atribua uma porta de problema de qualidade a cada porta de entrada que pode conter dados inválidos.
4. Opcionalmente, configure objetos de dados de destino para registros válidos e inválidos. Conecte as portas de saída de transformação de Exceção aos objetos de dados de destino no mapeamento.
5. Crie o objeto de dados de destino para registros inválidos. Escolha para gerar uma tabela de registros inválidos e adicione-a ao mapeamento. Quando você gerar uma tabela de registros inválidos, a ferramenta Developer também gerará uma tabela de problemas de qualidade de dados. Adicione a tabela de problemas de qualidade ao mapeamento.
6. Adicione o mapeamento para um fluxo de trabalho.
7. Configure uma tarefa Humana para atribuir revisão manual de registros inválidos aos usuários. Os usuários podem analisar e atualizar os registros inválidos na ferramenta Analyst.

## Mapeamentos de Exceção de Registro Inválido

Ao criar um mapeamento que identifica exceções de registros inválidos, você configura o mapeamento para gravar registros em um ou mais destinos de banco de dados, com base na qualidade dos dados nos registros.

A figura a seguir mostra um exemplo de mapeamento de Exceção de Registro Inválido:



O mapeamento contém os seguintes objetos:

#### Fonte de dados

Uma fonte de dados de Funcionários que contém os registros para analisar a Qualidade de Dados.

#### Maplet

O Bad\_Records\_Maplet contém transformações que verificam e adicionam problemas de qualidade e pontuações de registro aos registros de origem. As regras são transformações que analisam os dados e encontram problemas de qualidade. Por exemplo, você pode incluir uma transformação de Rotulador para comparar dados de entrada às tabelas de referência. Dependendo dos resultados, você pode configurar a transformação de Rotulador para retornar problemas de qualidade como colunas adicionais nas linhas. Você pode configurar uma transformação de Decisão que usa instruções IF, THEN, ELSE para examinar os dados e aplicar problemas de qualidade e pontuações do registro aos dados de entrada.

#### Transformação de Exceção

A transformação de Exceção determina quais registros gravar nos destinos de dados, incluindo a tabela de registros inválidos e a tabela de problemas.

#### Tabela de registros válidos

A transformação de Exceção grava todos os registros de boa qualidade na tabela target\_Employees.

#### Tabela de Registros Inválidos

A transformação de Exceção grava todos os registros de má qualidade na tabela target\_EmployeeBadRecords. Os registros inválidos exigem revisão manual.

#### Tabela de problemas

A transformação de Exceção grava os problemas de qualidade na tabela target\_EmployeeBadRecords\_ISSUES. Quando você exibe os registros inválidos na ferramenta Analyst, a interface do usuário vincula os problemas de qualidade aos registros inválidos.

Opcionalmente, a transformação de Exceção pode gravar registros rejeitados em uma tabela de registros rejeitados. Escolha criar um grupo de saída separado para registros rejeitados na exibição **Configuração** da transformação.

## Problemas de Qualidade de Exceção de Registro Inválido

Problemas de qualidade são strings de texto que descrevem o tipo de problema de qualidade de dados que causou uma baixa pontuação do registro. A transformação de Exceção de Registro Inválido recebe

problemas de qualidade associados a cada linha de origem que contém uma baixa pontuação do registro. Você pode configurar diferentes tipos de transformações que determinam problemas de qualidade e pontuações do registro.

Por exemplo, você pode criar uma transformação de Decisão que examina o número de telefone. A transformação de Decisão gera a pontuação do registro e os problemas de qualidade de número de telefone.

A seguinte estratégia de decisão identifica números de telefone de tamanho incorreto em uma transformação de Decisão:

```
IF LENGTH(Phone_Number) > 10 THEN
    Score:=50
    Phone_Quality_Issue:='Phone num too long'
ELSEIF LENGTH(Phone_Number) < 10 THEN
    Score:=50
    Phone_Quality_Issue:=' Phone num too short'
ELSE
    Score:=90
ENDIF
```

Quando você configurar a transformação de Exceção, deverá associar Phone\_Quality\_Issue à porta Phone\_Number. As portas são grupos de entrada diferentes.

A transformação de Exceção lê as pontuações geradas pela transformação de Decisão e atribui registros com uma pontuação de "50" ao grupo de registros inválidos das portas de saída. Ele grava Phone\_Quality\_Issue ao grupo de Problemas das portas de saída.

## Tarefas Humanas

Quando você configurar um fluxo de trabalho que contém uma transformação de Exceção, inclua o mapeamento em uma tarefa de Mapeamento. Adicione uma tarefa Humana ao mesmo fluxo de trabalho. A tarefa Humana exige que um ou mais usuários corrijam os registros de exceção na ferramenta Analyst.

A tarefa de Mapeamento identifica os registros na fonte de dados que contém problemas de qualidade de dados não resolvidos. Os analistas de dados usam a ferramenta Analyst para solucionar os problemas e para atualizar o status de qualidade de dados de cada registro.

Quando você configura uma tarefa Humana, cria uma ou mais instâncias de tarefa e uma ou mais etapas de tarefa. Uma instância de tarefa representa o conjunto de dados no qual um usuário deve trabalhar. Uma etapa de tarefa representa o tipo de trabalho que um usuário realiza nos registros da respectiva instância de tarefa. Você pode criar várias instâncias de tarefa para que diferentes usuários trabalhem em diferentes partes dos dados na ferramenta Analyst.

Um usuário pode atualizar o status dos registros inválidos na ferramenta Analyst de uma das seguintes maneiras:

- Se um registro for válido, o usuário atualizará os metadados da tabela para confirmar o registro para armazenamento persistente no banco de dados.
- Se um registro não for válido, o usuário atualizará os metadados da tabela de forma que o registro seja removido do banco de dados em um estágio posterior do fluxo de trabalho.
- Se o status do registro não for confirmado, o usuário atualizará os metadados da tabela de forma que o registro retorne para o fluxo de trabalho para processamento adicional em uma tarefa de Mapeamento.

Para obter mais informações sobre tarefas Humanas, consulte o *Guia de Fluxo de Trabalho do Informatica Developer*.



# Portas de Exceção de Registro Inválido

Configure as portas de entrada e saída na guia **Portas** da transformação de Exceção de Registro Inválido.

A transformação de Exceção de Registro Inválido contém grupos de portas de entrada e saída.

A figura a seguir mostra a guia **Portas**:

	Nome	Tipo	Precisão	Escala	Entrada	Saída	Padrão	Descrição
<b>Entradas</b>								
<b>Dados (3)</b>								
1	EmployeeID	decimal	10	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
2	Employee_Name	decimal	10	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
3	Phone_Nuber	string	10	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
<b>Problemas de Qualidade (0)</b>								
<b>Controle (1)</b>								
1	Pontuação	double	15	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
<b>Saídas</b>								
<b>Saída Padrão (4)</b>								
1	Pontuação	double	15	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>		
2	EmployeeID	decimal	10	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>		
3	Employee_Name	decimal	10	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>		
4	Phone_Nuber	string	10	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>		
<b>Registros Inválidos (6)</b>								
1	Workflow_ID	string	64	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>		
2	Row_Identifier	bigint	19	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>		
3	Record_Status	string	20	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>		
4	EmployeeID	decimal	10	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>		
5	Employee_Name	decimal	10	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>		
6	Phone_Nuber	string	10	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>		

## Portas de Entrada de Transformação de Exceção de Registro Inválido

Uma transformação de Exceção de Registro Inválido tem grupos de entrada separados para os dados, os problemas de qualidade e a pontuação do registro.

A transformação de Exceção de Registro Inválido tem os seguintes grupos de entrada:

### Dados

Os campos de dados de origem.

### Problemas de Qualidade

Contém portas que descrevem problemas de qualidade do registro. As portas de problema de qualidade contêm strings, como "Excess\_Characters" ou "Bad\_Data\_Format". Você pode ter vários problemas de qualidade em cada registro. A transformação não associa portas no grupo de Problemas de Qualidade a campos de dados de origem até que você atribua os problemas às portas de dados na exibição

**Atribuição de Problema.**

### Controle

A pontuação do registro. A transformação de Exceção analisa a pontuação do registro para determinar se as linhas de entrada são exceções. Se você não se conectar à porta de Pontuação, a transformação de Exceção identificará uma linha como uma exceção, caso uma porta de Problemas de Qualidade contenha dados.

## Saída de Transformação de Exceção de Registro Inválido

Uma transformação de Exceção de Registro Inválido tem vários grupos de saída.

A transformação de Exceção de Registro Inválido tem os seguintes grupos de saída:

### Saída Padrão

Os registros de boa qualidade que você não precisa examinar em busca de problemas de qualidade de dados.

Cada registro no grupo de Saída Padrão contém uma porta de Pontuação que representa a qualidade de dados do registro.

### Registros Inválidos

As exceções que você precisa examinar em busca de problemas de qualidade de dados.

Cada registro no grupo de Registros Inválidos contém um ID de fluxo de trabalho, um identificador de linha e uma porta de status de registro.

### Problemas

Os problemas de qualidade de registros no grupo de Registros Inválidos. Os problemas de qualidade são elementos de metadados que a ferramenta Analyst exibe quando você analisa os registros inválidos.

Cada registro no grupo Problemas contém um ID de fluxo de trabalho e uma porta de identificador de linha que identifica a qual linha do registro inválido os problemas pertencem.

### Registros Rejeitados

Grupo opcional que contém registros que você pode remover do banco de dados. Cada registro no grupo de Registros Rejeitados contém uma pontuação de registro baixa na porta de Pontuação.

## Exibição Configuração de Exceção de Registro Inválido

A exibição **Configuração** especifica os limites superiores e inferiores usados pela transformação para identificar os registros válidos e inválidos. A exibição **Configuração** também identifica as tabelas de destino dos registros com pontuações que estão acima ou abaixo dos limites.

A figura a seguir mostra a exibição **Configuração** da transformação de Exceção:

Você pode configurar as seguintes propriedades na exibição **Configuração**:

O limite inferior para o intervalo de pontuação de registros inválidos. A transformação identifica os registros com pontuações abaixo do limite inferior como registros rejeitados.

O limite superior para o intervalo de pontuação de registros inválidos. A transformação identifica registros com pontuações maiores ou iguais ao limite superior como registros válidos.

Os tipos de registros de saída. Na configuração padrão, a transformação grava os registros válidos na Saída Padrão e grava os registros inválidos na tabela de Registros Inválidos. Por padrão, a transformação não grava os registros rejeitados em uma tabela de banco de dados.

Os tipos de registro que a transformação grava nas portas de saída padrão. O padrão é Bons Registros.

Os tipos de registro que a transformação grava nas portas de saída de registros inválidos. O padrão é Registros Inválidos.

Cria um grupo de saída separado para os registros rejeitados. A opção está desmarcada por padrão.

Cria uma tabela de banco de dados para conter os dados de registro inválido. Quando você seleciona a opção, a transformação de Exceção cria a tabela de banco de dados, adiciona um objeto de dados ao repositório do Modelo e adiciona uma instância do objeto à tela de mapeamento. Você pode gerar a tabela de registros inválidos para uma instância de transformação de Exceção em um mapeamento. Quando você gera a tabela de registros inválidos, a ferramenta Developer também cria uma tabela de problemas para armazenar os metadados descritivos sobre os registros.

**Nota:** A ferramenta Developer adiciona um sufixo de 12 caracteres a cada nome de coluna nas tabelas de registro inválido. Se você usar um banco de dados Oracle, o nome da coluna de origem não poderá conter mais de 18 caracteres.

## Gerando a Tabela de Registros Inválidos e a Tabela de Problemas

Ao adicionar a transformação a um mapeamento, você pode gerar a tabela de Registros Inválidos e a tabela de Problemas. A ferramenta Developer adiciona as tabelas ao repositório do Modelo.

1. Clique em **Gerar tabela de registros inválidos** para gerar a tabela.  
A caixa de diálogo **Criar Objeto de Dados Relacionais** é exibida.
2. Procure as conexões de banco de dados. Selecione uma conexão com o banco de dados que contém a tabela.
3. Insira um nome para a tabela de Registros Inválidos. A ferramenta Developer aplica o nome que você inserir para a tabela de Registros Inválidos e para a tabela de Problemas.

A ferramenta Developer anexa a seguinte cadeia ao nome de tabela de Problemas:

`_ISSUE`

Se você se conectar a um banco de dados Oracle, o nome de tabela de Registros Inválidos não deverá conter mais de 24 caracteres.

4. Insira um nome para o objeto de dados de Registros Inválidos no repositório do Modelo.
5. Clique em **Concluir**.  
A ferramenta Developer adiciona as tabelas à tela de mapeamento e ao repositório do Modelo.

## Atribuição de Problema de Exceção de Registro Inválido

Você deve atribuir portas e prioridades aos problemas de qualidade de dados.

A figura a seguir mostra a exibição **Atribuição de Problema**:

Atribuição de Problema			
Atribuir portas e prioridades a problemas de qualidade			
Problema de Qualidade	Entrada		Priorid...
EmployeeID_Quality_Issue	EmployeeID		1
Name_Quality_Issue	Employee_Name		1
Phone_Quality_Issue			1

A exibição **Atribuição de Problema** contém os seguintes campos:

### Problema de Qualidade

Cada porta de problema de qualidade que você definiu no grupo de entrada Problemas de Qualidade é exibida na coluna **Problema de Qualidade**.

### Entrada

A coluna **Entrada** contém portas de dados que você atribui a problemas de qualidade na exibição **Atribuição de Problema**. Associar uma porta de entrada a cada problema de qualidade de porta. Cada

porta de entrada que contém de dados de qualidade inválida deve ter pelo menos uma porta de problema de qualidade correspondente que indica um tipo de problema. Você pode selecionar a porta Phone\_Number para Phone\_Quality\_Issue, por exemplo. Você pode atribuir uma porta a mais de um problema de qualidade.

#### **Prioridade do Problema**

As prioridades de problemas determinam quais problemas de qualidade são mais importantes quando você atribui a mesma porta de entrada a vários problemas de qualidade. Se ocorrer mais de um problema de qualidade em uma porta de entrada, o Serviço de Integração de Dados aplicará o problema de qualidade de prioridade mais alta. Se existir mais de um problema de qualidade para uma porta de entrada e os problemas tiverem a mesma prioridade, o Serviço de Integração de Dados aplicará o problema de qualidade mais alto na lista. Insira uma prioridade de 1 a 99, em que 1 representa a prioridade mais alta.

Defina prioridades de problemas para filtrar os registros na ferramenta Analyst.

## Atribuindo Portas para Problemas de Qualidade

Atribua uma porta para associar a cada problema de qualidade. A ferramenta Developer cria portas no grupo de saída de Problemas para cada associação que você adiciona na exibição de **Atribuição de Problema**.

1. Para cada problema de qualidade, clique no campo **Entrada** para exibir uma lista de portas de entrada.
2. Selecione uma porta de entrada para associar ao problema de qualidade.  
Você pode escolher a mesma porta para mais de um problema.
3. Clique na coluna **Problema** e selecione uma prioridade do problema de qualidade.

## Propriedades Avançadas da Transformação de Exceção

Configure propriedades que determinam como o Serviço de Integração de Dados processa dados para a transformação de Exceção.

Você pode configurar níveis de rastreamento para logs.

Configure a seguinte propriedade na guia **Avançado**:

#### **Nível de Rastreamento**

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.

# Configurando uma Transformação de Exceção de Registros Inválidos

Quando você configurar uma Transformação de Exceção de Registros Inválidos, configure as portas de entrada e os problemas de qualidade que podem ocorrer em cada porta. Defina os limites superior e inferior para determinar a qualidade de dados. Configure onde gravar as exceções e os registros rejeitados.

1. Crie uma transformação de Exceção de Registros Inválidos reutilizável ou não reutilizável.
  - Para criar uma transformação reutilizável, selecione **Arquivo > Novo > Transformação** e selecione uma transformação de Exceção de Registros Inválidos.
  - Para criar uma transformação não reutilizável, abra um mapeamento e adicione uma transformação de Exceção à tela de mapeamento. Selecione uma transformação de Exceção de Registros Inválidos no assistente.
2. Clique em **Avançar** ou em **Concluir**.

Se você clicar em **Avançar**, poderá atualizar os valores de limite padrão e as opções de roteamento de dados antes de criar a transformação.
3. Configure as portas de entrada.
  - Se você criar uma transformação reutilizável, selecione a guia **Portas** e adicione portas aos dados que você deseja conectar à transformação.
  - Se você criar uma transformação não reutilizável, adicione outros objetos à tela de mapeamento e arraste as portas de entrada para a transformação.
4. Selecione a exibição **Configuração**.
5. Configure os limites de pontuação superior e inferior.
6. Na seção **Opções de Roteamento de Dados**, configure a saída padrão e as propriedades da tabela de exceção para definir onde a transformação grava cada tipo de registro.

Configure onde gravar os Registros Válidos, Registros Inválidos e Registros Rejeitados. Você pode gravá-las na saída padrão ou na tabela de Registros Inválidos.
7. Abra a exibição **Atribuição de Problema**. Atribua problemas de qualidade de dados às portas de dados.

Atribua uma prioridade a cada problema. Se uma porta contém valores com vários problemas, a transformação exibe o problema de prioridade mais alta.
8. Selecione a opção para gerar uma tabela de registros inválidos. Insira as informações de conexão de banco de dados e nome de tabela. A tabela deve vir do esquema padrão.
  - Ao gerar uma tabela de registros inválidos, você gera uma tabela para os registros e uma tabela adicional para os problemas de qualidade de dados que se relacionam aos registros. A transformação cria um objeto de banco de dados no repositório do Modelo.
9. Conecte as portas de saída da transformação a um ou mais destinos de dados. Conecte as portas de saída aos objetos de dados que correspondem às opções de saída definidas na exibição **Configuração**.
  - Se criar uma transformação reutilizável, adicione a transformação a um mapeamento e conecte as portas de saída.
  - Se você criar uma transformação não reutilizável, a transformação conectará as portas à tabela de registros inválidos. Você conecta portas de saída a qualquer outro destino de dados.

# Exemplo de Mapeamento de Exceção de Registro Inválido

Uma organização conduz um projeto de dados para analisar dados de clientes novos. A organização precisa verificar se os dados de contato do cliente são válidos. O exemplo a seguir mostra como definir uma transformação de Exceção de Registro Inválido que recebe registros de um mapplet que faz uma análise de qualidade de dados de registros de clientes.

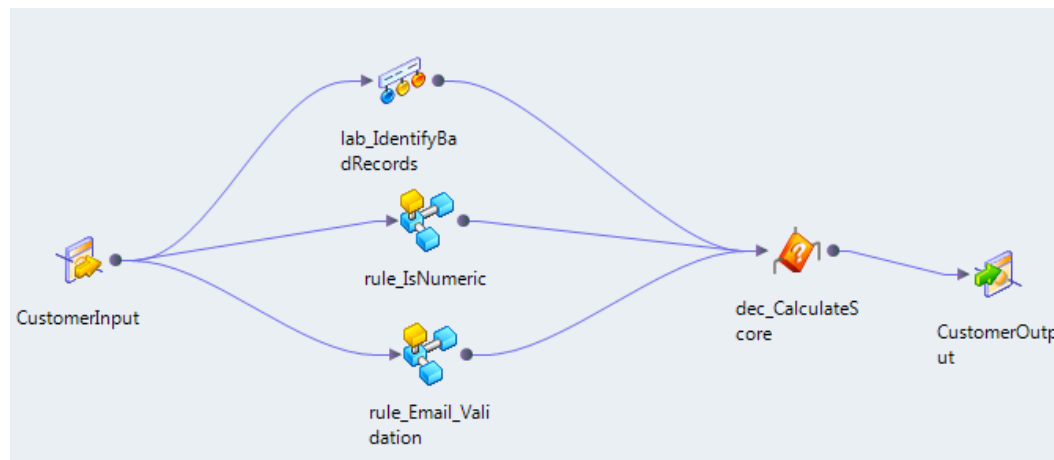
Crie um mapplet com transformações de qualidade de dados que avaliem o formato e a precisão dos dados de clientes. O mapplet inclui transformações que geram uma pontuação do registro com base nos resultados da análise de qualidade de dados. As transformações também definem os problemas de qualidade de dados com base nos resultados da análise.

## Mapplet de Exceção de Registro Inválido

Crie um mapplet que contenha transformações de qualidade de dados para verificar os valores de determinados campos. As transformações verificam as tabelas de referência e os conjuntos de conteúdo para determinar se os campos dos registros são válidos. As transformações aplicam uma pontuação do registro para cada registro com base nos resultados. A transformação de Exceção recebe os registros de mapplet e roteia cada registro para a saída apropriada com base na pontuação do registro.

O mapplet consiste em transformações de Rotulador, Decisão e Expressão.

A figura a seguir mostra os objetos no mapplet:



O mapplet executa as seguintes tarefas:

- Uma transformação de Rotulador verifica a localidade, o estado, o código de país, o código postal e os dados de código postal que recebe nas portas de entrada. A transformação contém uma estratégia para cada porta. As estratégias comparam os dados de origem para tabelas de referência e identificam os valores que não são válidos.
- Um mapplet de transformação de Expressão verifica se o número de telefone é numérico e verifica se o número contém 10 dígitos.
- Uma transformação de Rotulador e um mapplet de transformação de Expressão verificam se o endereço de e-mail é válido. A transformação de Expressão verifica a estrutura da string de e-mail. A transformação de Rotulador verifica o endereço IP em uma tabela de referência de sufixos de endereço IP internacionais.

- Uma transformação de Decisão recebe a saída da transformação e os mapplets. Ela calcula uma pontuação de registro geral para o registro de contato do cliente.

Crie um Mapeamento de exceção de registro inválido que inclua o mapplet. O mapeamento de exceção de registro inválido inclui uma transformação de Exceção que grava exceções de registros inválidos em uma tabela de banco de dados. Um analista de dados pesquisa e atualiza registros de exceção na tabela de registros inválidos com a ferramenta Analyst.

## Grupos de Entrada de Exemplo de Exceção de Registro Inválido

A transformação de Exceção tem três grupos de entrada. A transformação tem um grupo de Dados que recebe os dados de origem. Ela tem o grupo Problemas de Qualidade, que recebe os problemas de qualidade de dados encontrados pelas transformações de qualidade de dados. Ela também tem um grupo de Controle, que contém a pontuação do registro para a linha.

A figura a seguir mostra os grupos de entrada na transformação de Exceção:

	Nome	Tipo	Precisão	Escala	Entrada	Saída
	Entradas					
	Dados (11)					
1	CUST_ID	decimal	8	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
2	COMPANY	string	49	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
3	CONTACT	string	19	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
4	TITLE	string	35	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
5	ADDR1	string	47	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
6	ADDR2	string	47	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
7	ADDR3	string	42	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
8	ADDR4	string	36	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
9	COUNTRY	string	47	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
10	PHONE	string	34	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
11	EMAIL	string	25	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
	Problemas de Qualidade (6)					
1	CompanySt...	string	19	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
2	LocalityStatus	string	22	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
3	ProvinceSta...	string	19	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
4	CountryStat...	string	19	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
5	ZipStatus	string	8	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
6	EmailStatus	string	13	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>
	Controle (1)					
1	Score	double	15	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>

## Configuração de Exemplo de Exceção de Registro Inválido

Defina os limites superior e inferior na exibição **Configuração**. Identifique onde a transformação grava registros válidos, registros inválidos e registros rejeitados.

Aceite a configuração padrão para roteamento os registros válidos, registros inválidos e problemas.

A figura a seguir mostra a exibição **Configuração** da transformação de Exceção:



A tabela a seguir descreve as definições de configuração:

Clique nas tabelas **Gerar registros inválidos** para criar tabelas de Registros Inválidos e Problemas.

Adicione uma transformação de Gravação ao mapeamento e conecte as portas de saída padrão ao objeto de dados. O mapeamento também contém o objeto de banco de dados de registros inválidos e o objeto de banco de dados de problemas que você criou na exibição de **Configuração**.

A tabela de Registros Inválidos contém as exceções de registros com pontuações entre os limites inferior e superior.

A figura a seguir mostra os registros inválidos que a transformação de Exceção retorna:

Name: [excc\\_BadRecords.Bad\\_Output\\_DI](#)

	Workflow_ID	Row_Identifier	Record_Status	CUST_ID	COMPANY	CONTACT	TITLE	ADDR1	ADDR2
1	DummyWor ...	0	INVALID	7121657	WAITROSE	MRS LACI WINI...	DIRECTOR OF IS	380-394 N END...	FULHAM
2	DummyWor ...	1	INVALID	7121649	WAITROSE	MR NICHOLAS...	SENIOR PROJE...	EATON CNTR C...	EATON
3	DummyWor ...	2	INVALID	7121647	VARSHÉE SUPE...	MRS REUVEN C...	MANAGER	834 LONDON RD	THORNTON HE
4	DummyWor ...	3	INVALID	1002138	VICTORY SUPE...	MR RAY ARIAS	COMPUTER SPE...	22 FITCHBURG...	AYER
5	DummyWor ...	4	INVALID	1002137	VICTORY SUPE...	MR NEVILLE DE...	SR PROJ MANA...	21 TIMPANY BLV	GARDNER
6	DummyWor ...	5	INVALID	1002109	USDA FOREST...	MR MICHEAL F...	COMPUTER SPE...	1621 N KENT S...	ROSSLYN
7	DummyWor ...	6	INVALID	1002026	US ARMY CORP...	MR ROBERT EV...	PROGRAMMER...	20 MASS AVE NW	WASHINGTON
8	DummyWor ...	7	INVALID	1002062	US NAVY	MR JOHN WELCH	COMPUTER SPEC	DATA PROCESS...	OAK HARBOUR
9	DummyWor ...	8	INVALID	1062004	TRICOR INDUS...	MR MICHAEL G...	V.P.	8181 PROFESSI...	LANDOVER
10	DummyWor ...	9	INVALID	1001921	THE CORNER S...	MR DENIS LEE	MANAGER (\$\$...	102 MID STR S...	REDHILL
11	DummyWor ...	10	INVALID	7121217	THE CORNER S...	MRS TESSI SAN...	PROGRAMMER	192 BATTERSE...	BATTERSEA

A tabela de Registros Inválidos inclui todos os campos no registro de origem. Um registro inválido também inclui os seguintes campos:

#### Workflow\_ID

O nome do fluxo de trabalho que incluía a transformação de Exceção. O fluxo de trabalho contém a tarefa de Mapeamento da transformação de Exceção e a tarefa Humana para analisar os problemas. O Workflow\_ID conterá DummyWorkflowID se a transformação de Exceção não estiver em um fluxo de trabalho.

#### Row\_Identifier

Um número exclusivo que identifica cada linha.

#### Record\_Status

Um status do registro da ferramenta Analyst. Cada registro na tabela de Registros Inválidos recebe um status de *Inválido*. Você pode manter o status do registro quando atualiza registros na ferramenta Analyst.

## Tabela de Problemas

A tabela de Problemas contém uma linha para cada linha na tabela de Registros Inválidos. Cada linha contém os problemas que a análise de qualidade de dados encontrou para o registro de origem.

A figura a seguir mostra as colunas na tabela de Problemas:

Output

Name: [excc\\_BadRecords.Issues\\_DI](#)

	Workflow_ID	Row_Identifier	COMPANY	DQAPRIORITY_...	ADDR2	DQAPRIORITY_...	ADDR3	DQAPRIORITY_...	COUNTRY	DQ...
1	DummyWor ...	0		1	invalid_locality	2		3		3
2	DummyWor ...	1		1	invalid_locality	2	invalid_province	3		3
3	DummyWor ...	2		1	invalid_locality	2		3		3
4	DummyWor ...	3		1		2		3		3
5	DummyWor ...	4		1		2		3		3
6	DummyWor ...	5		1	invalid_locality	2		3		3
7	DummyWor ...	6		1		2		3		3
8	DummyWor ...	7		1	invalid_locality	2		3		3
9	DummyWor ...	8		1	invalid_locality	2		3		3
10	DummyWor ...	9		1		2		3		3
11	DummyWor ...	10		1	invalid_locality	2		3		3

A tabela de Problemas contém as seguintes colunas:

#### Workflow\_ID

Identifica o fluxo de trabalho que criou o registro. O fluxo de trabalho contém a tarefa de Mapeamento da transformação de Exceção e a tarefa Humana para analisar o problema.

#### Row\_Identifier

Identifica a linha do registro na tabela de banco de dados. O identificador de linha identifica qual linha da tabela de Registros Inválidos corresponde à linha na tabela de Problemas.

### Nome do Campo de Problema

O nome do campo é o nome do campo que pode causar problemas de qualidade. Quando o campo contém um erro, o valor da coluna é o texto do problema de qualidade. Na figura acima, o nome de campo ADDR2 contém o problema de qualidade `invalid_locality`.

### DQAPriority

A prioridade do problema. Quando vários problemas ocorrem para o mesmo campo, o problema com a prioridade mais alta aparece no Nome do Campo de Problema.

## Tabela de Registros Válidos

Cada registro da tabela de Registros Válidos tem uma pontuação do registro maior que o limite superior. Neste exemplo, o limite superior é 90.

A figura a seguir mostra os registros válidos que a transformação de Exceção retorna:

Output								
Name: <a href="#">exc_BadRecords.Output_DI</a>								
	Score	CUST_ID	COMPANY	CONTACT	TITLE	ADDR1	ADDR2	ADDR3
1	100	7121669	YEOVALE STOR...	MRS OPPORTU ...	OWNER/CONSULTANT	1 MARGROVE T...	BARNSTAPLE	DEVON
2	100	7121667	WHARFEDALE...	MRS PRAGER	GENERAL MANAGER	458 SOUTHCOA...	HULL	NORTH HI
3	100	1002195	WORLDSPAN	MR CHRISTOPH...	SENIOR ANALYST	N8J1-1 300 GA...	ATLANTA	GA
4	100	1002187	WINN DIXIE ST...	MS CORINA GA...	VICE PRESIDENT	1805 WAYNE M...	GOLDSBORO	NC
5	100	1002181	WINN DIXIE ST...	MR JENS GONY...	COMPUTER OPERATOR	506 W GANNO...	ZEBULON	NC
6	100	1002183	WINN DIXIE ST...	MS MERCIE VA...	MANAGER OF DBA	5715 GUNN HWY	TAMPA	FL
7	100	1000029	WINN DIXIE ST...	MR LYLE HICK...	MGR, DBA	3123 HWY 28 E	PINEVILLE	LA
8	100	1002178	WINN DIXIE ST...	MR HUSSEIN D...	AUTOMATED SYSTEMS MGR	2080 S FRONTA...	VICKSBURG	MS
9	100	1002177	WINN DIXIE ST...	MR RELUVEN RH...	CHIEF DBA	695 S SEMORA ...	ORLANDO	FL
10	100	1002180	WINN DIXIE ST...	MS GOLDI WEI...	MGR COORD	11957 S APOPK...	ORLANDO	FL

Os registros da tabela de Registros Válidos contêm a pontuação do registro e os campos de dados de origem.

## CAPÍTULO 8

# Transformação de Conversor de Maiúsculas/Minúsculas

Este capítulo inclui os seguintes tópicos:

- [Transformação de Conversor de Maiúsculas/Minúsculas - Visão Geral, 168](#)
- [Propriedades da Estratégia de Maiúsculas/Minúsculas, 169](#)
- [Configurando uma Estratégia de Conversor de Maiúsculas/Minúsculas, 169](#)
- [Propriedades Avançadas da Transformação de Conversor de Maiúsculas/Minúsculas, 170](#)
- [Transformação de Conversor de maiúsculas/minúsculas em um ambiente não nativo, 170](#)

## Transformação de Conversor de Maiúsculas/Minúsculas - Visão Geral

A transformação de Conversor de Maiúsculas/Minúsculas é uma transformação passiva que padroniza o caso dos caracteres alfabéticos em strings de entrada.

Você pode selecionar um formato de conversão de maiúsculas/minúsculas, como Maiúsculas, Minúsculas, Primeira letra da palavra em maiúscula e Primeira letra da frase em maiúscula. Também é possível reverter a formatação de maiúsculas/minúsculas atual de cada caractere nos dados de entrada.

A transformação de Conversor de Maiúsculas/Minúsculas pode usar os valores na coluna Válido de uma tabela de referência para definir a formatação de maiúsculas/minúsculas dos caracteres de entrada. Quando a transformação encontrar uma correspondência entre um valor de entrada e um valor válido, ela aplicará a formatação de maiúsculas/minúsculas do valor válido ao valor de entrada. Você pode usar tabelas de referência quando a conversão de maiúsculas e minúsculas é do tipo **Primeira Letra de Cada Palavra do Título em Maiúscula** ou **Primeira letra da sentença em maiúscula**.

É possível criar várias estratégias de conversão de maiúsculas/minúsculas em uma transformação de Conversor de Maiúsculas/Minúsculas. Cada estratégia usa um único tipo de conversão.

# Propriedades da Estratégia de Maiúsculas/Minúsculas

Você pode configurar propriedades para estratégias de conversão de maiúsculas e minúsculas.

Na exibição **Estratégias**, você pode configurar as seguintes propriedades de conversão de maiúsculas e minúsculas:

## Tipo de Conversão

Define o método de conversão de maiúsculas e minúsculas usado por uma estratégia. Você pode aplicar os seguintes tipos de conversão de maiúsculas e minúsculas:

- **Maiúscula.** Converte todas as letras em maiúsculas.
- **Primeira letra da sentença em maiúscula.** Coloca em maiúscula a primeira letra da string de dados do campo.
- **Alternar entre Maiúsculas e Minúsculas.** Converte letras minúsculas em maiúsculas e vice-versa.
- **Primeira Letra de Cada Palavra do Título em Maiúscula.** Coloca em maiúscula a primeira letra de cada substring.
- **Minúsculas.** Converte todas as letras em minúsculas.

O método padrão de conversão de maiúsculas e minúsculas é o uso de maiúsculas.

## Manter Palavras em Maiúsculas Inalteradas

Ignora o uso de strings em maiúsculas escolhido.

## Delimitadores

Define como o uso de maiúsculas e minúsculas funciona para a conversão de primeira letra de cada palavra do título em maiúscula. Por exemplo, escolha um traço como um delimitador para converter "smith-jones" em "Smith-Jones". O delimitador padrão é o caractere de espaço.

## Tabela de Referência

Aplica o formato de maiúsculas e minúsculas especificado por uma tabela de referência. Aplica-se somente se a opção de conversão de maiúsculas e minúsculas é **Primeira Letra de Cada Palavra do Título em Maiúscula** ou **Primeira letra da sentença em maiúscula**. Clique em **Novo** para adicionar uma tabela de referência à estratégia.

**Nota:** Se ocorrer uma correspondência com a tabela de referência no início de um token, o próximo caractere do token mudará para maiúsculas. Por exemplo, se a string de entrada for mcdonald e a tabela de referência tiver uma entrada para Mc, a string de saída será McDonald.

# Configurando uma Estratégia de Conversor de Maiúsculas/Minúsculas

Para alterar as maiúsculas e minúsculas das strings de entrada, defina as configurações na exibição **Estratégias** de uma transformação de Conversor de Maiúsculas/Minúsculas.

1. Selecione a exibição **Estratégias**.
2. Clique em **Novo**.  
O assistente de **Nova Estratégia** é aberto.

3. Opcionalmente, edite o nome e a descrição da estratégia.
4. Clique nos campos **Entradas** e **Saídas** para selecionar portas para a estratégia.
5. Configure as propriedades da estratégia. A estratégia de conversão padrão é **Maiúsculas**.
6. Clique em **Avançar**.
7. Opcionalmente, adicione tabelas de referência para personalizar as opções de maiúsculas e minúsculas para dados de entrada que correspondam a entradas da tabela de referência. A personalização de maiúsculas e minúsculas da tabela de referência aplica-se somente às estratégias de primeira letra de cada palavra do título em maiúscula e primeira letra da sentença em maiúscula.
8. Clique em **Concluir**.

## Propriedades Avançadas da Transformação de Conversor de Maiúsculas/Minúsculas

Configure propriedades que ajudam a determinar como o Serviço de Integração de Dados processa dados para a transformação de Conversor de Maiúsculas/Minúsculas.

Você pode configurar níveis de rastreamento para logs.

Configure a seguinte propriedade na guia **Avançado**:

### Nível de Rastreamento

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.

## Transformação de Conversor de maiúsculas/minúsculas em um ambiente não nativo

O processamento da transformação de Conversor de maiúsculas/minúsculas em um ambiente não nativo depende do mecanismo que executa a transformação.

Considere o suporte para os seguintes mecanismos de tempo de execução não nativos:

- Mecanismo Blaze. Suportado sem restrições.
- Mecanismo Spark. Com suporte sem restrições em mapeamentos em lote. Sem suporte em mapeamentos de fluxo.
- Mecanismo Databricks Spark. Sem suporte.

## CAPÍTULO 9

# Transformação de Classificador

Este capítulo inclui os seguintes tópicos:

- [Transformação de Classificador - Visão Geral, 171](#)
- [Modelos do Classificador, 172](#)
- [Algoritmos Classificadores, 172](#)
- [Opções de Transformação de Classificador, 172](#)
- [Estratégias de Classificador, 173](#)
- [Propriedades Avançadas da Transformação de Classificador, 173](#)
- [Configurando uma Estratégia de Classificador, 174](#)
- [Exemplo de Análise de Classificador, 174](#)
- [Transformação de Classificador em um ambiente não nativo, 179](#)

## Transformação de Classificador - Visão Geral

A transformação de Classificador é uma transformação passiva que analisa os campos de entrada e determina o tipo de informação em cada campo. Use uma transformação de Classificador quando campos de entrada contêm vários valores de texto.

Ao configurar a transformação de Classificador, selecione um modelo e um algoritmo classificadores. Um modelo classificador é um tipo de objeto de dados de referência. Um algoritmo classificador é um conjunto de regras que calcula o número de palavras similares em uma string e as posições relativas das palavras. A transformação compara a análise de algoritmo com o conteúdo do modelo classificador. A transformação retorna a classificação do modelo que identifica o tipo dominante de informação na string.

A transformação de Classificador pode analisar strings de comprimento significativo. Por exemplo, você pode usar a transformação para classificar o conteúdo de mensagens de e-mail, mensagens de mídia social e texto de documentos. Você passa o conteúdo de cada documento ou mensagem para um campo em uma coluna de fonte de dados e conecta a coluna a uma transformação de Classificador. Em cada caso, você prepara a fonte de dados para que cada campo inclua o conteúdo completo de um documento ou string que você deseja analisar.

# Modelos do Classificador

A transformação do classificador usa um objeto de dados de referência chamado modelo do classificador para analisar os dados de entrada. O modelo do classificador é selecionado quando você configura uma transformação do classificador. A transformação compara uma coluna de dados de entrada com os dados do modelo do classificador e retorna um rótulo que descreve o tipo de informações em cada campo de entrada.

Um modelo do classificador contém linhas de dados de referência e valores de rótulos. As linhas representam os dados de entrada na porta que você pode conectar à transformação do Classificador. Os valores de rótulos descrevem os tipos de informações que as linhas de dados contêm. Ao configurar um modelo do classificador, atribua um rótulo a cada linha de dados de referência no modelo.

Para vincular as linhas de dados de referência aos rótulos em um modelo do classificador, compile o modelo. O processo de compilação gera uma série de associações lógicas entre as linhas de dados e os valores de rótulo. Quando você executa um mapeamento que lê o modelo, o Serviço de Integração de Dados aplica a lógica do modelo aos dados de entrada da transformação do Classificador. O Serviço de Integração de Dados retorna os rótulos que descrevem mais com precisão as informações em cada campo de dados de entrada.

Você pode criar um modelo do classificador na Developer tool. O repositório do Modelo armazena o objeto do modelo do classificador. A Developer tool grava as linhas de dados, os rótulos e a compilação de dados em um arquivo na estrutura de diretórios da Informatica.

## Algoritmos Classificadores

Ao adicionar um modelo classificador a uma estratégia de transformação, você também seleciona um algoritmo classificador. O algoritmo determina como a transformação compara os dados do modelo classificador com os dados de entrada.

Você pode selecionar o algoritmo de **Naive Bayes** ou o algoritmo de **Entropia máxima**.

Considere os seguintes fatores ao selecionar um algoritmo:

- O algoritmo de Entropia máxima executa uma análise mais completa do que o algoritmo de Naive Bayes.
- Um mapeamento que usa o algoritmo de Naive Bayes é executado mais rapidamente do que um mapeamento que usa o algoritmo de Entropia máxima nos mesmos dados.
- Selecione o algoritmo de Entropia máxima com o modelo classificador que a Informatica inclui no Acelerador Central.

## Opções de Transformação de Classificador

A transformação de Classificador exibe opções configuráveis em uma série de guias ou exibições na ferramenta Developer.

Quando você abre uma transformação reutilizável, as opções aparecem em uma série de guias no editor de transformação. Quando você abre uma transformação não reutilizável em um mapeamento, as opções aparecem em uma série de exibições no editor de mapeamento. Selecione a guia de propriedades de mapeamento para ver as exibições em uma transformação não reutilizável.



Você pode selecionar as seguintes exibições:

#### **Geral**

Exiba e atualize o nome da transformação e a descrição.

#### **Portas**

Exiba as portas de entrada e saída na transformação.

**Nota:** Em uma transformação de Classificador reutilizável, as exibições Geral e Portas são combinadas na guia **Visão Geral**.

#### **Estratégias**

Adicione, remova ou edite uma estratégia.

#### **Dependências**

Exiba as portas de entrada e saída em cada estratégia.

#### **Avançado**

Defina o nível de detalhes que a transformação grava em arquivos de log.

## Estratégias de Classificador

Uma estratégia é um conjunto de operações de análise de dados que uma transformação executa em dados de entrada. Você cria pelo menos uma estratégia na transformação de Classificador. Uma estratégia de Classificador lê uma única porta de entrada.

Você define uma ou mais operações em uma estratégia. Uma operação de classificador identifica um modelo e um algoritmo classificadores para aplicar aos dados da porta de entrada. Cada operação grava em uma porta de saída diferente. Crie várias operações em uma estratégia quando quiser analisar uma porta de entrada de diferentes maneiras.

**Nota:** Quando quiser identificar os idiomas usados nos dados de origem, selecione o algoritmo de Entropia máxima na operação de classificador.

## Propriedades Avançadas da Transformação de Classificador

Configure propriedades que ajudam a determinar como o Serviço de Integração de Dados processa dados para a transformação de Classificador.

Você pode configurar níveis de rastreamento para logs.

Configure a seguinte propriedade na guia **Avançado**:

#### **Nível de Rastreamento**

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.

# Configurando uma Estratégia de Classificador

Você configura uma estratégia para identificar os tipos de informação nos dados. Cada estratégia analisa uma porta de entrada.

Em uma transformação não reutilizável, conecte as portas de entrada à transformação antes de configurar a estratégia.

1. Abra a transformação e selecione a exibição **Estratégias**.
2. Clique em **Nova Estratégia**.  
O assistente de criação de estratégia é aberto.
3. Digite um nome e uma descrição opcional para a estratégia.
4. Selecione uma porta de entrada no campo Entradas.
5. Verifique se o valor de precisão da porta de entrada é suficientemente alto para ler todos os campos na porta de entrada. A porta trunca os dados de entrada quando eles atingem o limite de precisão.
6. Selecione ou limpe a opção para adicionar dados de pontuação à saída da estratégia.
7. Clique em **Avançar**.
8. Confirme o tipo de operação de classificador e clique em **Avançar**.
9. Selecione um algoritmo classificador. Você pode selecionar os seguintes algoritmos:
  - Naive Bayes
  - Entropia máxima**Nota:** Para identificar o idioma usado nos dados de origem, selecione o algoritmo de entropia máxima.
10. Verifique a porta de saída.  
A transformação cria uma única porta de saída para cada operação em uma estratégia. Você pode editar o nome da porta e a precisão.
11. Selecione um modelo classificador.  
O assistente lista os objetos de modelo classificador no repositório do Modelo.
12. Clique em **Avançar** para adicionar outra operação à estratégia. Caso contrário, clique em **Concluir**.

## Exemplo de Análise de Classificador

Você é um administrador de dados em uma empresa de software que lançou um novo aplicativo para smartphone. A empresa quer entender a resposta do público ao aplicativo e a cobertura de mídia que ele recebe. A empresa solicita que você e sua equipe analisem os comentários sobre o aplicativo nas mídias sociais.

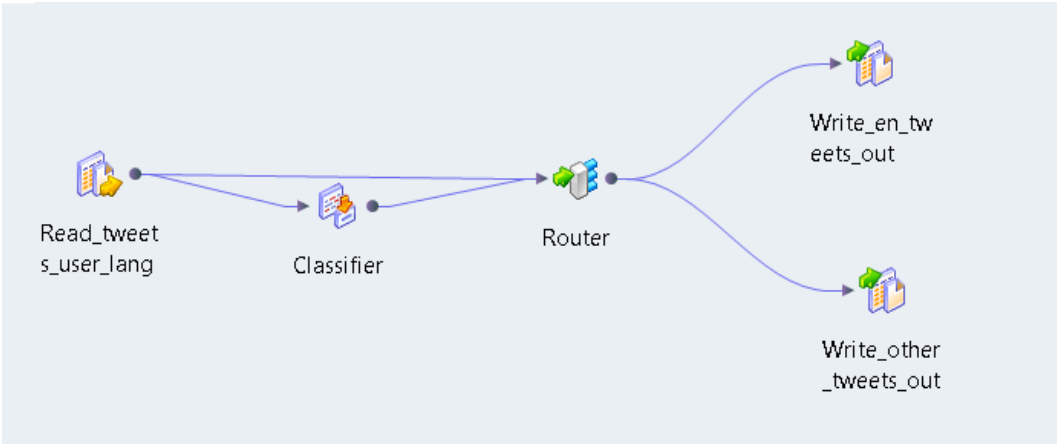
Você decide capturar dados de feeds do Twitter que falam sobre smartphones. Você usa a interface de programação de aplicativo do Twitter para filtrar o fluxo de dados do Twitter. Você cria uma fonte de dados que contém os dados do Twitter que deseja analisar.

Como os feeds do Twitter contêm mensagens em vários idiomas, é necessário identificar o idioma usado em cada mensagem. Você decide usar uma transformação de Classificador para analisar os idiomas. Você cria um mapeamento que identifica os idiomas nos dados de origem e grava as mensagens do Twitter em destinos de dados em inglês e em outros idiomas.

# Criar o Mapeamento Classificador

Você cria um mapeamento que lê uma fonte de dados, classifica os idiomas nos dados e grava os dados em destinos com base nos idiomas que eles contêm.

A seguinte imagem mostra o mapeamento na ferramenta Developer:



O mapeamento criado contém os seguintes objetos:

Nome do Objeto	Descrição
Read_tweet_user_lang	Fonte de dados. Contém as mensagens do Twitter.
Classificador	Transformação de Classificador. Identifica os idiomas usados nas mensagens do Twitter.
Roteador	Transformação de Roteador. Roteia as mensagens do Twitter para objetos de destino de dados de acordo com os idiomas que elas contêm.
Write_en_tweets_out	Destino de dados. Contém as mensagens do Twitter em inglês.
Write_other_tweets_out	Destino de dados. Contém mensagens do Twitter em outros idiomas.

## Amostra de Dados de Entrada

O seguinte fragmento de dados mostra uma amostra dos dados do Twitter analisados no mapeamento:

### Twitter Message

```
RT @GanaphoneS3: Faltan 10 minutos para la gran rifa de un iPhone 5...
RT @Clarified: How to Downgrade Your iPhone 4 From iOS 6.x to iOS 5.x (Mac)...
RT @jerseyjazz: The razor was the iPhone of the early 2000s
RT @KrissiDevine: Apple Pie that I made for Thanksgiving. http://t.com/s9ImzFxO
RT @sophieHz: Dan yang punya 2 kupon undian. Masuk dalam kotak undian yang berhadiah Samsung
RT @IsabelFreitas: o galaxy tem isso isso isso e a bateria ã melhor que do iPhone
RT @PremiusIpad: Faltan 15 minutos para la gran rifa de un iPhone 5...
RT @payyton3: I want apple cider
RT @wiesteronder: Retweet als je iets van Apple, Nike, Adidas of microsoft hebt!
```

## Configuração da Fonte de Dados

A fonte de dados contém uma única porta. Cada linha na porta contém uma única mensagem do Twitter.

A seguinte tabela descreve a configuração da fonte de dados:

Nome da Porta	Tipo de Porta	Precisão
texto	n/d	200

## Configuração da Transformação de Classificador

A transformação de Classificador usa uma única porta de entrada e uma porta de saída. A porta de entrada da transformação lê o campo de texto da fonte de dados. A porta de saída contém o idioma identificado para cada mensagem do Twitter no campo de texto. A transformação de Classificador usa códigos de país ISO para identificar o idioma.

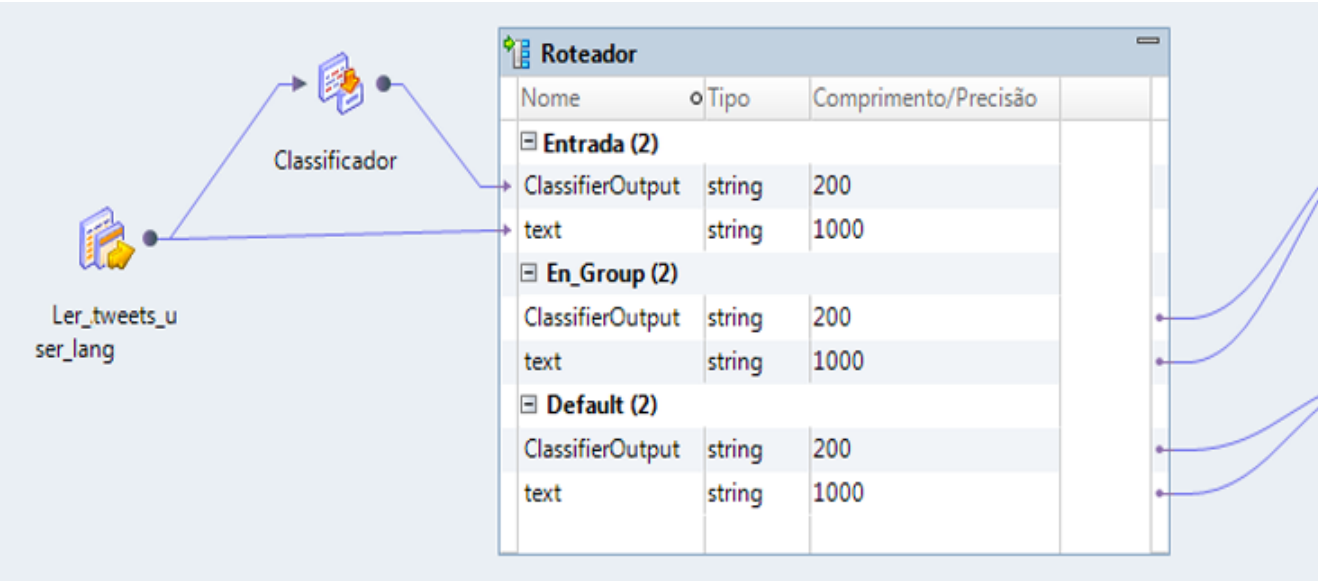
A seguinte tabela descreve a configuração da transformação de Classificador:

Nome da Porta	Tipo de Porta	Precisão	Estratégia
text_input	Entrada	200	Classifier1
Classifier_Output	Saída	2	Classifier1

# Configuração da Transformação de Roteador

A transformação de Roteador usa duas portas de entrada. Ela lê as mensagens do Twitter da fonte de dados e os códigos de país ISO da transformação de Classificador. A transformação de Roteador roteia os dados das portas de entrada para diferentes portas de saída com base em uma condição que você especifica.

A seguinte imagem mostra os grupos de portas da transformação de Roteador e as conexões de porta:



A seguinte tabela descreve a configuração da transformação de Roteador:

Nome da Porta	Tipo de Porta	Grupo de Portas	Precisão
Classifier_Output	Entrada	Entrada	2
texto	Entrada	Entrada	200
Classifier_Output	Entrada	Padrão	2
texto	Entrada	Padrão	200
Classifier_Output	Entrada	En_Group	2
texto	Entrada	En_Group	200

Você configura a transformação para criar fluxos de dados para mensagens no idioma inglês e para mensagens em outros idiomas. Para criar um fluxo de dados, adicione um grupo de portas de saída à transformação. Use as opções de **Grupos** na transformação para adicionar o grupo de portas.

Para determinar como a transformação roteia dados para cada fluxo de dados, você define uma condição em um grupo de portas. A condição identifica uma porta e especifica um possível valor na porta. Quando a transformação encontra um valor da porta de entrada que corresponde à condição, ela roteia os dados de entrada para o grupo de portas que se aplica à condição.

Defina a seguinte condição em En\_Group:

```
ClassifierOutput='en'
```

**Nota:** A transformação de Roteador lê dados de dois objetos no mapeamento. A transformação pode combinar os dados de cada grupo de saída, pois não altera a sequência de linha definida nos objetos de dados.

## Configuração de Destino de Dados

O mapeamento contém um destino de dados para mensagens do Twitter no idioma inglês e um destino para as mensagens em outros idiomas. Você conecta as portas de um grupo de saída da transformação de Roteador para um destino de dados.

A seguinte tabela descreve a configuração dos destinos de dados:

Nome da Porta	Tipo de Porta	Precisão
texto	n/d	200
Classifier_Output	n/d	2

## Resultado do Mapeamento Classificador

Quando você executa o mapeamento, a transformação de Classificador identifica o idioma de cada mensagem do Twitter. A transformação de Roteador grava o texto da mensagem em destinos de dados com base nas classificações de idioma.

O seguinte fragmento de dados mostra um exemplo dos dados de destino no idioma inglês:

ISO Country Code	Twitter Message
en	RT @Clarified: How to Downgrade Your iPhone 4 From iOS 6.x to iOS 5.x (Mac)...
en	RT @jerseyjazz: The razor was the iPhone of the early 2000s
en	RT @KrisiDevine: Apple Pie that I made for Thanksgiving. <a href="http://t.com/s9ImzFxO">http://t.com/s9ImzFxO</a>
en	RT @payyton3: I want apple cider

O seguinte fragmento de dados mostra um exemplo dos dados de destino identificados para outros idiomas:

ISO Country Code	Twitter Message
es	RT @GanaphoneS3: Faltan 10 minutos para la gran rifa de un iPhone 5...
id	RT @sophieHz: Dan yang punya 2 kupon undian. Masuk dalam kotak undian yang berhadiah Samsung Champ.
pt	RT @IsabelFreitas: o galaxy tem isso isso isso e a bateria ã melhor que do iPhone
es	RT @PremiusIpad: Faltan 15 minutos para la gran rifa de un iPhone 5...
nl	RT @wiesteronder: Retweet als je iets van Apple, Nike, Adidas of microsoft hebt! <a href="http://t.co/Je6Ts00H">http://t.co/Je6Ts00H</a>

# Transformação de Classificador em um ambiente não nativo

O processamento da transformação de Classificador em um ambiente não nativo depende do mecanismo que executa a transformação.

Considere o suporte para os seguintes mecanismos de tempo de execução não nativos:

- Mecanismo Blaze. Suportado sem restrições.
- Mecanismo Spark. Com suporte sem restrições em mapeamentos de fluxo e em lote.
- Mecanismo Databricks Spark. Sem suporte.

## CAPÍTULO 10

# Transformação de Comparação

Este capítulo inclui os seguintes tópicos:

- [Transformação de Comparação - Visão Geral, 180](#)
- [Estratégias de Correspondência de Campo, 180](#)
- [Estratégias de Correspondência de Identidade, 183](#)
- [Configurando uma Estratégia de Comparação, 184](#)
- [Propriedades Avançadas da Transformação de Comparação, 184](#)
- [Transformação de Comparação em um ambiente não nativo, 185](#)

## Transformação de Comparação - Visão Geral

A transformação de Comparação é uma transformação passiva que avalia a similaridade entre pares de strings de entrada e calcula o grau de similaridade de cada par como uma pontuação numérica.

Quando você configura a transformação, seleciona um par de colunas de entrada e atribui uma estratégia de correspondência a elas.

A transformação de Comparação gera pontuações de correspondência em um intervalo de 0 a 1, onde 1 indica uma correspondência perfeita.

**Nota:** As estratégias disponíveis na transformação de Comparação também estão disponíveis na transformação de Correspondência. Use a transformação de Comparação para definir as operações de comparação de correspondência que você adicionará a um mapplet de correspondência. Você pode adicionar várias transformações de Comparação ao mapplet. Use a transformação de Correspondência para definir comparações de correspondência em uma única transformação. Você pode incorporar um mapplet correspondente em uma transformação de Correspondência.

## Estratégias de Correspondência de Campo

A transformação de Comparação inclui estratégias predefinidas de correspondência de campo que comparam pares de campos de dados de entrada.



## Bigram

Use o algoritmo Bigram para comparar strings de texto longo, como endereços postais inseridos em um único campo.

O algoritmo Bigram calcula uma pontuação de correspondência para duas strings de dados com base na ocorrência de caracteres consecutivos em ambas as strings. O algoritmo procura pares de caracteres consecutivos que são comuns para ambas as strings. Ele divide o número de pares correspondentes nessas strings pelo número total de pares de caracteres.

### Exemplo de Bigram

Considere as seguintes strings:

- larder
- lerder

Essas strings geram os seguintes grupos de Bigram:

```
l a, a r, r d, d e, e r
l e, e r, r d, d e, e r
```

Observe que a segunda ocorrência da string "e r" dentro da string "lerder" não é correspondente, pois não há uma segunda ocorrência correspondente de "e r" na string "larder".

Para calcular a pontuação de correspondência do Bigram, a transformação divide o número de pares correspondentes (6) pelo número total de pares em ambas as strings (10). Nesse exemplo, as strings são 60% semelhantes, e a pontuação de correspondência é 0,60.

## Distância de Hamming

Use o algoritmo Distância de Hamming quando a posição dos caracteres de dados for um fator crítico, por exemplo, em campos numéricos ou de código como números de telefone, códigos postais ou códigos de produto.

O algoritmo Distância de Hamming calcula uma pontuação de correspondência para duas strings de dados computando o número de posições nas quais os caracteres diferem entre as strings de dados. Para strings de comprimento diferente, cada caractere adicional na string mais longa é contado como uma diferença entre as strings.

### Exemplo de Distância de Hamming

Considere as seguintes strings:

- Morlow
- Marlowes

Os caracteres realçados indicam as posições que o algoritmo de Hamming identifica como diferentes.

Para calcular a pontuação de correspondência de Hamming, a transformação divide o número de caracteres correspondentes (5) pelo comprimento da string mais longa (8). Nesse exemplo, as strings são 62,5% semelhantes, e a pontuação de correspondência é 0,625.

## Distância de Edição

Use o algoritmo Distância de Edição para comparar palavras ou strings de texto curto, como nomes.

O algoritmo Distância de Edição calcula o "custo" mínimo de transformar uma string em outra por meio da inserção, exclusão ou substituição de caracteres.

## Exemplo de Distância de Edição

Considere as seguintes strings:

- Levenston
- Levenshtein

Os caracteres realçados indicam as operações necessárias para transformar uma string em outra.

O algoritmo Distância de Edição divide o número de caracteres inalterados (8) pelo comprimento da string mais longa (11). Nesse exemplo, as strings são 72,7% semelhantes, e a pontuação de correspondência é 0,727.

## Distância de Jaro

Use o algoritmo Distância de Jaro para comparar duas strings quando a similaridade dos caracteres iniciais nas strings for uma prioridade.

A pontuação de correspondência Distância de Jaro reflete o grau de semelhança entre os quatro primeiros caracteres de ambas as strings e o número de transposições de caracteres identificadas. A transformação pesa a importância da correspondência entre os quatro primeiros caracteres usando o valor que você insere na propriedade Penalidade.

### Propriedades da Distância de Jaro

Ao configurar um algoritmo Distância de Jaro, é possível definir as seguintes propriedades:

#### Penalidade

Determina a penalidade da pontuação de correspondência se os quatro primeiros caracteres em duas strings comparadas não forem idênticos. A transformação subtrai o valor total da penalidade para uma incompatibilidade de primeiro caractere e subtrai frações da penalidade com base na posição dos outros caracteres incompatíveis. O valor da penalidade padrão é 0,20.

#### Distinção Entre Maiúsculas e Minúsculas

Determina se o algoritmo Distância de Jaro faz distinção entre maiúsculas e minúsculas ao comparar caracteres.

### Exemplo de Distância de Jaro

Considere as seguintes strings:

- 391859
- 813995

Se você usar o valor de Penalidade padrão de 0,20 para analisar essas strings, o algoritmo Distância de Jaro retornará uma pontuação de correspondência de 0,513. Essa pontuação de correspondência indica que as strings são 51,3% semelhantes.

## Distância de Hamming Invertida

Use o algoritmo Distância de Hamming Invertido para calcular a porcentagem de posições de caractere que diferem entre as duas strings quando se faz a leitura da direita para a esquerda.

O algoritmo Distância de Hamming calcula uma pontuação de correspondência para duas strings de dados computando o número de posições nas quais os caracteres diferem entre as strings de dados. Para strings de comprimento diferente, o algoritmo conta cada caractere adicional na string mais longa como uma diferença entre as strings.

### Exemplo de Distância de Hamming Invertido

Considere as seguintes strings, que usam alinhamento da direita para a esquerda para imitar o algoritmo Hamming Invertido:

- 1-999-9999
- **011-01**-999-9991

Os caracteres realçados indicam as posições que o algoritmo Distância de Hamming Invertido identifica como diferentes.

Para calcular a pontuação de correspondência do Hamming Invertido, a transformação divide o número de caracteres correspondentes (9) pelo comprimento da string mais longa (15). Nesse exemplo, a pontuação de correspondência é 0,6, indicando que as strings são 60% semelhantes.

## Estratégias de Correspondência de Identidade

A transformação de Comparação inclui estratégias predefinidas de correspondência de identidade que você pode usar para localizar correspondências para indivíduos, endereços ou entidades corporativas.

A tabela a seguir descreve a operação de correspondência que cada estratégia de correspondência de identidade executa:

Estratégia de Correspondência de Identidade	Operação de Correspondência
Endereço	Identifica um endereço de correspondência.
Contato	Identifica um contato dentro de uma organização em um único local.
Entidade Corporativa	Identifica uma organização por seu nome corporativo legal.
Divisão	Identifica uma organização em um endereço.
Família	Identifica uma família por um nome de família e endereço ou número de telefone.
Campos	Identifica campos personalizados que você selecionar.
Membros da Família	Identifica os membros da mesma família na mesma residência.
Indivíduo	Identifica um indivíduo pelo nome e ID ou data de nascimento.
Organização	Identifica uma organização pelo nome.
Nome da Pessoa	Identifica uma pessoa pelo nome.
Residente	Identifica uma pessoa em um endereço.

Estratégia de Correspondência de Identidade	Operação de Correspondência
Contato Amplo	Identifica um contato dentro de uma organização, independentemente da localização.
Família Ampla	Identifica os membros da mesma família independentemente da localização.

**Nota:** As estratégias de correspondência de identidade leem arquivos de dados de referência chamados . Entre em contato com o usuário do Informatica Administrator para obter informações sobre os arquivos de dados de preenchimento instalados no sistema.

## Configurando uma Estratégia de Comparação

Para configurar uma estratégia de comparação, edite as configurações na exibição **Estratégias** de uma transformação de Comparação.

1. Selecione a exibição **Estratégias**.
2. Selecione uma estratégia de comparação na seção **Estratégias**.
3. Na seção **Campos**, clique duas vezes em uma célula na coluna **Campos Disponíveis** para selecionar uma entrada.

**Nota:** Você deve selecionar uma entrada para cada linha que exiba um nome de entrada em negrito na coluna **Campos de Entrada**.

## Propriedades Avançadas da Transformação de Comparação

Configure propriedades que ajudam a determinar como o Serviço de Integração de Dados processa dados para a transformação de Comparação.

Você pode configurar níveis de rastreamento para logs.

Configure a seguinte propriedade na guia **Avançado**:

### Nível de Rastreamento

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.

# Transformação de Comparação em um ambiente não nativo

O processamento da transformação de Comparação em um ambiente não nativo depende do mecanismo que executa a transformação.

Considere o suporte para os seguintes mecanismos de tempo de execução não nativos:

- Mecanismo Blaze. Suportado sem restrições.
- Mecanismo Spark. Com suporte sem restrições em mapeamentos em lote. Sem suporte em mapeamentos de fluxo.
- Mecanismo Databricks Spark. Sem suporte.

## CAPÍTULO 11

# Transformação de Consolidação

Este capítulo inclui os seguintes tópicos:

- [Transformação de Consolidação - Visão Geral, 186](#)
- [Mapeamentos de Consolidação, 187](#)
- [Portas da Transformação de Consolidação, 187](#)
- [Exibições da Transformação de Consolidação, 187](#)
- [Estratégias Simples, 190](#)
- [Estratégias Baseadas em Linhas, 191](#)
- [Estratégias Avançadas, 192](#)
- [Funções de Consolidação Simples, 192](#)
- [Funções de Consolidação Baseadas em Linhas, 196](#)
- [Exemplo de Mapeamento de Consolidação, 200](#)
- [Configurando uma Transformação de Consolidação, 201](#)
- [Transformação de Consolidação em um ambiente não nativo, 202](#)

## Transformação de Consolidação - Visão Geral

A transformação de Consolidação é uma transformação ativa que analisa grupos de registros relacionados e cria um registro consolidado para cada grupo. Use a transformação de Consolidação para consolidar grupos de registros gerados por transformações como as de Gerador de Chaves, Correspondência e Associação.

A transformação de Consolidação gera registros consolidados aplicando estratégias a grupos de registros relacionados. A transformação contém uma porta de saída que indica qual registro é o consolidado. Você pode optar por limitar a saída da transformação para incluir apenas registros consolidados.

Por exemplo, você pode consolidar grupos de registros duplicados de funcionários gerados pela transformação de Correspondência. A transformação de Consolidação pode criar um registro consolidado que contém dados mesclados de todos os registros no grupo.

Você pode configurar a transformação de Consolidação para usar diferentes tipos de estratégias com base nos seus requisitos de consolidação. Use estratégias simples para criar um registro consolidado de vários registros. Ao usar estratégias simples, você especifica uma estratégia para cada porta. Use uma estratégia baseada em linhas para analisar linhas no grupo de registros e criar um registro consolidado com os valores de uma das linhas. Use uma estratégia avançada para criar um registro consolidado aplicando uma expressão criada por você.

# Mapeamentos de Consolidação

Para consolidar registros, crie um mapeamento que cria grupos de registros relacionados. Adicione uma transformação de Consolidação a um mapeamento e configure a transformação para consolidar cada grupo de registros em um único registro principal.

Conecte uma transformação de Consolidação a outras transformações de acordo com os objetivos comerciais e os requisitos de dados. Para consolidar registros correspondentes, você pode conectar a transformação de Consolidação a uma transformação de Correspondência. Para consolidar registros como parte do gerenciamento de registros de exceção, conecte a transformação de Consolidação a uma transformação de Exceção. Por exemplo, se você usar uma transformação de Gerador de Chaves para agrupar registros, poderá conectar uma transformação de Consolidação diretamente a essa transformação de Gerador de Chaves. A transformação de Consolidação cria um registro consolidado para cada grupo criado pela transformação de Gerador de Chaves.

## Saída de mapeamento em ambientes nativos e Hadoop

Quando você executa um mapeamento de consolidação em um ambiente nativo e em um ambiente Hadoop, a transformação de Consolidação pode gerar resultados diferentes. Como o mapeamento é executado em vários nós no Hadoop, os registros de entrada podem entrar na transformação de Consolidação em uma ordem diferente daquela no ambiente nativo. Como resultado, a transformação pode gerar diferentes conjuntos de registros sobreviventes em cada ambiente para o mesmo conjunto de dados de entrada. Em cada caso, os cálculos da transformação e os resultados consolidados são precisos para a ordem de linhas de entrada.

Para gerar os mesmos registros sobrevivente em ambientes nativos e Hadoop, configure a transformação de Consolidação para classificar os registros na seguinte ordem:

- Primeiro, classifique os registros na porta Agrupar por.
- Em seguida, classifique os registros na ordem em que as portas de entrada aparecem na transformação.

# Portas da Transformação de Consolidação

A ferramenta Developer cria uma porta de saída para cada porta de entrada que você adicionar. Não é possível adicionar as portas de saída manualmente à transformação. A transformação de Consolidação também inclui uma porta de saída **IsSurvivor** que indica os registros consolidados.

Uma das portas de entrada que você adicionar à transformação de Consolidação deverá conter chaves de grupo. A transformação de Consolidação requer informações de chave de grupo, pois as estratégias de consolidação processam grupos de registros em vez de conjuntos de dados inteiros.

Quando você adiciona uma porta de entrada, a ferramenta Developer cria o nome da porta de saída adicionando o sufixo "1" ao nome da porta de entrada. A transformação inclui também a porta de saída **IsSurvivor**, que indica se um registro é o registro consolidado. Para registros consolidados, a transformação de Consolidação grava a string "Y" na porta **IsSurvivor**. Para registros de entrada, a transformação de Consolidação grava a string "N" na porta **IsSurvivor**.

# Exibições da Transformação de Consolidação

A transformação de Consolidação contém exibições para portas, estratégias e propriedades avançadas.

## Exibição Estratégias da Transformação de Consolidação

A exibição **Estratégias** contém propriedades para estratégias simples, baseadas em linhas e avançadas.

A seguinte lista descreve os tipos de estratégias de consolidação:

### Estratégia Simples

Uma estratégia simples analisa todos os valores de uma porta em um grupo de registros e seleciona um valor. Você especifica uma estratégia simples para cada porta. A transformação de Consolidação usa os valores de porta selecionados por todas as estratégias simples para criar um registro consolidado. Exemplos de estratégias simples incluem o valor mais frequente em uma porta, o valor mais longo em uma porta ou o valor que não esteja em branco mais frequente em uma porta.

### Estratégia Baseada em Linhas

Uma estratégia baseada em linhas analisa linhas no grupo de registros e seleciona uma linha. A transformação de Consolidação usa os valores de porta dessa linha para criar um registro consolidado. Exemplos de estratégias baseadas em linhas incluem a contagem de caracteres mais alta, o número mais baixo de campos em branco ou a contagem mais alta dos campos mais frequentes.

### Estratégia Avançada

Uma estratégia avançada analisa um grupo de registros usando estratégias definidas por você. Você cria estratégias avançadas usando funções de consolidação em uma expressão. A transformação de Consolidação cria um registro consolidado com base na saída da expressão. A expressão criada por você também pode usar todas as funções disponíveis na transformação de Decisão.

## Propriedades Avançadas da Transformação de Consolidação

A transformação de Consolidação contém propriedades avançadas que determinam o comportamento de classificação, o modo de saída, o comportamento da memória cache e o nível de rastreamento.

Você pode configurar as seguintes propriedades avançadas:

### Classificação

Determina se a transformação classifica as linhas de entrada nos dados da porta **Agrupar por**. A propriedade é ativada por padrão.

Selecione a propriedade se as linhas de entrada não estiverem pré-classificadas.

### Classificação com Distinção entre Maiúsculas e Minúsculas

Determina se a operação de classificação faz distinção entre maiúsculas e minúsculas. A propriedade é ativada por padrão.

### Modo de Saída

Determina se a transformação grava todos os registros como saída ou grava os registros consolidados como saída. O valor padrão é Todos.

### Diretório de Arquivo de Cache

Especifica o diretório no qual o Serviço de Integração de Dados grava os dados temporários da transformação atual. O Serviço de Integração de Dados grava os arquivos temporários no diretório quando o volume de dados de entrada é maior do que a memória disponível do sistema. O Serviço de Integração de Dados exclui os arquivos temporários depois que o mapeamento é executado.

Você pode inserir um caminho de diretório na propriedade ou usar um parâmetro para identificar o diretório. Especifique um caminho local na máquina do Serviço de Integração de Dados. O Serviço de Integração de Dados deve ser capaz de gravar no diretório. O valor padrão é o parâmetro do sistema CacheDir.



### Tamanho do Arquivo de Cache

Determina a quantidade de memória do sistema que o Serviço de Integração de Dados usa para classificar os dados de entrada na transformação. O valor padrão é 400.000 bytes. Você pode usar um parâmetro para especificar o tamanho do arquivo de cache.

Antes de classificar os dados, o Serviço de Integração de Dados aloca a quantidade de memória especificada por você. Se a operação de classificação gerar uma quantidade maior de dados, o Serviço de Integração de Dados gravará o excesso de dados no diretório do arquivo de cache. Se a operação de classificação exigir mais memória do que a memória do sistema e o armazenamento de arquivos podem fornecer, o mapeamento falhará.

**Nota:** Se você inserir um valor igual ou maior que 65536, a transformação lerá o valor em bytes. Se você inserir um valor menor, a transformação lerá o valor em megabytes.

### Nível de Rastreamento

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.

## Tamanho do Arquivo de Cache

A propriedade do tamanho do arquivo de cache determina a quantidade de memória do sistema que o Serviço de Integração de Dados atribui à transformação de Consolidação para operações de classificação. Configure a propriedade com um valor menor ou igual à quantidade de RAM na máquina host do Serviço de Integração de Dados.

Para obter um desempenho melhor, especifique um tamanho de arquivo de cache de pelo menos 16 MB.

Antes de começar uma operação de classificação, o Serviço de Integração de Dados aloca a memória especificada pela propriedade do tamanho do arquivo de cache. O Serviço de Integração de Dados transmite todos os dados de entrada para a transformação de Consolidação antes de realizar uma operação de classificação.

Se o volume de dados de entrada for maior do que o tamanho do arquivo de cache, o Serviço de Integração de Dados gravará dados no diretório do arquivo de cache. Quando ele grava dados no diretório do arquivo de cache, o Serviço de Integração de Dados consome espaço em disco que representa pelo menos duas vezes o volume de dados de entrada.

Utilize a seguinte fórmula para determinar o tamanho dos dados recebidos:

$$[\text{number\_of\_input\_rows} * (\text{Sum}(\text{column\_size}) + 16)]$$

A seguinte tabela lista os possíveis tipos de dados e os valores de tamanho de coluna a serem aplicados nos cálculos de dados do arquivo de cache:

Tipo de Dados	Tamanho da Coluna
Binário	Precisão + 8. Arredondar para o múltiplo de 8 mais próximo.
Data/Hora	29
Decimal, alta precisão desativada (precisão total)	16
Decimal, alta precisão ativada (precisão <=18)	24
Decimal, alta precisão ativada (precisão >18, <=28)	32

Tipo de Dados	Tamanho da Coluna
Decimal, alta precisão ativada (precisão >28)	16
Decimal, alta precisão ativada (escala negativa)	16
Duplos	16
Real	16
Número inteiro	16
Cadeia, Texto	Modo Unicode: 2*(precisão + 5) Modo ASCII: precisão + 9

## Estratégias Simples

Uma estratégia simples analisa uma porta em um grupo de registros e retorna um valor. Você especifica uma estratégia simples para cada porta. A transformação de Consolidação usa os valores de porta selecionados por todas as estratégias simples para criar um registro consolidado.

Quando você configura uma estratégia na exibição **Estratégias** da transformação, a estratégia exibe o seguinte texto como o método de consolidação:

*Usar padrão.*

A estratégia padrão é "ID da linha mais alta".

Você pode escolher uma das seguintes estratégias simples:

### Média

Analisa uma porta no grupo de registros e retorna a média de todos os valores.

Para os tipos de dados String e Data/hora, a estratégia retorna o valor de ocorrência mais frequente.

### Mais Longo

Analisa uma porta no grupo de registros e retorna o valor com a contagem de caracteres mais alta. Se a contagem de caracteres mais alta é compartilhada por dois ou mais valores, a estratégia retorna o primeiro valor qualificado.

### Máximo

Analisa uma porta no grupo de registro e retorna o valor mais alto.

Para o tipo de dados String, a estratégia retorna a string mais longa. Para o tipo de dados Data/hora, a estratégia retorna a data mais recente.

### Mínimo

Analisa uma porta no grupo de registros e retorna o valor mais baixo.

Para o tipo de dados String, a estratégia retorna a string mais curta. Para o tipo de dados Data/hora, a estratégia retorna a data mais recente.

### Mais frequente

Analisa uma porta no grupo de registros e retorna o valor de ocorrência mais frequente, incluindo valores em branco ou nulos. Se o número mais alto de ocorrências é compartilhado por dois ou mais valores, a estratégia retorna o primeiro valor qualificado.

### Mais frequente não em branco

Analisa uma porta no grupo de registros e retorna o valor de ocorrência mais frequente, excluindo valores em branco ou nulos. Se o número mais alto de ocorrências não em branco é compartilhado por dois ou mais valores, a estratégia retorna o primeiro valor qualificado.

### Mais Curto

Analisa uma porta no grupo de registros e retorna o valor com a contagem de caracteres mais baixa. Se a contagem de caracteres mais baixa é compartilhada por dois ou mais valores, a estratégia retorna o primeiro valor qualificado.

### ID da linha mais alta

Analisa uma porta no grupo de registros e retorna o valor com a ID da linha mais alta.

## Estratégias Baseadas em Linhas

Uma estratégia baseada em linhas analisa linhas no grupo de registros e seleciona uma linha. A transformação de Consolidação usa os valores de porta dessa linha para criar um registro consolidado. A estratégia padrão é "mais dados".

Escolha uma das seguintes estratégias baseadas em linhas:

### Mais dados

Seleciona a linha com a contagem de caracteres mais alta. Se a contagem de caracteres mais alta é compartilhada por duas ou mais linhas, a estratégia retorna o último valor qualificado.

### Mais preenchido

Seleciona a linha com o número mais alto de colunas não em branco. Se o número mais alto de colunas não em branco é compartilhado por duas ou mais linhas, a estratégia retorna o último valor qualificado.

### Modal exato

Seleciona a linha com a contagem mais alta dos valores não em branco mais frequentes. Por exemplo, considere uma linha que tem três portas contendo os valores mais frequentes no grupo de registros. A contagem dos valores mais frequentes para essa linha é "3".

Se a contagem mais alta dos valores não em branco mais frequentes é compartilhada por duas ou mais linhas, a estratégia retorna o último valor qualificado.

### Exemplo de Estratégia Baseada em Linhas

A tabela a seguir mostra um grupo de registros de amostra. A última coluna descreve os motivos pelos quais estratégias específicas baseadas em linhas selecionam diferentes linhas neste grupo de registros.

ID do Produto	Nome	SobrenomeCódigo Postal	Seleção de Estratégia
2106	Bartholomew	28516	A estratégia Mais Dados seleciona esta linha porque ela contém mais caracteres do que as outras linhas.

ID do Produto	Nome	Sobrenome	Código Postal	Seleção de Estratégia
2236	Bart	Smith	28579	A estratégia Mais Preenchido seleciona esta linha porque ela tem mais colunas não em branco do que as outras linhas.
2236	<Blank>	Smith	28516	A estratégia Modal Exato seleciona esta linha porque ela contém a contagem mais alta dos valores mais frequentes.

## Estratégias Avançadas

Você pode usar estratégias avançadas para criar estratégias de consolidação a partir de funções predefinidas. É possível usar funções de consolidação e outras funções da Informatica.

Você pode criar expressões que contêm funções de consolidação simples ou funções de consolidação baseadas em linhas. Use as funções de consolidação simples para criar um registro consolidado com base em valores de porta no grupo de registros. Use as funções de consolidação baseadas em linhas para selecionar uma linha do grupo de registros.

As expressões de consolidação devem preencher todas as portas de saída na transformação de Consolidação. Se as expressões de consolidação não usarem todas as portas de saída, a transformação fará os mapeamentos falharem.

Você pode usar uma estratégia simples ou baseada em linhas como um modelo para uma estratégia avançada. Configure uma estratégia simples ou baseada em linhas e selecione Avançado. A transformação de Consolidação gera uma expressão com funções que executam a estratégia. Você pode adicionar mais funções para implementar requisitos adicionais.

## Funções de Consolidação Simples

As funções de consolidação simples selecionam um valor de todos os valores de porta em um grupo de registros. Ao usar uma função de consolidação simples, você fornece a função com uma porta e o grupo por porta.

### CONSOL\_AVG

Analisa uma porta no grupo de registros e retorna a média de todos os valores.

#### Sintaxe

```
CONSOL_AVG(string, group by)
```

A tabela a seguir descreve os argumentos para este comando:

Argumento	Required/Optional	Descrição
<i>string</i>	Obrigatório	Nome da porta de entrada.
<i>agrupar por</i>	Obrigatório	Nome da porta de entrada que contém o identificador de grupo.

## Valor de Retorno

A média de todos os valores em uma porta.

Para os tipos de dados String e Data/hora, a função retorna o valor de ocorrência mais frequente.

## Exemplo

A seguinte expressão usa a função `CONSOL_AVG` para localizar o valor médio da porta de entrada `SalesTotal`:

```
SalesTotal1:= CONSOL_AVG(SalesTotal, GroupKey)
```

Nesta expressão, a função `CONSOL_AVG` usa a porta `GroupKey` para identificar um grupo de registros. Dentro desse grupo de registros, a função analisa a porta `SalesTotal` e retorna o valor médio. A expressão grava o valor médio na porta de saída `SalesTotal1`.

# CONSOL\_LONGEST

Analisa uma porta no grupo de registros e retorna o valor com a contagem de caracteres mais alta.

## Sintaxe

```
CONSOL_LONGEST(string, group by)
```

A tabela a seguir descreve os argumentos para este comando:

Argumento	Required/Optional	Descrição
<i>string</i>	Obrigatório	Nome da porta de entrada.
<i>agrupar por</i>	Obrigatório	Nome da porta de entrada que contém o identificador de grupo.

## Valor de Retorno

O valor de porta com a contagem de caracteres mais alta.

Se a contagem de caracteres mais alta é compartilhada por dois ou mais valores, a estratégia retorna o primeiro valor qualificado.

## Exemplo

A seguinte expressão usa a função `CONSOL_LONGEST` para analisar a porta de entrada `FirstName` e localizar o valor com a contagem de caracteres mais alta:

```
FirstName1:= CONSOL_LONGEST(FirstName, GroupKey)
```

Nesta expressão, a função `CONSOL_LONGEST` usa a porta `GroupKey` para identificar um grupo de registros. Dentro desse grupo de registros, a função analisa a porta `FirstName` e retorna o valor mais longo. A expressão grava esse valor na porta de saída `FirstName1`.

# CONSOL\_MAX

Analisa uma porta no grupo de registro e retorna o valor mais alto.

## Sintaxe

```
CONSOL_MAX(string, group by)
```

A tabela a seguir descreve os argumentos para este comando:

Argumento	Required/Optional	Descrição
<i>string</i>	Obrigatório	Nome da porta de entrada.
<i>agrupar por</i>	Obrigatório	Nome da porta de entrada que contém o identificador de grupo.

### Valor de Retorno

O valor de porta mais alto.

Para o tipo de dados String, a função retorna a string mais longa. Para o tipo de dados Data/hora, a função retorna a data mais recente.

### Exemplo

A seguinte expressão usa a função `CONSOL_MAX` para analisar a porta de entrada `SalesTotal` e localizar o valor mais alto:

```
SalesTotal1:= CONSOL_MAX(SalesTotal, GroupKey)
```

Nesta expressão, a função `CONSOL_MAX` usa a porta `GroupKey` para identificar um grupo de registros. Dentro desse grupo de registros, a função analisa a porta `SalesTotal` e retorna o valor mais alto. A expressão grava esse valor na porta de saída `SalesTotal1`.

## CONSOL\_MIN

Analisa uma porta no grupo de registros e retorna o valor mais baixo.

### Sintaxe

```
CONSOL_MIN(string, group by)
```

A tabela a seguir descreve os argumentos para este comando:

Argumento	Required/Optional	Descrição
<i>string</i>	Obrigatório	Nome da porta de entrada.
<i>agrupar por</i>	Obrigatório	Nome da porta de entrada que contém o identificador de grupo.

### Valor de Retorno

O valor de porta mais baixo.

Para o tipo de dados String, a função retorna a string mais curta. Para o tipo de dados Data/hora, a função retorna a data mais antiga.

### Exemplo

A seguinte expressão usa a função `CONSOL_MIN` para analisar a porta de entrada `SalesTotal` e localizar o valor mais baixo:

```
SalesTotal1:= CONSOL_MIN(SalesTotal, GroupKey)
```

Nesta expressão, a função `CONSOL_MIN` usa a porta `GroupKey` para identificar um grupo de registros. Dentro desse grupo de registros, a função analisa a porta `SalesTotal` e retorna o valor mais baixo. A expressão grava esse valor na porta de saída `SalesTotal1`.

## CONSOL\_MOSTFREQ

Analisa uma porta no grupo de registros e retorna o valor de ocorrência mais frequente, incluindo valores em branco ou nulos.

### Sintaxe

```
CONSOL_MOSTFREQ(string, group by)
```

A tabela a seguir descreve os argumentos para este comando:

Argumento	Required/Optional	Descrição
<i>string</i>	Obrigatório	Nome da porta de entrada.
<i>agrupar por</i>	Obrigatório	Nome da porta de entrada que contém o identificador de grupo.

### Valor de Retorno

O valor de ocorrência mais frequente, incluindo valores em branco ou nulos.

Se o número mais alto de ocorrências é compartilhado por dois ou mais valores, a estratégia retorna o primeiro valor qualificado.

### Exemplo

A seguinte expressão usa a função `CONSOL_MOSTFREQ` para analisar a porta de entrada `Company` e localizar o valor de ocorrência mais frequente:

```
Company1:= CONSOL_MOSTFREQ(Company, GroupKey)
```

Nesta expressão, a função `CONSOL_MOSTFREQ` usa a porta `GroupKey` para identificar um grupo de registros. Dentro desse grupo de registros, a função analisa a porta `Company` e retorna o valor de ocorrência mais frequente. A expressão grava esse valor na porta de saída `Company1`.

## CONSOL\_MOSTFREQ\_NB

Analisa uma porta no grupo de registros e retorna o valor de ocorrência mais frequente, excluindo valores em branco ou nulos.

### Sintaxe

```
CONSOL_MOSTFREQ_NB(string, group by)
```

A tabela a seguir descreve os argumentos para este comando:

Argumento	Required/Optional	Descrição
<i>string</i>	Obrigatório	Nome da porta de entrada.
<i>agrupar por</i>	Obrigatório	Nome da porta de entrada que contém o identificador de grupo.

### Valor de Retorno

O valor de ocorrência mais frequente, excluindo valores em branco ou nulos.

Se o número mais alto de ocorrências é compartilhado por dois ou mais valores, a estratégia retorna o primeiro valor qualificado.

## Exemplo

A seguinte expressão usa a função `CONSOL_MOSTFREQ_NB` para analisar a porta de entrada `Company` e localizar o valor de ocorrência mais frequente:

```
Company1:= CONSOL_MOSTFREQ_NB(Company, GroupKey)
```

Nesta expressão, a função `CONSOL_MOSTFREQ_NB` usa a porta `GroupKey` para identificar um grupo de registros. Dentro desse grupo de registros, a função analisa a porta `Company` e retorna o valor de ocorrência mais frequente. A expressão grava esse valor na porta de saída `Company1`.

## CONSOL\_SHORTEST

Analisa uma porta no grupo de registros e retorna o valor com a contagem de caracteres mais baixa.

### Sintaxe

```
CONSOL_SHORTEST(string, group by)
```

A tabela a seguir descreve os argumentos para este comando:

Argumento	Required/Optional	Descrição
<i>string</i>	Obrigatório	Nome da porta de entrada.
<i>agrupar por</i>	Obrigatório	Nome da porta de entrada que contém o identificador de grupo.

### Valor de Retorno

O valor de porta com a contagem de caracteres mais baixa.

Se a contagem de caracteres mais baixa é compartilhada por dois ou mais valores, a estratégia retorna o primeiro valor qualificado.

## Exemplo

A seguinte expressão usa a função `CONSOL_SHORTEST` para analisar a porta de entrada `FirstName` e localizar o valor com a contagem de caracteres mais baixa:

```
FirstName1:= CONSOL_SHORTEST(FirstName, GroupKey)
```

Nesta expressão, a função `CONSOL_SHORTEST` usa a porta `GroupKey` para identificar um grupo de registros. Dentro desse grupo de registros, a função analisa a porta `FirstName` e retorna o valor mais curto. A expressão grava esse valor na porta de saída `FirstName1`.

# Funções de Consolidação Baseadas em Linhas

Use as funções de consolidação baseadas em linhas para selecionar um registro em um grupo de registros. Você deve usar as funções de consolidação baseadas em linhas dentro de instruções `IF-THEN-ELSE`.



## CONSOL\_GETROWFIELD

Lê a linha identificada por uma função de consolidação baseada em linhas e retorna o valor da porta especificada. Use um argumento numérico para especificar uma porta.

Você deve usar a função `CONSOL_GETROWFIELD` em conjunto com uma das seguintes funções de consolidação baseadas em linhas:

- `CONSOL_MODEXACT`
- `CONSOL_MOSTDATA`
- `CONSOL_MOSTFILLED`

Para cada porta de entrada em uma função de consolidação baseada em linhas, você deve usar uma instância da função `CONSOL_GETROWFIELD`.

### Sintaxe

```
CONSOL_GETROWFIELD(value)
```

A tabela a seguir descreve os argumentos para este comando:

Argumento	Required/ Optional	Descrição
<i>valor</i>	Obrigatório	Número que indica uma porta de entrada na função de consolidação baseada em linhas. Use "0" para especificar a porta mais à esquerda na função. Use os números subsequentes para indicar outras portas.

### Valor de Retorno

O valor da porta especificada. A função lê esse valor de uma linha identificada por uma função de consolidação baseada em linhas.

### Exemplo

A seguinte expressão usa a função `CONSOL_GETROWFIELD` em conjunto com a função `CONSOL_MOSTDATA`:

```
IF (CONSOL_MOSTDATA(First_Name, Last_Name, GroupKey, GroupKey))
THEN
First_Name1 := CONSOL_GETROWFIELD(0)
Last_Name1 := CONSOL_GETROWFIELD(1)
GroupKey1 := CONSOL_GETROWFIELD(2)
ELSE
First_Name1 := First_Name
Last_Name1 := Last_Name
GroupKey1 := GroupKey
ENDIF
```

Nesta expressão, a função `CONSOL_MOSTDATA` analisa as linhas em um grupo de registros e identifica uma única linha. As funções `CONSOL_GETROWFIELD` usam números consecutivos para ler os valores de porta dessa linha e gravar os valores nas portas de saída.

## CONSOL\_MODEXACT

Identifica a linha com a contagem mais alta dos valores mais frequentes.

Por exemplo, considere uma linha que tem três portas contendo os valores mais frequentes no grupo de registros. A contagem dos valores mais frequentes para essa linha é "3".

Você deve usar essa função em conjunto com a função `CONSOL_GETROWFIELD`. A função `CONSOL_GETROWFIELD` retorna os valores da linha identificada pela função `CONSOL_MODEXACT`.

## Sintaxe

```
CONSOL_MODALSEXACT(string1, [string2, ..., stringN,]  
group by)
```

A tabela a seguir descreve os argumentos para este comando:

Argumento	Required/Optional	Descrição
<i>string</i>	Obrigatório	Nome da porta de entrada.
<i>agrupar por</i>	Obrigatório	Nome da porta de entrada que contém o identificador de grupo.

## Valor de Retorno

TRUE para a linha que tem a pontuação de contagem mais alta dos campos mais frequentes, FALSE para todas as outras linhas.

## Exemplo

A seguinte expressão usa a função `CONSOL_MODALSEXACT` para localizar a linha que contém a contagem mais alta dos campos mais frequentes:

```
IF (CONSOL_MODALSEXACT(First_Name,Last_Name,GroupKey,GroupKey))  
THEN  
First_Name1 := CONSOL_GETROWFIELD(0)  
Last_Name1 := CONSOL_GETROWFIELD(1)  
GroupKey1 := CONSOL_GETROWFIELD(2)  
ELSE  
First_Name1 := First_Name  
Last_Name1 := Last_Name  
GroupKey1 := GroupKey  
ENDIF
```

Nesta expressão, a função `CONSOL_MODALSEXACT` analisa as linhas em um grupo de registros e identifica uma única linha. As funções `CONSOL_GETROWFIELD` usam números consecutivos para ler os valores de porta dessa linha e gravar os valores nas portas de saída.

# CONSOL\_MOSTDATA

Identifica a linha que contém mais caracteres de todas as portas

Você deve usar essa função em conjunto com a função `CONSOL_GETROWFIELD`. A função `CONSOL_GETROWFIELD` retorna os valores da linha identificada pela função `CONSOL_MOSTDATA`.

## Sintaxe

```
CONSOL_MOSTDATA(string1, [string2, ..., stringN,]  
group by)
```

A tabela a seguir descreve os argumentos para este comando:

Argumento	Required/Optional	Descrição
<i>string</i>	Obrigatório	Nome da porta de entrada.
<i>agrupar por</i>	Obrigatório	Nome da porta de entrada que contém o identificador de grupo.

## Valor de Retorno

TRUE para a linha que contém mais caracteres de todas as portas, FALSE para todas as outras linhas.

## Exemplo

A seguinte expressão usa a função `CONSOL_MOSTDATA` para localizar a linha que contém mais caracteres:

```
IF (CONSOL_MOSTDATA(First_Name,Last_Name,GroupKey,GroupKey))
THEN
First_Name1 := CONSOL_GETROWFIELD(0)
Last_Name1 := CONSOL_GETROWFIELD(1)
GroupKey1 := CONSOL_GETROWFIELD(2)
ELSE
First_Name1 := First_Name
Last_Name1 := Last_Name
GroupKey1 := GroupKey
ENDIF
```

Nesta expressão, a função `CONSOL_MOSTDATA` analisa as linhas em um grupo de registros e identifica uma única linha. As funções `CONSOL_GETROWFIELD` usam números consecutivos para ler os valores de porta dessa linha e gravar os valores nas portas de saída.

## CONSOL\_MOSTFILLED

Identifica a linha que contém o número mais alto de campos não em branco.

Você deve usar essa função em conjunto com a função `CONSOL_GETROWFIELD`. A função `CONSOL_GETROWFIELD` retorna os valores da linha identificada pela função `CONSOL_MOSTFILLED`.

### Sintaxe

```
CONSOL_MOSTFILLED(string1, [string2, ..., stringN,]
group by)
```

A tabela a seguir descreve os argumentos para este comando:

Argumento	Required/Optional	Descrição
<i>string</i>	Obrigatório	Nome da porta de entrada.
<i>agrupar por</i>	Obrigatório	Nome da porta de entrada que contém o identificador de grupo.

### Valor de Retorno

TRUE para a linha que contém o número mais alto de campos não em branco, FALSE para todas as outras linhas.

## Exemplo

A seguinte expressão usa a função `CONSOL_MOSTFILLED` para localizar a linha que contém mais caracteres:

```
IF (CONSOL_MOSTFILLED(First_Name,Last_Name,GroupKey,GroupKey))
THEN
First_Name1 := CONSOL_GETROWFIELD(0)
Last_Name1 := CONSOL_GETROWFIELD(1)
GroupKey1 := CONSOL_GETROWFIELD(2)
ELSE
First_Name1 := First_Name
Last_Name1 := Last_Name
GroupKey1 := GroupKey
ENDIF
```

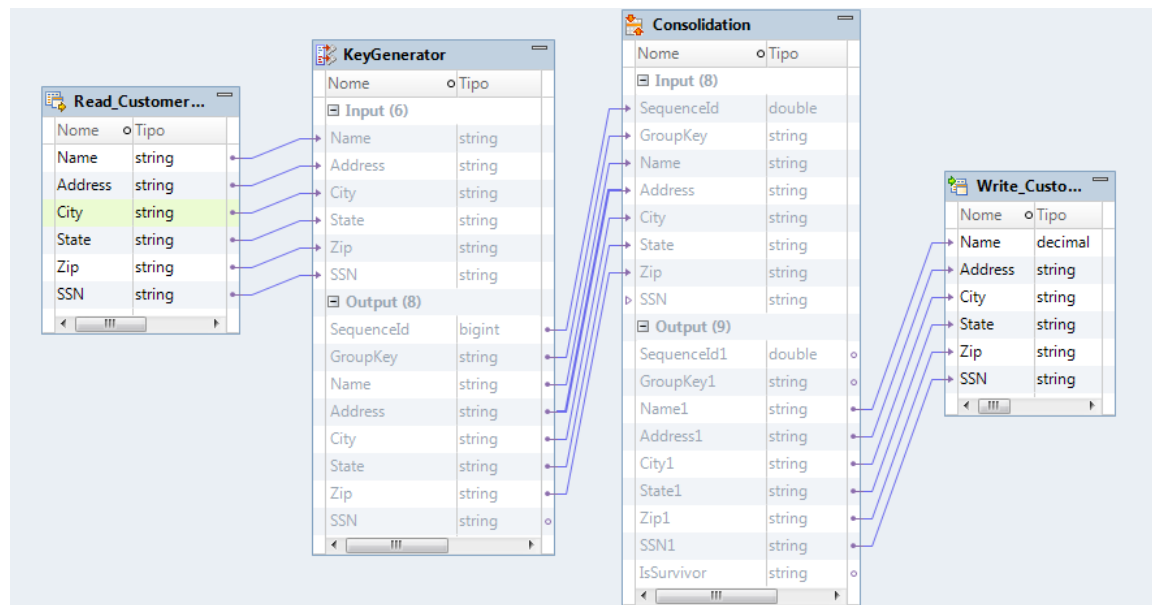
Nesta expressão, a função `CONSOL_MOSTFILLED` analisa as linhas em um grupo de registros e identifica uma única linha. As funções `CONSOL_GETROWFIELD` usam números consecutivos para ler os valores de porta dessa linha e gravar os valores nas portas de saída.

## Exemplo de Mapeamento de Consolidação

Sua organização precisa consolidar registros duplicados de clientes. Para consolidar os registros de clientes, você agrupa os dados com uma transformação de Gerador de Chaves e usar a transformação de Consolidação para consolidar os registros.

Crie um mapeamento com uma fonte de dados contendo registros de clientes, uma transformação de Gerador de Chaves, uma transformação de Consolidação e um destino de dados. O mapeamento agrupa os registros de clientes, consolida os grupos e grava um único registro consolidado.

A seguinte figura mostra o mapeamento:



### Dados de Entrada

Os dados de entrada que você quer analisar contêm informações dos clientes.

A seguinte tabela contém os dados de entrada para este exemplo:

Nome	Endereço	Cidade	Estado	CEP	SSN
Dennis Jones	100 All Saints Ave	Nova York	NY	10547	987-65-4320
Dennis Jones	1000 Alberta Rd	Nova York	NY	10547	987-65-4320
D Jones	100 All Saints Ave	Nova York	NY	10547-1521	

### Transformação de Gerador de Chaves

Use a transformação de Gerador de Chaves para agrupar os dados de entrada com base na porta de código postal.

A transformação retorna os seguintes dados:

SequenceId	GroupKey	Nome	Endereço	Cidade	Estado	Código Postal	SSN
1	10547	Dennis Jones	100 All Saints Ave	Nova York	NY	10547	987-65-4320
2	10547	Dennis Jones	1000 Alberta Rd	Nova York	NY	10547	
3	10547	D Jones	100 All Saints Ave	Nova York	NY	10547-1521	987-65-4320

## Transformação de Consolidação

Use a transformação de Consolidação para gerar o registro consolidado.

Configure a transformação de Consolidação para usar o tipo de estratégia baseada em linhas. Selecione a estratégia Modal Exato para escolher a linha com a contagem mais alta dos valores mais frequentes. A estratégia Modal Exato usa os valores dessa linha para gerar um registro consolidado. O registro consolidado é aquele com o valor "Y" na porta IsSurvivor.

A transformação retorna os seguintes dados:

GroupKey	Nome	Endereço	Cidade	Estado	Código Postal	SSN	IsSurvivor
10547	Dennis Jones	100 All Saints Ave	Nova York NY	10547	987-65-4320	N	
10547	Dennis Jones	1000 Alberta Rd	Nova York NY	10547		N	
10547	D Jones	100 All Saints Ave	Nova York NY	10547-1521	987-65-4320	N	
10547	D Jones	100 All Saints Ave	Nova York NY	10547-1521	987-65-4320	Y	

## Saída do Mapeamento de Consolidação

Configure a transformação de Consolidação de forma que a saída do mapeamento contenha apenas registros consolidados.

Neste exemplo, você está razoavelmente confiante de que os valores mais frequentes selecionados pela estratégia Modal Exato são os valores de porta corretos. Para gravar apenas registros consolidados no destino de mapeamento, selecione a exibição **Avançado** e defina o modo de saída como "Somente Sobreviventes".

Quando você executar o mapeamento, a saída do mapeamento conterá apenas registros consolidados.

## Configurando uma Transformação de Consolidação

Ao configurar a transformação de Consolidação, você escolhe os tipos de estratégia, escolhe estratégias ou grava expressões, seleciona uma porta de agrupamento e configura opções avançadas.

1. Selecione a exibição **Consolidação**.
2. Escolha um tipo de estratégia.
3. Configure a estratégia.
  - Para o tipo de estratégia simples, selecione uma estratégia para cada porta.
  - Para o tipo de estratégia baseada em linhas, selecione uma estratégia.
  - Para o tipo de estratégia avançada, crie uma expressão que use funções de consolidação.
4. No campo Agrupar por, selecione a porta que contém o identificador de grupo.
5. Ative a classificação na exibição **Avançado** se os dados de entrada não estiverem classificados.
6. Configure a saída para conter registros consolidados ou todos os registros.

# Transformação de Consolidação em um ambiente não nativo

O processamento da transformação de Validador de Endereço em um ambiente não nativo depende do mecanismo que executa a transformação.

Considere o suporte para os seguintes mecanismos de tempo de execução não nativos:

- Mecanismo Blaze. Suportado com restrições.
- Mecanismo Spark. Com suporte com restrições em mapeamentos em lote. Sem suporte em mapeamentos de fluxo.
- Mecanismo Databricks Spark. Sem suporte.

## Transformação de Consolidação no mecanismo Blaze

A transformação de Consolidação é suportada com as seguintes restrições:

- A transformação pode processar registros em uma ordem diferente em cada ambiente.
- A transformação pode identificar um registro diferente como o sobrevivente em cada ambiente.

## Transformação de Consolidação no mecanismo Spark

A transformação de Consolidação é suportada com as seguintes restrições:

- A transformação pode processar registros em uma ordem diferente em cada ambiente.
- A transformação pode identificar um registro diferente como o sobrevivente em cada ambiente.

## Transformação de Consolidação no mecanismo Databricks Spark

A transformação de Consolidação em um ambiente não nativo depende do mecanismo que executa a transformação.

Considere o suporte para os seguintes mecanismos de tempo de execução não nativos:

- Mecanismo Databricks Spark. Sem suporte.

## CAPÍTULO 12

# Transformação de Mascaramento de Dados

Este capítulo inclui os seguintes tópicos:

- [Transformação de Mascaramento de Dados - Visão Geral, 203](#)
- [Técnicas de Mascaramento, 204](#)
- [Regras de Mascaramento, 215](#)
- [Formatos Especiais de Máscara, 220](#)
- [Arquivo de Valor Padrão, 224](#)
- [Configuração da Transformação de Mascaramento de Dados, 224](#)
- [Propriedades de Tempo de Execução da Transformação de Mascaramento de Dados, 226](#)
- [Exemplo de Mascaramento de Dados, 227](#)
- [Propriedades Avançadas da Transformação de Mascaramento de Dados, 230](#)
- [Transformação de Mascaramento de dados em um ambiente não nativo, 230](#)

## Transformação de Mascaramento de Dados - Visão Geral

A transformação de Mascaramento de Dados converte dados de produção confidenciais em dados de teste realistas para ambientes que não são de produção. A transformação de Mascaramento de Dados modifica os dados de origem com base em técnicas de mascaramento que você configura para cada coluna.

Crie dados mascarados para desenvolvimento, testes e treinamento de software e mineração de dados. É possível manter os relacionamentos de dados nos dados mascarados e manter a integridade referencial entre tabelas de bancos de dados.

A transformação de Mascaramento de Dados fornece regras de mascaramento com base no tipo de dados de origem e na técnica de mascaramento que você configura para uma coluna. No caso de strings, é possível restringir quais caracteres em uma string serão substituídos. Você pode restringir quais caracteres aplicar em uma máscara. No caso de números e datas, é possível fornecer uma variedade de números para os dados mascarados. É possível configurar uma variedade fixa ou de diferença percentual do número original. O Data Integration Service substitui caracteres com base na localidade que você configura para a transformação.

# Técnicas de Mascaramento

O tipo de mascaramento é o tipo de mascaramento de dados a ser aplicado à coluna selecionada.

Você pode selecionar uma das seguintes técnicas de mascaramento para uma coluna de entrada:

## **Aleatório**

Gera resultados aleatórios e não repetíveis para os mesmos dados de origem e regras de mascaramento. Você pode mascarar tipos de dados de data, numéricos e de string. O mascaramento aleatório não precisa de um valor de semente. Os resultados do mascaramento aleatório são não determinísticos.

## **Expressão**

Aplica uma expressão a uma coluna de origem para criar ou mascarar dados. Você pode mascarar todos os tipos de dados.

## **Chave**

Substitui os dados de origem por valores repetíveis. A transformação de Mascaramento de Dados produz resultados determinísticos para os mesmos dados de origem, regras de mascaramento e valor de semente. Você pode mascarar tipos de dados de data, numéricos e de string.

## **Substituição**

Substitui uma coluna de dados por dados semelhantes, mas não relacionados, de um dicionário. Você pode mascarar o tipo de dados de string.

## **Dependentes**

Substitui os valores de uma coluna de origem com base nos valores de outra coluna de origem. Você pode mascarar o tipo de dados de string.

## **Tokenização**

Substitui dados de origem por dados gerados com base em critérios de mascaramento personalizados. A transformação de Mascaramento de Dados aplica regras especificadas em um algoritmo personalizado. Você pode mascarar o tipo de dados de string.

## **Formatos Especiais de Máscara**

Número de cartão de crédito, endereço de e-mail, endereço IP, telefone, número de Seguridade Social, número de Seguro Social ou URL. A transformação de Mascaramento de Dados aplica regras internas para mascarar de forma inteligente esses tipos comuns de dados confidenciais.

## **Sem Mascaramento**

A transformação de Mascaramento de Dados não altera os dados de origem.

O padrão é Sem Mascaramento.

## Mascaramento Aleatório

O mascaramento aleatório gera dados mascarados aleatórios não determinísticos. A transformação de Mascaramento de Dados retorna diferentes valores quando o mesmo valor de origem ocorre em diferentes linhas. Você pode definir regras de mascaramento que afetam o formato de dados retornado pela transformação de Mascaramento de Dados. Mascare valores numéricos, de string e de data com um mascaramento aleatório.

## Mascaramento de Valores de String



Configure mascaramento aleatório para gerar saídas aleatórias para colunas de string. Para configurar os limites de cada caractere na string de saída, configure um formato de máscara. Configure caracteres de filtro para definir quais caracteres de origem devem ser mascarados e os caracteres com os quais se mascara esses caracteres de origem.

É possível aplicar as seguintes regras de mascaramento para uma porta de string:

#### **Faixa**

Configure a extensão mínima e máxima da string. A Transformação de Mascaramento de Dados retorna uma string de caracteres aleatórios entre as extensões mínima e máxima de string.

#### **Formato de Máscara**

Para cada caractere dos dados de entrada, defina o tipo de caractere a ser substituído. É possível limitar cada caractere a um tipo de caractere alfabético, numérico ou alfanumérico.

#### **Caracteres da String de Origem**

Defina os caracteres na string de origem que você deseja mascarar. Por exemplo, mascare o caractere de sinal de número (#) sempre que ocorrer nos dados de entrada. A Transformação de Mascaramento de Dados mascara todos os caracteres de entrada quando os caracteres da string de origem estiverem em branco.

#### **Caracteres de Substituição da String de Resultados**

Substitua os caracteres na string de destino pelos caracteres que você definir em Caracteres da String de Resultado. Por exemplo, insira os caracteres a seguir para configurar cada máscara para que contenha caracteres alfabéticos maiúsculos de A a Z:

ABCDEFGHIJKLMNOPQRSTUVWXYZ

## **Mascaramento de Valores Numéricos**

Ao mascarar dados numéricos, é possível configurar uma série de valores de saída de uma coluna. A Transformação de Mascaramento de Dados retorna um valor entre os valores mínimo e máximo do intervalo, dependendo da precisão da porta. Para definir a série, configure os intervalos mínimo e máximo ou um intervalo de embaçamento com base em uma variância a partir do valor de origem primário.

Você pode configurar os seguintes parâmetros de mascaramento para dados numéricos:

#### **Faixa**

Defina um intervalo de valores de saída. A Transformação de Mascaramento de Dados retorna dados numéricos entre os valores mínimo e máximo.

#### **Faixa de Embaçamento**

Defina um intervalo de valores de saída que estejam dentro de uma variância fixa ou uma variância de porcentagem dos dados de origem. A Transformação de Mascaramento de Dados retorna dados numéricos que estão próximos ao valor dos dados de origem. É possível configurar um intervalo e um intervalo de embaçamento.

## **Mascaramento de Valores de Data**

Para mascarar valores de data com o mascaramento aleatório, configure um intervalo de datas de saída ou escolha uma variância. Quando configurar uma variância, escolha uma parte da data para desfocar. Escolha o ano, mês, dia, hora, minuto ou segundo. A transformação de Mascaramento de Dados retorna uma data que esteja dentro do intervalo que você configurou.

Você pode configurar as seguintes regras de mascaramento quando mascarar um valor datetime:

### Faixa

Configura os valores mínimo e máximo para retornar, em referência ao valor datetime selecionado.

### Embaçamento

Mascara uma data com base em uma variância que você aplica a uma unidade da data. A Transformação de Mascaramento de Dados retorna uma data dentro de uma variância. Você pode embaçar o ano, mês, dia, hora, minuto ou segundo. Escolha uma variância baixa e alta para aplicar.

## Mascaramento de Expressão

O mascaramento de expressão aplica uma expressão a uma porta para alterar os dados ou criar novos dados. Ao configurar o mascaramento de expressão, crie uma expressão no Editor de Expressão. Selecione portas de entrada e de saída, funções, variáveis e operadores para criar expressões.

É possível concatenar dados de várias portas para criar um valor para outra porta. Por exemplo, é necessário criar um nome de logon. A origem tem colunas de nome e de sobrenome. Mascare o nome e o sobrenome dos arquivos de pesquisa. Na Transformação de Mascaramento de Dados, crie outra porta chamada Logon. Para a porta de Logon, configure uma expressão para concatenar a primeira letra do nome com a primeira letra do sobrenome:

```
SUBSTR (FIRSTNM, 1, 1) || LASTNM
```

Selecione funções, portas, variáveis e operadores na interface de apontar e clicar para minimizar erros ao construir expressões.

O Editor de Expressão exibe as portas de saída que não estão configuradas para mascaramento de expressão. Você não pode usar a saída de uma expressão como entrada de outra. Se você adicionar o nome da porta de saída manualmente à expressão, pode obter resultados inesperados.

Ao criar uma expressão, verifique se ela retorna um valor que combine com o tipo de dados da porta. A Transformação de Mascaramento de Dados retorna zero se o tipo de dados da porta de expressão for numérico e o tipo de dados da expressão não for o mesmo. A Transformação de Mascaramento de Dados retorna um valor nulo se o tipo de dados da porta de expressão for uma string e o tipo de dados da expressão não for o mesmo.

## Mascaramento Repetível de Expressão

Configure o mascaramento repetível de expressão quando uma coluna de origem ocorrer em mais de uma tabela e você precisar mascarar-la com o mesmo valor.

Quando você configura o mascaramento de expressão repetível, a transformação de Mascaramento de Dados salva os resultados de uma expressão em uma tabela de armazenamento. Se a coluna ocorrer em outra tabela de origem, a transformação de Mascaramento de Dados retornará o valor mascarado da tabela de armazenamento e não da expressão.

## Nome do Dicionário

Ao configurar um mascaramento de expressão repetível, é necessário inserir um nome de dicionário. O nome do dicionário é uma chave que permite que várias Transformações de Mascaramento de Dados gerem os mesmos valores mascarados a partir dos mesmos valores de origem. Defina o mesmo nome de dicionário em cada Transformação Mascaramento de Dados. O nome do dicionário pode ser qualquer texto.

## Tabela de Armazenamento

A tabela de armazenamento contém os resultados do mascaramento de expressão repetível entre sessões. Uma linha da tabela de armazenamento contém a coluna de origem e um par de valores mascarados. A

tabela de armazenamento para mascaramento de expressão é separada da tabela de armazenamento para mascaramento de substituição.

Toda vez que a transformação de Mascaramento de Dados mascara um valor com uma expressão repetível, ela pesquisa a tabela de armazenamento pelo nome de dicionário, local, nome de coluna e valor de entrada. Se ela encontrar uma linha na tabela de armazenamento, ela retorna o valor mascarado da tabela de armazenamento. Se a transformação de Mascaramento de Dados não encontrar uma linha, ela gerará um valor mascarado da expressão para a coluna.

É necessário criptografar as tabelas de armazenamento para mascaramento de expressão quando você tiver dados descritos no armazenamento e usar o mesmo nome de dicionário como chave.

## Criptografando Tabelas de Armazenamento para Mascaramento de Expressão

Você pode usar funções de codificação de linguagem de transformação para criptografar tabelas de armazenamento. É necessário criptografar tabelas de armazenamento quando você tiver ativado criptografia de armazenamento.

1. Crie um mapeamento com a tabela de armazenamento IDM\_EXPRESSION\_STORAGE como origem.
2. Crie uma transformação de Mascaramento de Dados.
3. Aplique a técnica de mascaramento de expressão nas portas de valor mascarado.
4. Use a seguinte expressão na porta MASKEDVALUE:

```
Enc_Base64 (AES_Encrypt (MASKEDVALUE, Key) )
```

5. Vincule as portas ao destino.

## Exemplo

Por exemplo, a tabela Funcionários contém as colunas a seguir:

```
FirstName  
LastName  
LoginID
```

Na transformação de Mascaramento de Dados, mascare LoginID com uma expressão que combine FirstName e LastName. Configure a máscara da expressão para ser repetível. Insira um nome de dicionário como uma chave de mascaramento repetível.

A tabela Computer\_Users contém uma coluna LoginID, mas não a FirstName ou LastName:

```
Dept  
LoginID  
Password
```

Para mascarar LoginID em Computer\_Users com o mesmo LoginID que Funcionários, configure o mascaramento de expressão para a coluna LoginID. Ative o mascaramento repetível e insira o mesmo nome de dicionário que você definiu para a tabela Funcionários em LoginID. O Serviço de Integração recupera os valores LoginID da tabela de armazenamento.

Crie uma expressão padrão para usar quando o Serviço de Integração não puder encontrar uma linha na tabela de armazenamento para LoginID. A tabela Computer\_Users não tem as colunas FirstName ou LastName, portanto a expressão cria um LoginID menos significativo.

## Script de Tabela de Armazenamento

A Informatica fornece scripts que podem ser executados para criar a tabela de armazenamento. Os scripts estão no seguinte local:

```
<PowerCenter installation directory>\client\bin\Extensions\DataMasking
```

O diretório contém um script para bancos de dados Sybase, Microsoft SQL Server, IBM DB2 e Oracle. Cada script recebe o nome <Expression\_<database type>.

## Regras e Diretrizes para Mascaramento de Expressão

Use as seguintes regras e diretrizes para mascaramento de expressão:

- Você não pode usar a saída de uma expressão como entrada de outra. Se você adicionar o nome da porta de saída manualmente à expressão, pode obter resultados inesperados.
- Utilize o método de apontar e clicar para construir expressões. Selecione funções, portas, variáveis e operadores na interface de apontar e clicar para minimizar erros ao construir expressões.
- Se a transformação de Mascaramento de Dados estiver configurada para mascaramento repetido, e a tabela de armazenamento não existir, o Serviço de Integração substituirá os dados de origem por valores padrão.

## Mascaramento de Chave

Uma coluna configurada para mascaramento de chaves retorna dados mascarados determinísticos sempre que o valor de origem e o valor de semente são iguais. A Transformação Mascaramento de Dados retorna valores exclusivos para a coluna.

Quando você configura uma coluna para mascaramento de chaves, a transformação de Mascaramento de Dados cria um valor de semente para a coluna. É possível alterar o valor de semente para produzir dados repetíveis entre diferentes Transformações de Mascaramento de Dados. Por exemplo, configure mascaramento de chave para impor integridade referencial. Use o mesmo valor de semente para mascarar uma chave primária em uma tabela e o valor de chave externa em outra tabela.

Você pode definir regras de mascaramento que afetam o formato de dados retornado pela transformação de Mascaramento de Dados. Mascare valores de string e numéricos com mascaramento de chave.

## Valores de String de Mascaramento

Você pode configurar o mascaramento de chaves a fim de gerar saída repetível para as strings. Configure um formato de máscara para definir limitações para cada caractere na string de saída. Configure os caracteres da string de origem que definem quais caracteres de origem devem ser mascarados. Configure caracteres de substituição da string de resultados para limitar os dados mascarados a determinados caracteres.

É possível configurar as seguintes regras de mascaramento para strings de mascaramento de chaves:

### Semente

Aplique um valor de semente para gerar dados mascarados determinísticos para uma coluna. Você pode inserir um número entre 1 e 1.000.

### Formato de Máscara

Defina o tipo de caractere a ser substituído para cada caractere dos dados de entrada. É possível limitar cada caractere a um tipo de caractere alfabético, numérico ou alfanumérico.

### Caracteres da String de Origem

Defina os caracteres na string de origem que você deseja mascarar. Por exemplo, mascare o caractere de sinal de número (#) sempre que ele ocorrer nos dados de entrada. A transformação de Mascaramento de Dados mascara todos os caracteres de entrada quando os Caracteres da String de Origem estão em branco. A transformação de Mascaramento de Dados nem sempre retornará dados exclusivos se o número de caracteres da string de origem for inferior ao número de caracteres da string de resultados.

### Caracteres da String de Resultados

Substitua os caracteres da string de destino pelos caracteres que você definir em Caracteres da String de Resultados. Por exemplo, insira os seguintes caracteres para configurar cada máscara para que contenha todos os caracteres alfabéticos em maiúsculo:

ABCDEFGHIJKLMNOPQRSTUVWXYZ

## Mascaramento de Valores Numéricos

Configure o mascaramento de chave para os dados de origem numéricos, para gerar a saída determinística. Quando você configurar uma coluna para o mascaramento de chaves numéricas, atribua um valor de semente aleatório para a coluna. Quando a transformação de Mascaramento de Dados mascara os dados de origem, ela aplica um algoritmo de mascaramento que exige a semente.

Você pode alterar o valor da semente para uma coluna, a fim de produzir resultados repetíveis se o mesmo valor de origem ocorrer em uma coluna diferente. Por exemplo, você deseja manter uma relação de chave externa primária entre duas tabelas. Em cada transformação de Mascaramento de Dados, insira o mesmo valor de semente para a coluna de chave primária, que para a coluna de chave externa. A transformação de Mascaramento de Dados produz resultados determinísticos para os mesmos valores numéricos. A integridade referencial é mantida entre as tabelas.

## Mascarando Valores de Data e Hora

Quando você pode configurar o mascaramento de chaves para valores de data e hora, a transformação de Mascaramento de Dados requer um número aleatório como semente. Você pode alterar a semente para corresponder ao valor de semente de outra coluna, a fim de retornar valores datetime repetíveis entre as colunas.

A transformação de Mascaramento de Dados pode mascarar datas entre 1753 e 2400 com o mascaramento de chaves. Se o ano da origem for bissexto, a transformação de mascaramento de dados retorna um ano que também é bissexto. Se o mês da origem contém 31 dias, a transformação de mascaramento de dados retorna um mês que tenha 31 dias. Se o mês da origem for fevereiro, a transformação de mascaramento de dados retorna fevereiro.

A transformação de mascaramento de dados sempre gera datas válidas.

## Mascaramento de Substituição

O mascaramento de substituição substitui uma coluna de dados por dados semelhantes, mas não relacionados. Use o mascaramento de substituição para substituir dados de produção por dados de teste realistas. Quando você configurar o mascaramento de substituição, defina o dicionário que contém os valores substitutos.

A transformação de Mascaramento de Dados realizará uma pesquisa no dicionário que você configurar. A transformação de Mascaramento de Dados substitui dados da origem por dados do dicionário. Arquivos de

dicionário podem conter dados de string, valores de data e hora, inteiros e números de ponto flutuante. Insira valores de data e hora no seguinte formato:

mm/dd/yyyy

Você pode substituir dados com valores repetíveis ou não repetíveis. Se você escolher valores repetíveis, a transformação de Mascaramento de Dados produzirá resultados deterministas para dados da mesma origem e valor de semente. Você deve configurar um valor de semente para substituir dados por resultados deterministas. O Serviço de Integração mantém uma tabela de armazenamento de valores de origem mascarados para mascaramento repetível.

Você pode substituir mais de uma coluna de dados por valores mascarados da mesma linha de dicionário. Configure o mascaramento de substituição para uma coluna de entrada. Configure o mascaramento de dados dependentes para as outras colunas que recebem dados mascarados da mesma linha do dicionário.

## Dicionários

Um dicionário é uma tabela de referência que contém os dados substitutos e um número de série para cada linha na tabela. Crie uma tabela de referência para mascaramento de substituição a partir de um arquivo simples ou uma tabela relacional importada para o repositório do Modelo.

A transformação de Mascaramento de Dados gera um número para recuperar uma linha de dicionário pelo número de série. A transformação de Mascaramento de Dados gera uma chave de hash para mascaramento de substituição repetível ou um número aleatório para mascaramento não repetível. Você poderá configurar uma condição de pesquisa adicional se configurar o mascaramento de substituição repetível.

Você pode configurar um dicionário para mascarar mais de uma porta na transformação de Mascaramento de Dados.

Quando a transformação de Mascaramento de Dados recupera dados de substituição de um dicionário, ela não verifica se o valor dos dados substitutos é igual ao valor original. Por exemplo, a transformação de Mascaramento de Dados pode substituir o nome John pelo mesmo nome (John) a partir de um arquivo de dicionário.

O seguinte exemplo mostra uma tabela de dicionário que contém nome e sexo:

SNO	GENDER	FIRSTNAME
1	M	Adam
2	M	Adeel
3	M	Adil
4	F	Alice
5	F	Alison

Neste dicionário, o primeiro campo da linha é o número de série e o segundo é o gênero. O Serviço de Integração sempre procura um registro do dicionário pelo número de série. Você pode adicionar o sexo como uma condição de pesquisa se configurar o mascaramento repetível. O Serviço de Integração recupera uma linha do dicionário usando uma chave hash, e encontra uma linha com um sexo que corresponde ao sexo nos dados da origem.

Utilize as seguintes regras e diretrizes ao criar uma tabela de referência:

- Cada registro na tabela deve ter um número de série.

- Os números de série são inteiros sequenciais, começando no um. Os números de série não podem ter um número ausente na sequência.
- A coluna do número de série pode estar em qualquer local de uma linha da tabela. Ela pode ter qualquer rótulo.

Se você usar uma tabela de arquivo simples para criar a tabela de referência, use as seguintes regras e diretrizes:

- A primeira linha da tabela de arquivo simples deve ter rótulos de coluna para identificar os campos em cada registro. Os campos estão separados por vírgulas. Se a primeira linha não contiver rótulos de coluna, o Serviço de Integração usa os valores dos campos na primeira linha como nomes de coluna.
- Se você criar uma tabela de arquivo simples no Windows e copiá-la para uma máquina UNIX, verifique se o formato do arquivo está correto para o UNIX. Por exemplo, o Windows e o UNIX usam caracteres diferentes para o marcador de final de linha.

## Tabelas de Armazenamento

A transformação de Mascaramento de Dados mantém tabelas de armazenamento para substituição repetida entre sessões. Uma linha da tabela de armazenamento contém a coluna de origem e um par de valores mascarados. Toda vez que a transformação de Mascaramento de Dados mascara um valor com um valor substituto repetível, ela pesquisa a tabela de armazenamento por nome de dicionário, local, nome de coluna, valor de entrada e semente. Se ela encontrar uma linha, ela retorna o valor mascarado da tabela de armazenamento. Se a transformação de Mascaramento de Dados não encontrar uma linha, ela recuperará uma linha do dicionário com uma chave de hash.

O formato de nome do dicionário na tabela de armazenamento não é o mesmo para um dicionário de arquivo simples e um dicionário relacional. O nome do dicionário de um arquivo simples é identificado pelo nome do arquivo. Já o nome do dicionário relacional tem a seguinte sintaxe:

```
<Connection object>_<dictionary table name>
```

O Informatica fornece scripts que você pode executar para criar uma tabela de armazenamento relacional. Os scripts estão no seguinte local:

```
<PowerCenter Client installation directory>\client\bin\Extensions\DataMasking
```

O diretório contém um script para bancos de dados Sybase, Microsoft SQL Server, IBM DB2 e Oracle. Cada script recebe o nome Substitution\_<tipo do banco de dados>. Você poderá criar uma tabela em um banco de dados diferente se configurar as instruções SQL e as restrições de chaves primárias.

É necessário criptografar as tabelas de armazenamento para mascaramento de substituição quando você tiver dados descriptografados no armazenamento e usar o mesmo valor de semente e dicionário para criptografar as mesmas colunas.

## Criptografando Tabelas de Armazenamento para Mascaramento de Substituição

Você pode usar funções de codificação de linguagem de transformação para criptografar tabelas de armazenamento. É necessário criptografar tabelas de armazenamento quando você tiver ativado criptografia de armazenamento.

1. Crie um mapeamento com a tabela de armazenamento IDM\_SUBSTITUTION\_STORAGE como origem.
2. Crie uma transformação de Mascaramento de Dados.
3. Aplique a técnica de mascaramento de substituição nas portas de valor de entrada e de valor mascarado.
4. Use a seguinte expressão na porta INPUTVALUE:

```
Enc_Base64(AES_Encrypt(INPUTVALUE, Key))
```

5. Use a seguinte expressão na porta MASKEDVALUE:

```
Enc_Base64(AES_Encrypt(MASKEDVALUE, Key))
```

6. Vincule as portas ao destino.

## Propriedades do Mascaramento de Substituição

Você pode configurar estas regras de mascaramento para o mascaramento de substituição:

- **Saída Repetível.** Retorna resultados deterministas entre sessões. A transformação de Mascaramento de Dados armazena valores mascarados na tabela de armazenamento.
- **Valor de Semente.** Aplique um valor semente para gerar dados mascarados determinísticos para uma coluna. Digite um número entre 1 e 1.000.
- **Saída Exclusiva.** Força a transformação de Mascaramento de Dados a criar valores de saída exclusivos para valores de entrada exclusivos. Não há dois valores de entrada mascarados para o mesmo valor de saída. O dicionário deve ter linhas exclusivas suficientes para ativar a saída exclusiva. Quando você desativa a saída exclusiva, a transformação de Mascaramento de Dados pode não mascarar valores de entrada para valores de saída exclusivos. O dicionário pode conter menos linhas.
- **Porta Exclusiva.** A porta usada para identificar registros exclusivos para mascaramento de substituição. Por exemplo, você deseja mascarar nomes em uma tabela chamada Cliente. Se você selecionar a coluna da tabela que contém os nomes como a porta exclusiva, a transformação de Mascaramento de Dados substituirá os nomes duplicados pelo mesmo valor mascarado. Se você selecionar a coluna Customer\_ID como a porta exclusiva, a transformação de Mascaramento de Dados substituirá cada nome por um valor exclusivo.
- **Informações do Dicionário.** Configure a tabela de referência que contém os valores de dados substitutos. Clique em **Selecionar Origem** para selecionar uma tabela de referência.
  - **Nome do Dicionário.** Exibe o nome da tabela de referência que você selecionar.
  - **Coluna Saída.** Escolha a coluna para retornar à transformação de Mascaramento de Dados.
- **Condição de pesquisa.** Configure uma condição de pesquisa para qualificar ainda mais a linha do dicionário a ser usada no mascaramento de substituição. A condição de pesquisa é semelhante à cláusula WHERE em uma consulta SQL. Ao configurar uma condição de pesquisa, você compara o valor de colunas na origem ao de colunas no dicionário.

Por exemplo, você quer mascarar o nome. Os dados na origem e no dicionário têm uma coluna de nomes e outra de sexo. Você pode adicionar uma condição segundo a qual cada nome feminino seja substituído por outro nome feminino do dicionário. A condição de pesquisa compara o gênero na origem ao gênero no dicionário.

  - **Porta de entrada.** Coluna de dados da origem a ser usada na pesquisa.
  - **Coluna do Dicionário.** A coluna do dicionário a ser comparada à porta de entrada.

## Regras e Diretrizes para Mascaramento de Substituição

Use as seguintes regras e diretrizes para mascaramento de substituição:

- Se uma tabela de armazenamento não existir para uma máscara de substituição repetível exclusiva, a sessão falhará.
- Se o dicionário não contiver linhas, a transformação de Mascaramento de Dados retorna uma mensagem de erro.



- Quando a transformação de Mascaramento de Dados encontra um valor de entrada com o local, o dicionário e a semente na tabela de armazenamento, ela recupera o valor mascarado, mesmo que a linha não esteja mais no dicionário.
- Se você excluir um objeto de conexão ou modificar o dicionário, trunque a tabela de armazenamento. Caso contrário, talvez obtenha resultados inesperados.
- Se o número de valores no dicionário for menor que o número de valores exclusivos nos dados de origem, a transformação de Mascaramento de Dados não poderá mascarar os dados com valores repetíveis exclusivos. A transformação de Mascaramento de Dados retorna uma mensagem de erro.

## Mascaramento dependente

O mascaramento dependente substitui várias colunas de dados de origem por dados da mesma linha de dicionário.

Quando a transformação de Mascaramento de Dados executa o mascaramento de substituição para várias colunas, os dados mascarados podem conter combinações de campos irreais. Você pode configurar o mascaramento dependente para substituir os dados para várias colunas de entrada da mesma linha de dicionário. Os dados mascarados recebem combinações válidas, como "Nova York, Nova York" ou "Chicago, Illinois".

Ao configurar o mascaramento dependente, você deve primeiro configurar uma coluna de entrada para o mascaramento de substituição. Configure outras colunas de entrada para serem dependentes dessa coluna de substituição. Por exemplo, escolha a coluna de Código Postal para mascaramento de substituição e as colunas de Cidade e Estado para serem dependentes da coluna de Código Postal. O mascaramento dependente garante que os valores substituídos de cidade e estado sejam válidos para o valor de código postal substituído.

**Nota:** Você só pode configurar uma coluna para mascaramento dependente depois de configurar uma coluna para mascaramento de substituição.

Configure as seguintes regras de mascaramento ao configurar uma coluna para mascaramento dependente:

### Coluna de Dependentes

O nome da coluna de entrada que você configurou para o mascaramento de substituição. A transformação de Mascaramento de Dados recupera dados substitutos de um dicionário usando as regras de mascaramento para aquela coluna. A coluna configurada para mascaramento de substituição se torna a coluna principal para recuperação de dados mascarados do dicionário.

### Coluna de Saída

O nome da coluna do dicionário que contém o valor da coluna que está sendo configurada com mascaramento dependente.

## Exemplo de mascaramento dependente

Um dicionário de mascaramento de dados pode conter linhas de endereço com os seguintes valores:

SNO	STREET	CITY	STATE	CEP	COUNTRY
1	32 Apple Lane	Chicago	IL	61523	US
2	776 Ash Street	Dallas	TX	75240	US

SNO	STREET	CITY	STATE	CEP	COUNTRY
3	2229 Big Square	Atleeville	TN	38057	US
4	6698 Cowboy Street	Houston	TX	77001	US

É necessário mascarar os dados de origem com combinações válidas de cidade, estado e código postal do dicionário Endereço.

Configure a porta Código Postal para mascaramento de substituição. Insira as seguintes regras de mascaramento para a porta Código Postal:

Regra	Valor
Nome do Dicionário	Endereço
Coluna de Número de Série	SNO
Coluna de Saída	CEP

Configure a porta Cidade para mascaramento dependente. Insira as seguintes regras de mascaramento para a porta Cidade:

Regra	Valor
Coluna de Dependentes	CEP
Coluna de Saída	Cidade

Configure a porta Estado para mascaramento dependente. Insira as seguintes regras de mascaramento para a porta Estado:

Regra	Valor
Coluna de Dependentes	CEP
Coluna de Saída	Estado

Quando a transformação de Mascaramento de Dados mascara o código postal, ela retorna da linha de dicionário a cidade e o estado corretos para o código postal em questão.

## Mascaramento de Tokenização

Use a técnica de mascaramento de tokenização para mascarar dados de string de origem com base em critérios que você especifica em um algoritmo. Por exemplo, você pode criar um algoritmo que contém um endereço de e-mail falso para substituir entradas de campos nos dados de origem.

Você pode configurar o formato dos dados mascarados usando o mascaramento de Tokenização. É necessário atribuir um nome de tokenizador ao algoritmo de mascaramento antes de usá-lo. O nome do tokenizador faz referência ao algoritmo de mascaramento (JAR) usado. Especifique o nome do tokenizador ao aplicar a técnica de mascaramento de tokenização.

## Configurando o Mascaramento de Tokenização

Realize as seguintes tarefas antes de usar a técnica de mascaramento de tokenização:

1. Navegue até o diretório `tokenprovider` no caminho: `<Informatica_home>\services\shared`.
2. Abra o seguinte arquivo XML: `com.informatica.products.ilm.tx-tokenizerprovider.xml`.
3. Adicione o nome do tokenizador e o nome totalmente qualificado do arquivo de classe para cada tokenizador que você deseja usar. Implemente a classe do tokenizador na classe `com.informatica.products.ilm.tx-tokenprovider-<Build-Number>.jar` do diretório `tokenprovider`. Para cada tokenizador, insira as informações no arquivo XML como no seguinte exemplo:

```
<TokenizerProvider>
<Tokenizer Name="CCTokenizer"
ClassName="com.informatica.tokenprovider.CCTokenizer"/>
</TokenizerProvider>
```

Onde:

- Nome do Tokenizador é o nome definido pelo usuário entre aspas.
- ClassName é o nome definido pelo usuário para o atributo CLASSNAME. Implemente-o a partir de `com.informatica.products.ilm.tx-tokenprovider-<Build-Number>.jar`.

Após a configuração, você poderá usar a técnica de mascaramento de Tokenização. Insira o nome do tokenizador para especificar o algoritmo a ser usado quando você criar um mapeamento.

## Regras de Mascaramento

Regras de mascaramento são as opções que você configura após escolher a técnica de mascaramento.

Quando você escolhe a técnica de mascaramento aleatório ou de chaves, pode configurar o formato de máscara, os caracteres da string de origem e os caracteres da string de resultado. Você pode configurar intervalo ou embaçamento com o mascaramento aleatório.

A seguinte tabela descreve as regras de mascaramento que você pode configurar para cada técnica de mascaramento:

Regra de Mascaramento	Descrição	Técnica de Mascaramento	Tipo de Dados de Origem
Formato de Máscara	Máscara que limita cada caractere em uma string de saída a um caractere alfabético, numérico ou alfanumérico.	Aleatório e Chave	String
Caracteres da String de Origem	Conjunto de caracteres de origem para mascarar ou excluir do mascaramento.	Aleatório e Chave	String
Caracteres de Substituição da String de Resultados	Um conjunto de caracteres para incluir ou excluir em uma máscara.	Aleatório e Chave	String

Regra de Mascaramento	Descrição	Técnica de Mascaramento	Tipo de Dados de Origem
Intervalo	Um intervalo de valores de saída. - Numérico. A transformação de Mascaramento de Dados retorna dados numéricos entre os valores mínimo e máximo. - String. Retorna uma string de caracteres aleatórios entre os comprimentos de string mínimo e máximo. - Data/Hora. Retorna uma data e uma hora dentro da data e hora mínimas e máximas.	Aleatório	Numérico String Date/Time
Embaçamento	Intervalo de valores de saída com uma variação fixa ou percentual dos dados de origem. A transformação de Mascaramento de Dados retorna dados que estão próximos ao valor dos dados de origem. As colunas de data e hora exigem uma variação fixa. As colunas exigem uma variação fixa.	Aleatório	Numérico Date/Time

## Formato de Máscara

Configure um formato de máscara para limitar cada caractere da coluna de saída a um caractere alfabético, numérico ou alfanumérico. Use os caracteres a seguir para definir um formato de máscara:

A, D, N, X, +, R

**Nota:** O formato de máscara contém caracteres em maiúsculas. Quando você insere um caractere de máscara em minúscula, a transformação de Mascaramento de Dados converte o caractere em letra maiúscula.

A tabela a seguir descreve os caracteres do formato de máscara:

Caractere	Descrição
A	Caracteres alfabéticos. Por exemplo, caracteres em ASCII de a a z e de A a Z.
Dia	Dígitos de 0 a 9. A transformação de Mascaramento de Dados retorna um "X" para caracteres diferentes dos dígitos de 0 a 9.
N	Caracteres alfanuméricos. Por exemplo, caracteres em ASCII de a a z, de A a Z e de 0 a 9.
X	Qualquer caractere. Por exemplo, alfanumérico ou símbolo.
+	Sem mascaramento.
R	Caracteres restantes. R especifica que os caracteres restantes na string podem ser qualquer tipo de caractere. R deve aparecer como o último caractere da máscara.

Por exemplo, um nome de departamento tem o formato a seguir:

nnn-<department\_name>

É possível configurar uma máscara para forçar os três primeiros caracteres para serem numéricos, o nome do departamento para ser alfanumérico e o traço para permanecer na saída. Configure o seguinte formato de máscara:

DDD+AAAAAAAAAAAAAAAA

A Transformação Mascaramento de Dados substitui os três primeiros caracteres por caracteres numéricos. O quarto caractere não é substituído. A transformação de Mascaramento de Dados substitui os caracteres restantes por caracteres alfabéticos.

Se você não definir um formato de máscara, a Transformação Mascaramento de Dados substituirá cada caractere de origem por qualquer caractere. Se o formato de máscara for mais extenso do que a string de entrada, a Transformação Mascaramento de Dados ignora os caracteres a mais no formato de máscara. Se o formato de máscara for menor que a string de origem, a transformação de Mascaramento de Dados mascarará os caracteres restantes com formato R.

**Nota:** Não é possível configurar um formato de máscara com a opção de intervalo.

## Caracteres da String de Origem

Caracteres da string de origem são caracteres que podem ser ou não mascarados. A posição dos caracteres na string de origem não importa. Os caracteres de origem diferenciam maiúsculas e minúsculas.

Você pode configurar qualquer número de caracteres. Se Caracteres estiver em branco, a transformação de Mascaramento de Dados substituirá todos os caracteres de origem na coluna.

Selecione uma das seguintes opções para caracteres da string de origem:

### Apenas máscara

A transformação de Mascaramento de Dados mascara caracteres que você tenha configurado como caracteres da string de origem. Por exemplo, se você inserir os caracteres A, B e c, a transformação de Mascaramento de Dados substituirá A, B ou c por outro caractere quando ocorrer nos dados de origem. Outros caracteres de origem diferentes de A, B ou c não serão alterados. A máscara diferencia maiúsculas e minúsculas.

### Mascarar tudo exceto

Mascara todos os caracteres, exceto os caracteres da sequência de origem que ocorram na string de origem. Por exemplo, se você inserir o caractere de origem do filtro "-" e selecionar Mascarar Todos Exceto, a transformação de Mascaramento de Dados não substituirá o caractere "-" quando ele ocorrer na sequência de origem. Os outros caracteres de origem serão alterados.

## Exemplo de String de Origem

Um arquivo de origem apresenta uma coluna denominada Dependentes. A coluna Dependentes contém mais de um nome separados por vírgulas. É necessário mascarar a coluna Dependentes e manter a vírgula nos dados de teste para delimitar os nomes.

No caso da coluna de Dependências, selecione Caracteres da String de Origem. Escolha Não Mascarar e digite "," para o caractere de origem a ignorar. Não digite aspas.

A Transformação Mascaramento de Dados substitui todos os caracteres na string de origem, exceto a vírgula.

## Caracteres de Substituição da String de Resultados

Caracteres de substituição da string de resultados são escolhidos por você como caracteres substitutos nos dados mascarados. Quando você configura caracteres de substituição da string de resultados, a transformação de Mascaramento de Dados substitui caracteres na string de origem por caracteres de substituição da string de resultados. Para não gerar a mesma saída para valores de entrada diferentes, configure uma variedade grande de caracteres substitutos ou mascare poucos caracteres da origem. A posição de cada caractere na string não é importante.

Selecione uma das seguintes opções para caracteres de substituição da string de resultados:

### Usar Apenas

Mascare a origem só com os caracteres que você definir como caracteres de substituição da string de resultados. Por exemplo, se você inserir os caracteres A, B e c, a transformação de Mascaramento de Dados substituirá cada caractere na coluna de origem por A, B ou c. A palavra "casa" poderá ser substituída por "BAcA".

### Usar tudo exceto

Mascare a origem com quaisquer caracteres, exceto aqueles que você definir como caracteres de substituição da string de resultados. Por exemplo, se você inserir os caracteres de substituição da string de resultados A, B e c, os dados mascarados nunca conterão os caracteres A, B e c.

## Exemplo de caracteres de substituição da string de resultados

Para substituir todas as vírgulas da coluna Dependentes por pontos e vírgulas, realize as seguintes tarefas:

1. Configure as vírgulas como caracteres de string de origem e selecione Apenas Máscara.  
A Transformação Mascaramento de Dados mascara apenas as vírgulas quando essas ocorrem na coluna Dependentes.
2. Configure os pontos-e-vírgulas como caracteres de substituição de string de resultado e selecione Apenas Usar.  
A Transformação Mascaramento de Dados substitui cada vírgula da coluna Dependentes por um ponto-e-vírgula.

## Intervalo

Defina um intervalo para dados numéricos, de data ou de string. Quando você define um intervalo para valores numéricos ou de data, a transformação de Mascaramento de Dados mascara os dados de origem com um valor entre os valores mínimo e máximo. Quando configura um intervalo para uma string, você configura um intervalo de tamanhos de string.

### Intervalo de Strings

Quando você configura um mascaramento de string aleatório, a transformação de Mascaramento de Dados gera strings com comprimento variável em relação ao comprimento da string de origem. Se preferir, você pode configurar uma largura de string máxima e uma mínima. Os valores que você inserir como largura máxima ou mínima deverão ser números inteiros positivos. Cada largura deverá ser menor ou igual à precisão da porta.

### Intervalo numérico

Defina os valores mínimo e máximo para uma coluna numérica. O valor máximo deve ser menor que ou igual à precisão da porta. O intervalo padrão vai desde um até o comprimento da precisão da porta.

## Intervalo de datas

Configure os valores mínimo e máximo para o valor datetime. Os campos mínimos e máximos contêm as datas mínima e máxima padrão. O formato padrão de datetime é MM/DD/AAAA HH24:MI:SS. O datetime máximo deve ser posterior ao mínimo.

## Embaçamento

O Embaçamento cria um valor de saída dentro de uma variância fixa ou percentual a partir do valor dos dados de origem. Configure o embaçamento para retornar um valor aleatório próximo do valor original. Você pode embaçar valores numéricos e de data.

### Embaçando Valores Numéricos

Selecione uma variância fixa ou percentual para embaçar um valor de origem numérico. O valor baixo de embaçamento é uma variância que está abaixo do valor de origem. O valor alto de embaçamento é uma variância que está acima do valor de origem. Os valores baixo e alto devem ser maiores ou iguais a zero. Quando a Transformação Mascaramento de Dados retorna dados mascarados, os dados numéricos ficam no intervalo que você define.

A tabela a seguir descreve os resultados de mascaramento dos valores de intervalo de embaçamento quando o valor de origem de entrada é 66:

Tipo de Embaçamento	Inferior	Alto	Resultado
Determinado	0	10	Entre {66} e {76}
Determinado	10	0	Entre {56} e {66}
Determinado	10	10	Entre 56 e 76
Por cento	0	50	Entre 66 e 99
Por cento	50	0	Entre 33 e 66
Por cento	50	50	Entre 33 e 99

### Embaçando Valores de Data

Masque uma data como variância da data de origem, configurando o embaçamento. Selecione uma unidade da data na qual aplicar a variância. É possível selecionar o ano, o mês, o dia ou a hora. Insira os limites mínimo e máximo para definir uma variância que esteja acima e abaixo da unidade da data de origem. A Transformação de Mascaramento de Dados aplica a variância e retorna uma data que esteja dentro da variância.

Por exemplo, para restringir a data mascarada para uma data que esteja a dois anos da data de origem, selecione o ano como unidade. Insira 2 como o limite mínimo e máximo. Se uma data de origem for 02/02/2006, a Transformação de Mascaramento de Dados retornará uma data entre 02/02/2004 e 02/02/2008.

Por padrão, a unidade de embaçamento é o ano.

# Formatos Especiais de Máscara

Formatos especiais de máscara são máscaras que você pode aplicar a tipos de dados comuns. Com um formato de máscara especial, a transformação de Mascaramento de Dados retorna um valor mascarado que tem um formato realista, mas não é um valor válido.

Por exemplo, quando você mascara um número de Seguridade Social, a transformação de Mascaramento de Dados retorna um número da previdência social com formato correto, mas que não é válido. Você pode configurar mascaramentos repetíveis para números de Seguridade Social.

Configure máscaras especiais para os seguintes tipos de dados:

- Números de Seguridade Social
- Números de cartão de crédito
- Números de telefone
- Endereços de URL
- Endereços de e-mail
- Endereços IP
- Números de Seguro Social

Quando o tipo ou o formato dos dados de origem é inválido para uma máscara, o Data Integration Service aplica uma máscara padrão aos dados. O Serviço de Integração aplica valores mascarados do arquivo de valores padrão. Você pode editar o arquivo de valores padrão para alterar esses valores.

## Mascaramento de Número de Cartão de Crédito

A transformação Mascaramento de Dados gera um número de cartão de crédito logicamente válido, quando mascara um número de cartão de crédito válido. O comprimento do número do cartão de crédito de origem deve ser 13 a 19 dígitos. O número do cartão de crédito de entrada deve ter uma soma de verificação válida e baseada nas regras do setor de cartão de crédito.

O número do cartão de crédito de origem pode conter números, espaços e hífen. Se o cartão de crédito tiver caracteres incorretos ou o comprimento errado, o Serviço de Integração escreve um erro no log da sessão. O Serviço de Integração aplica uma máscara padrão de número do cartão de crédito quando os dados de origem são inválidos.

A transformação Mascaramento de Dados não mascara o Número de Identificação Bancária (BIN) de seis dígitos. Por exemplo, a transformação Mascaramento de Dados pode mascarar o número de cartão de crédito 4539 1596 8210 2773 como 4539 1516 0556 7067. A transformação de Mascaramento de Dados cria um número mascarado que tem uma soma de verificação válida.

## Mascaramento de Endereço de E-mail

Use a transformação de Mascaramento de Dados para mascarar o endereço de e-mail que contém o valor da string. A transformação de Mascaramento de Dados pode mascarar um endereço de e-mail com caracteres ASCII aleatórios ou substituir o endereço de e-mail com um endereço realista de e-mail.

Você pode aplicar os seguintes tipos de mascaramento com o endereço de e-mail:

### **Mascaramento de e-mail padrão**

A transformação de Mascaramento de Dados retorna caracteres ASCII aleatórios quando mascara um endereço de e-mail. Por exemplo, a transformação de Mascaramento de Dados pode mascarar `Georgesmith@yahoo.com` como `KtrlupQAPyk@vdSKh.BIC`. É o padrão.



### **Mascaramento de e-mail avançado**

A transformação de Mascaramento de Dados mascara o endereço de e-mail com outro endereço realista de e-mail derivado das portas de saída da transformação ou das colunas de dicionário.

## **Mascaramento de E-mail Avançado**

Com o tipo de mascaramento de e-mail avançado, você pode mascarar o endereço de e-mail com outro endereço de e-mail realista. A transformação de Mascaramento de Dados cria o endereço de e-mail das colunas de dicionário ou das portas de saída da transformação.

Você pode criar a parte local no endereço de e-mail a partir do mapeamento de portas de saída. Ou pode criar a parte local no endereço de e-mail a partir da tabela relacional ou das colunas de arquivos simples.

A transformação de Mascaramento de Dados pode criar o nome de domínio para o endereço de e-mail a partir de um valor constante ou de um valor aleatório no dicionário de domínio.

É possível criar um mascaramento de e-mail avançado com base nas seguintes opções:

### **Endereço de E-mail com Base em Portas Dependentes**

Você pode criar um endereço de e-mail com base nas portas de saída da transformação de Mascaramento de Dados. Selecione as portas de saída da transformação para as colunas de nome e sobrenome. A transformação de Mascaramento de Dados mascara o nome, o sobrenome ou ambos, com base nos valores especificados para o comprimento do nome e do sobrenome.

### **Endereço de E-mail com Base em um Dicionário**

Você pode criar um endereço de e-mail com base nas colunas de um dicionário. Selecione uma tabela de referência como origem para o dicionário.

Selecione as colunas do dicionário para o nome e o sobrenome. A transformação de Mascaramento de Dados mascara o nome, o sobrenome ou ambos, com base nos valores especificados para o comprimento do nome e do sobrenome.

## **Parâmetros de Configuração para um Tipo de Mascaramento de Endereço de E-mail Avançado**

Especifique os parâmetros de configuração quando você configurar o mascaramento de endereço de e-mail avançado.

Você pode especificar os seguintes parâmetros de configuração:

### **Delimitador**

Você pode selecionar um delimitador, como um ponto, hífen ou sublinhado, para separar o nome e o sobrenome no endereço de e-mail. Se você não desejar separar o nome e o sobrenome no endereço de e-mail, deixe o delimitador em branco.

### **Coluna de FirstName**

Selecione uma porta de saída da transformação de Mascaramento de Dados ou uma coluna de dicionário para mascarar o nome no endereço de e-mail.

### **Coluna de LastName**

Selecione uma porta de saída da transformação de Mascaramento de Dados ou uma coluna de dicionário para mascarar o sobrenome no endereço de e-mail.

### **Tamanho das colunas de FirstName ou LastName**

Restringe o comprimento de caracteres para mascarar as colunas de nome e de sobrenome. Por exemplo, os dados de entrada são Timothy para nome e Smith para sobrenome. Selecione 5 como o

tamanho da coluna de nome. Selecione 1 como o tamanho da coluna de sobrenome e um ponto como o delimitador. A transformação de Mascaramento de Dados gera o seguinte endereço de e-mail:

```
timot.s@<domain_name>
```

#### DomainName

Você pode usar um valor constante, como gmail.com, para o nome de domínio. Ou você pode especificar um outro arquivo de dicionário que contém uma lista de nomes de domínio. O dicionário de domínio pode ser um arquivo simples ou uma tabela relacional.

## Mascaramento de Endereço IP

A transformação de Mascaramento de Dados mascara um endereço IP, transformando-o em outro endereço IP, dividindo-o em quatro números separados por um ponto final. O primeiro número é a rede. A Transformação Mascaramento de Dados mascara o número de rede dentro dos limites da rede.

A Transformação Mascaramento de Dados mascara um endereço IP de Classe A como um endereço IP de Classe A e um endereço 10.x.x.x como um endereço 10.x.x.x. A Transformação Mascaramento de Dados não mascara o endereço de rede de classe e privado. Por exemplo, a Transformação Mascaramento de Dados pode mascarar 11.12.23.34 como 75.32.42.52. e 10.23.24.32 como 10.61.74.84.

**Nota:** Quando você mascara vários endereços IP, a transformação de Mascaramento de Dados pode retornar valores não únicos, pois não mascara a classe ou a rede privada dos endereços IP.

## Mascaramento de Número de Telefone

A transformação de Mascaramento de Dados mascara um número de telefone sem alterar o formato do número de telefone original. Por exemplo, a transformação de Mascaramento de Dados pode mascarar o número de telefone (607)382 0658 como (408)256 3106.

Os dados de origem podem conter números, espaços, hífen e parênteses. O Serviço de Integração não mascara caracteres alfabéticos nem caracteres especiais.

A transformação de Mascaramento de Dados pode mascarar dados do tipo string, integer e bigint.

## Mascaramento de CPF

A transformação de Mascaramento de Dados gera um número de Seguridade Social que não é válido com base na High Group List mais recente da Administração de Seguridade Social. A High Group List contém números válidos emitidos pela Administração de Seguridade Social.

A High Group List padrão é um arquivo de texto no seguinte local:

```
<Installation Directory>\infa_shared\SrcFiles\highgroup.txt
```

Para usar o arquivo da High Group List em fluxos de trabalho, copie o arquivo de texto para o diretório de origem que você configura para o Data Integration Service.

A transformação de Mascaramento de Dados gera números de Seguridade Social que não estão na High Group List. A Administração da Seguridade Social atualiza a High Group List todo mês. Baixe a versão mais recente da lista no seguinte local:

```
http://www.socialsecurity.gov/employer/ssns/highgroup.txt
```

## Formato de Número da Previdência Social

A transformação de Mascaramento de Dados aceita qualquer formato SSN que contenha nove dígitos. Os dígitos podem ser delimitados por qualquer conjunto de caracteres. Por exemplo, a transformação de Mascaramento de Dados aceita o formato a seguir: `+54-*9944$#789-,*()''`.

## Requisito de Código de Área

A transformação de Mascaramento de Dados retorna um número de previdência social que não é válido com o mesmo formato da origem. Os primeiros três dígitos do SSN definem o código de área. A Transformação Mascaramento de Dados não mascara o código de área. Ela mascara o número do grupo e o número de série. O SSN de origem deve conter um código de área válido. A transformação mascaramento de dados localiza o código de área na Lista de Grupos Altos e determina um intervalo de novos números que ela pode aplicar como dados mascarados. Se o SSN não for válido, a transformação mascaramento de dados não mascara os dados de origem.

## Mascaramento de CPF Repetível

A transformação de Mascaramento de Dados retorna números de CPF deterministas com mascaramento repetível. A transformação de Mascaramento de Dados não pode retornar todos os números de CPF exclusivos porque não pode retornar números de CPF válidos emitidos pela Administração do Seguro Social.

## Mascaramento de Endereço de URL

A transformação de Mascaramento de Dados procura a string `'://'` e analisa a subsequência à direita para analisar uma URL. A URL de origem deve conter a string `'://'`. A URL de origem pode conter caracteres alfabéticos e numéricos.

A transformação de Mascaramento de Dados não mascara o protocolo da URL. Por exemplo, se a URL for `http://www.yahoo.com`, a transformação de Mascaramento de Dados poderá retornar `http://MgL.aHjCa.VsD/`. A transformação de Mascaramento de Dados pode gerar uma URL que não é válida.

**Nota:** A transformação de Mascaramento de Dados sempre retorna caracteres ASCII para uma URL.

## Mascaramento de Número de Previdência Social

A transformação de Mascaramento de Dados mascara um CPF de nove dígitos. Os dígitos podem ser delimitados por qualquer conjunto de caracteres.

Se o número não contiver delimitadores, o número mascarado também não conterá delimitadores. Caso contrário, o número mascarado terá o seguinte formato:

`xxx-xxx-xxx`

## Números SIN Repetíveis

Você pode configurar a transformação de Mascaramento de Dados para retornar valores SIN repetíveis. Quando você configura uma porta para mascaramento SIN repetível, a transformação de Mascaramento de Dados retorna dados mascarados determinísticos sempre que o valor SIN de origem e o valor de semente são iguais.

Para retornar números SIN repetíveis, ative **Valores Repetíveis** e insira um número de semente. A transformação de Mascaramento de Dados retorna valores exclusivos para cada SIN.

## Dígito de Início SIN

Você pode definir o primeiro dígito do SIN mascarado.

Ative **Dígito de Início** e insira o dígito. A transformação de Mascaramento de Dados cria números SIN mascarados que começam com o número que você inseriu.

## Arquivo de Valor Padrão

Quando o tipo ou o formato dos dados de origem é inválido para uma máscara, o Data Integration Service aplica uma máscara padrão aos dados. O Serviço de Integração aplica valores mascarados do arquivo de valores padrão. Você pode editar o arquivo de valores padrão para alterar esses valores.

O arquivo de valores padrão é um arquivo XML no seguinte local:

```
<Diretório de Instalação>\infa_shared\SrcFiles\defaultValue.xml
```

Para usar o arquivo de valores padrão em fluxos de trabalho, copie o arquivo de valores padrão no diretório de origem que você configura para o Data Integration Service.

O arquivo defaultValue.xml contém os seguintes pares de nome-valor:

```
<?xml version="1.0" standalone="yes" ?>
<defaultValue
  default_char = "X"
  default_digit = "9"
  default_date = "11/11/1111 00:00:00"
  default_email = "abc@xyz.com"
  default_ip = "99.99.9.999"
  default_url = "http://www.xyz.com"
  default_phone = "999 999 999 9999"
  default_ssn = "999-99-9999"
  default_cc = "9999 9999 9999 9999"
  default_sin = "999-999-999"
  default_seed = "500"/>
```

### TÓPICOS RELACIONADOS:

- [“Configurar o Data Integration Service” na página 225](#)

## Configuração da Transformação de Mascaramento de Dados

Use as etapas a seguir para configurar uma transformação de Mascaramento de Dados.

1. Configure opções de execução para o Data Integration Service.
2. Crie a transformação.
3. Defina as portas de entrada.
4. Configure regras de mascaramento para cada porta que você deseja alterar.
5. Visualize os dados para verificar os resultados.

## Configurar o Data Integration Service

Você pode configurar opções de execução para o Data Integration Service no Informatica Administrator (a ferramenta Administrator).

Configure opções de execução para definir os seguintes diretórios padrão:

- Diretório inicial. Contém o diretório de origem e o diretório de cache.
- Diretório de origem. Contém os arquivos de origem para fluxos de trabalho. Por exemplo, o diretório de origem pode conter os arquivos highgrp.txt e defaultvalue.xml.
- Diretório de cache. Contém arquivos de cache para mascaramento de substituição.

Para definir valores para as opções de execução, abra a ferramenta Administrator e selecione o Data Integration Service no **Navegador do Domínio**. Clique na exibição **Propriedades** e, em seguida, clique em **Edit** na seção **Opções de Execução**.

### TÓPICOS RELACIONADOS:

- [“Arquivo de Valor Padrão” na página 224](#)

## Criando uma Transformação de Mascaramento de Dados

Crie uma transformação de Mascaramento de Dados na Developer tool.

Antes de criar a transformação de Mascaramento de Dados, crie a origem. Importe um arquivo simples ou uma tabela de banco de dados relacional como um objeto de dados físicos.

1. Selecione um projeto ou pasta na exibição **Object Explorer**.
2. Clique em **Arquivo > Novo > Transformação**.  
A caixa de diálogo **Novo** é exibida.
3. Selecione a transformação de Mascaramento de Dados.
4. Clique em **Avançar**.
5. Insira um nome para a transformação.
6. Clique em **Concluir**.  
A transformação aparece no editor.

## Definindo as Portas

Adicione portas de entrada de Mascaramento de Dados na exibição **Visão Geral**. Quando você cria uma porta de entrada, a ferramenta Developer cria uma porta de saída correspondente por padrão. A porta de saída tem o mesmo nome que a porta de entrada.

1. Na exibição **Visão Geral**, clique em **Novo** para adicionar uma porta.
2. Configure o tipo de dados, precisão e escala para a coluna.  
Você deve configurar o tipo de dados da coluna antes de definir regras de mascaramento para a coluna.
3. Para configurar o mascaramento de dados para a porta, clique na seta na coluna de tipo de mascaramento da exibição **Visão Geral**.

## Configurando o Mascaramento de Dados para Cada Porta

Selecione uma técnica de mascaramento e as regras de mascaramento correspondentes para uma porta na caixa de diálogo Mascaramento de Dados. A caixa de diálogo Mascaramento de Dados é exibida quando você clica na coluna de mascaramento de dados da guia **Portas**.

1. Ative **Aplicar Mascaramento** para configurar o mascaramento para a porta selecionada.  
A ferramenta Developer exibe uma lista de técnicas de mascaramento que você pode usar com base no tipo de dados da porta a ser mascarada.
2. Selecione uma técnica de mascaramento da lista.  
A ferramenta Developer exibe diferentes regras de mascaramento com base na técnica de mascaramento escolhida. Alguns dos formatos especiais de máscara não têm regras de mascaramento a serem configuradas.
3. Configure as regras de mascaramento.
4. Clique em **OK** para aplicar a configuração de mascaramento de dados para a porta.  
Quando você define o mascaramento de dados para uma porta, a ferramenta Developer cria uma porta de saída chamada **out-<nome da porta>**. O **<nome da porta>** é o mesmo nome que a porta de entrada. A transformação de Mascaramento de Dados retorna os dados mascarados na porta **out-<nome da porta>**.

## Visualizando os Dados Mascarados

Você pode comparar os dados mascarados aos dados originais quando exibe os resultados da transformação de Mascaramento de Dados no **Visualizador de Dados**.

1. Após configurar as portas de transformação de Mascaramento de Dados e as regras de mascaramento, crie um mapeamento que inclua a origem do objeto de dados físicos e a transformação de Mascaramento de Dados.
2. Conecte a origem à transformação de Mascaramento de Dados.
3. Verifique se a origem tem dados em um local compartilhado que o Serviço de Integração de Dados possa acessar.
4. Clique na transformação de Mascaramento de Dados para selecioná-la no mapeamento.
5. Clique em **Visualizador de Dados** e em **Executar**.  
A ferramenta Developer exibe os dados de todas as portas de saída da transformação de Mascaramento de Dados. As portas que têm o prefixo de **saída** (output) contêm dados mascarados. Você pode comparar os dados mascarados aos dados originais na exibição **Visualizador de Dados**.

## Propriedades de Tempo de Execução da Transformação de Mascaramento de Dados

Você pode configurar propriedades de tempo de execução da transformação de Mascaramento de Dados para melhorar o desempenho.

Configure as seguintes propriedades de tempo de execução:

### **Tamanho do Cache**

O tamanho do cache do dicionário na memória principal. Aumente o tamanho de memória para melhorar o desempenho. O tamanho mínimo recomendado é 32 MB para 100.000 registros. O padrão é 8 MB.

### **Diretório de Cache**

A localização do cache do dicionário. Você deve ter permissões de gravação para o diretório. O padrão é CacheDir.

### **Tabela de armazenamento compartilhado**

Habilita o compartilhamento de tabela de armazenamento entre instâncias de transformação de Mascaramento de Dados. Ative a tabela de armazenamento compartilhado quando duas instâncias da transformação de Mascaramento de Dados usarem a mesma coluna do dicionário para a conexão de banco de dados, valor de semente e localidade. Você também pode habilitar a tabela de armazenamento compartilhado quando duas portas na mesma transformação de Mascaramento de Dados usarem a mesma coluna do dicionário para a conexão, semente, e a localidade. Desative a tabela de armazenamento compartilhado quando as Transformações de Mascaramento de Dados ou as portas não compartilharem a coluna do dicionário. Por padrão, essa opção fica desativada.

### **Intervalo de Confirmação de Armazenamento**

O número de linhas a serem confirmadas de uma vez para a tabela de armazenamento. Aumente o valor para melhorar o desempenho. Configure o intervalo de confirmação quando não configurar a tabela de armazenamento compartilhado. O padrão é 100.000.

### **Criptografar Armazenamento**

Criptografa as tabelas de armazenamento, tais como IDM\_SUBSTITUTION\_STORAGE e IDM\_EXPRESSION\_STORAGE. Verifique se você criptografou dados em tabelas de armazenamento antes de ativar a propriedade de criptografia de armazenamento. Desmarque essa opção se você não deseja criptografar as tabelas de armazenamento. Por padrão, essa opção fica desativada.

### **Chave de Criptografia de Armazenamento**

A transformação de Mascaramento de Dados criptografa o armazenamento com base na chave de criptografia de armazenamento. Use a mesma chave de criptografia para cada execução de sessão da mesma instância da transformação de Mascaramento de Dados.

### **Nome do Proprietário de Dicionário de Substituição**

Nome do proprietário da tabela de dicionário de substituição quando você seleciona o tipo de mascaramento de substituição. Se o usuário do banco de dados especificado na conexão de banco de dados não for o proprietário da tabela de dicionário de substituição em uma sessão, você precisará especificar o proprietário da tabela.

### **Nome do Proprietário de Armazenamento**

Nome do proprietário da tabela para IDM\_SUBSTITUTION\_STORAGE ou IDM\_EXPRESSION\_STORAGE quando você seleciona o tipo de mascaramento de expressão repetível ou de substituição repetível exclusiva.

## **Exemplo de Mascaramento de Dados**

Um desenvolvedor precisa criar dados de teste para aplicativos de clientes. Esses dados devem conter dados realistas de clientes que possam ser acessados por outros desenvolvedores no ambiente de desenvolvimento da empresa.

O desenvolvedor cria um serviço de dados que retorna dados mascarados dos clientes, como ID de cliente, número de cartão de crédito e renda. O mapeamento inclui uma transformação de Mascaramento de Dados que transforma os dados dos clientes.

A seguinte figura mostra o mapeamento:



O mapeamento contém as seguintes transformações:

- Read\_Customer\_Data. Contém as informações de cartão de crédito e renda do cliente.
- Transformação Customer\_Data\_Masking. Mascara todas as colunas, exceto FirstName e LastName. A transformação de Mascaramento de Dados passa as colunas mascaradas para o destino.
- Customer\_TestData. Transformação de saída que recebe os dados mascarados dos clientes.

## Dados Read\_Customer

Os dados dos clientes contêm as seguintes colunas:

Coluna	Tipo de dados
CustomerID	Número inteiro
LastName	String
FirstName	String
CreditCard	String
Income	Número inteiro
Join_Date	Datetime (DD/MM/YYYY)

A seguinte tabela mostra exemplos de dados de clientes:

CustomerID	LastName	FirstName	CreditCard	Income	JoinDate
0095	Bergeron	Barbara	4539-1686-3069-3957	12000	12/31/1999
0102	Brosseau	Derrick	5545-4091-5232-8948	4000	03/03/2011
0105	Anderson	Lauren	1234-5678-9012-3456	5000	04/03/2009
0106	Boonstra	Pauline	4217-9981-5613-6588	2000	07/07/2007
0107	Chan	Brian	4533-3156-8865-3156	4500	06/18/1995

## Transformação de Mascaramento de Dados de Clientes

A transformação de Mascaramento de Dados mascara todas as colunas na linha do cliente, exceto o nome e o sobrenome.

A transformação de Mascaramento de Dados executa os seguintes tipos de mascaramento:

- Mascaramento de chaves



- Mascaramento aleatório
- Mascaramento de cartão de crédito

A seguinte tabela mostra as regras de mascaramento para cada porta na transformação de Mascaramento de Dados:

Porta de Entrada	Tipo de Mascaramento	Regras de Mascaramento	Descrição
CustomerID	Chave	A semente é 934. A ID de cliente não tem nenhum formato de máscara. Os caracteres de substituição da string de resultados são 1234567890.	A máscara CustomerID é determinística. A ID de cliente mascarada contém números.
LastName	Sem Mascaramento	-	-
FirstName	Sem Mascaramento	-	-
CreditCard	CreditCard	-	A transformação de Mascaramento de Dados mascara o número de cartão de crédito com outro número que tenha uma soma de verificação válida.
Income	Aleatório	Embaçamento Porcentagem Limite inferior = 1 Limite superior = 10	A renda mascarada está dentro de dez por cento da renda de origem.
JoinDate	Aleatório	Embaçamento Unidade = Ano Limite Inferior = 5 HighBound=5	A data mascarada está dentro de 5 anos da data original.

## Resultados de Dados de Teste de Clientes

A transformação Customer\_TestData recebe dados realistas de clientes da transformação de Mascaramento de Dados.

O destino Customer\_TestData recebe os seguintes dados:

out-CustomerID	out-LastName	outFirstName	out-CreditCard	out-Income	out-JoinDate
3954	Bergeron	Barbara	4539-1625-5074-4106	11500	03/22/2001
3962	Brosseau	Derrick	5545-4042-8767-5974	4300	04/17/2007
3964	Anderson	Lauren	1234-5687-2487-9053	5433	09/13/2006
3965	Boonstra	Pauline	4217-9935-7437-4879	1820	02/03/2010
3966	Chan	Brian	4533-3143-4061-8001	4811	10/30/2000

A renda está dentro de dez por cento da renda original. A data de associação está dentro de cinco anos da data original.

# Propriedades Avançadas da Transformação de Mascaramento de Dados

Configure propriedades que ajudam a determinar como o Serviço de Integração de Dados processa dados para a transformação de Mascaramento de Dados.

Você pode configurar níveis de rastreamento para logs.

Configure a seguinte propriedade na guia **Avançado**:

## Nível de Rastreamento

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.

## Transformação de Mascaramento de dados em um ambiente não nativo

O processamento da transformação de Mascaramento de dados em um ambiente não nativo depende do mecanismo que executa a transformação.

Considere o suporte para os seguintes mecanismos de tempo de execução não nativos:

- Mecanismo Blaze. Suportado com restrições.
- Mecanismo Spark. Com suporte com restrições em mapeamentos de fluxo e em lote.
- Mecanismo Databricks Spark. Sem suporte.

## Transformação de Mascaramento de dados no mecanismo Blaze

A transformação de Mascaramento de dados é suportada com as seguintes restrições.

A validação do mapeamento falha nas seguintes situações:

- A transformação é configurada para mascaramento de expressão repetível.
- A transformação é configurada para mascaramento exclusivo de substituição repetível.

Você pode usar as seguintes técnicas de mascaramento neste mecanismo:

Cartão de Crédito  
E-mail  
Expressão  
Endereço IP  
Chave  
Telefone  
Aleatório  
SIN  
SSN  
Tokenização

URL  
Substituição aleatória  
Substituição repetível  
Dependente com Substituição aleatória  
Dependente com Substituição repetível

## Transformação de Mascaramento de dados no mecanismo Spark

A transformação de Mascaramento de dados é suportada com as seguintes restrições.

A validação do mapeamento falha nas seguintes situações:

- A transformação é configurada para mascaramento de expressão repetível.
- A transformação é configurada para mascaramento exclusivo de substituição repetível.

Você pode usar as seguintes técnicas de mascaramento neste mecanismo:

Cartão de Crédito  
E-mail  
Expressão  
Endereço IP  
Chave  
Telefone  
Aleatório  
SIN  
SSN  
Tokenização  
URL  
Substituição aleatória  
Substituição repetível  
Dependente com Substituição aleatória  
Dependente com Substituição repetível

Para otimizar o desempenho da transformação de Mascaramento de dados, configure as seguintes propriedades de configuração do mecanismo Spark na conexão do Hadoop:

### **spark.executor.cores**

Indica o número de núcleos que cada processo do executor usa para executar tasklets no mecanismo Spark.

Defina como: `spark.executor.cores=1`

### **spark.executor.instances**

Indica o número de instâncias que cada processo do executor usa para executar tasklets no mecanismo Spark.

Defina como: `spark.executor.instances=1`

**spark.executor.memory**

Indica a quantidade de memória que cada processo do executor usa para executar tasklets no mecanismo Spark.

Defina como: `spark.executor.memory=3G`

## Transformação de Mascaramento de dados em um mapeamento de fluxo

Os mapeamentos de fluxo têm as mesmas regras de processamento que os mapeamentos em lote no mecanismo Spark.

## CAPÍTULO 13

# Transformação do Processador de Dados

Este capítulo inclui os seguintes tópicos:

- [Visão geral da Transformação do Processador de Dados, 233](#)
- [Exibições da Transformação do Processador de Dados, 234](#)
- [Portas da Transformação do Processador de Dados, 235](#)
- [Componente de Inicialização, 237](#)
- [Referências, 238](#)
- [Configurações da Transformação do Processador de Dados, 239](#)
- [Eventos, 246](#)
- [Logs, 247](#)
- [Desenvolvimento da Transformação do Processador de Dados, 250](#)
- [Importação e Exportação da Transformação do Processador de Dados, 254](#)
- [Transformação de Processador de Dados em um ambiente não nativo, 257](#)

## Visão geral da Transformação do Processador de Dados

A transformação do Processador de Dados processa formatos de arquivo não estruturados e semi-estruturados em um mapeamento. Configure a transformação para processar formatos de mensagens, páginas HTML, XML, JSON e documentos PDF. Você também pode converter formatos estruturados como ACORD, HIPAA, HL7, EDI-X12, EDIFACT e SWIFT.

Um mapeamento usa uma transformação do Processador de Dados para alterar os documentos de um formato para outro. A transformação do Processador de Dados processa arquivos de qualquer formato em um mapeamento. Quando cria uma transformação do Processador de Dados, você define os componentes que convertem os dados.

Uma transformação do Processador de Dados pode conter vários componentes para processar dados. Cada componente pode conter outros componentes.

Por exemplo, você poderá receber faturas do cliente em arquivos do Microsoft Word. Configure uma transformação do Processador de Dados para analisar os dados de cada arquivo do Word. Extraia os dados do cliente para uma tabela de Cliente. Extraia informações sobre o pedido para uma tabela de Pedidos.

Quando cria uma transformação do Processador de Dados, você define um XMap, um Script ou uma Biblioteca. Um XMap converte um arquivo hierárquico de entrada em um arquivo hierárquico de saída de outra estrutura. Uma Biblioteca transforma um tipo de mensagem de indústria em um documento XML com uma estrutura hierárquica, ou de XML em um formato padrão da indústria. Um Script pode analisar documentos de origem em formato hierárquico, converter do formato hierárquico em outros formatos de arquivo ou mapear um documento hierárquico para outro formato hierárquico.

Defina Scripts na transformação do Processador de Dados no Editor do IntelliScript. Você pode definir os seguintes tipos de Scripts:

- **Analisador.** Converte documentos de origem em XML. A saída de um Analisador é sempre XML. A entrada pode ter qualquer formato, como texto, HTML, Word, PDF ou HL7.
- **Serializador.** Converte um arquivo XML em um documento de saída de qualquer formato. A saída de um Serializador pode ter qualquer formato, como documento de texto, documento HTML ou PDF.
- **Mapeador.** Converte um documento de origem XML em outra estrutura ou esquema XML. Você pode transformar os mesmos documentos XML como em um XMap.
- **Transformador.** Modifica os dados em qualquer formato. Adiciona, remove, converte ou altera o texto. Use Transformadores com um Analisador, um Mapeador ou um Serializador. Também você pode executar um Transformador como um componente autônomo.
- **Streamer.** Divide grandes documentos de entrada, como fluxos de dados multigigabytes, em segmentos. O Streamer processa documentos que contêm várias mensagens ou registros, como arquivos HIPAA ou EDI.

Para obter mais informações, consulte o *Guia do Usuário do Data Transformation*.

## Exibições da Transformação do Processador de Dados

A transformação do Processador de Dados tem várias exibições que você pode acessar ao configurar essa transformação e executá-la na ferramenta Developer.

Algumas das exibições da transformação do Processador de Dados não aparecem na ferramenta Developer por padrão. Para alterar as exibições da transformação, clique em **Janela > Mostrar Exibição > Outros > Informatica**. Selecione as exibições que você deseja ver.

A transformação do Processador de Dados tem as seguintes exibições fixas:

### Exibição Visão Geral

Configure portas e defina o componente de inicialização.

### Exibição de Referência

Adicione ou remova esquemas da transformação.

### Exibição de Configurações

Defina configurações de transformação para codificação, controle de saída e geração de XML.

### Exibição Objetos

Adicione, modifique ou exclua objetos de Script, XMap e Biblioteca da transformação.

Também é possível acessar as seguintes exibições para a transformação do Processador de Dados:

#### **Exibição Origem Hexadecimal do Processador de Dados**

Mostra um documento de entrada em formato hexadecimal.

#### **Exibição de Eventos do Processador de Dados**

Mostra informações sobre eventos que ocorrem quando você executa a transformação na ferramenta Developer. Mostra eventos de inicialização, execução e resumo.

#### **Exibição Ajuda de Scripts**

Mostra a ajuda contextual para o editor de Script.

#### **Exibição Visualizador de Dados**

Mostra dados de entrada de exemplo, executa a transformação e exibe os resultados da saída.

## Portas da Transformação do Processador de Dados

Defina as portas da transformação do Processador de Dados na exibição **Visão Geral** da transformação.

Uma transformação do Processador de Dados pode ler a entrada de um arquivo, um buffer ou um buffer com fluxo de um leitor de arquivos complexos. Você pode usar um leitor de arquivos simples como buffer para ler um arquivo inteiro de uma só vez. Você também pode ler um arquivo de entrada a partir de um banco de dados.

As portas de saída que você cria dependem de se você deseja retornar uma string, arquivos complexos ou linhas de dados relacionais a partir da transformação.

## Portas de Entrada da Transformação do Processador de Dados

Quando você cria uma transformação do Processador de Dados, a ferramenta Developer cria uma porta de entrada padrão. Quando você define uma porta de entrada adicional em um componente de inicialização do Script, a ferramenta Developer cria uma porta de entrada adicional na transformação.

O tipo de entrada determina o tipo de dados que o Serviço de Integração de Dados transmite para a transformação do Processador de Dados. O tipo da entrada determina se a entrada é um caminho de arquivo de origem ou dados.

Configure um dos seguintes tipos de entrada:

#### **Buffer**

A transformação do Processador de Dados recebe linhas de dados de origem na porta de Entrada. Use o tipo de entrada de buffer ao configurar a transformação para receber dados de um arquivo simples ou de uma transformação Informatica.

#### **Arquivo**

A transformação do Processador de Dados recebe o caminho do arquivo de origem na porta de Entrada. O componente de inicialização do Processador de Dados abre o arquivo de origem. Use o tipo de entrada de arquivo para analisar arquivos binários, como arquivos do Microsoft Excel ou do Microsoft Word. Também é possível usar o tipo de entrada de Arquivo para arquivos extensos que podem exigir muita memória do sistema para processamento com uma porta de entrada de buffer.

### Parâmetro de Serviço

A transformação do Processador de Dados recebe valores a serem aplicados a variáveis nas portas de parâmetros de serviço. Quando você escolhe as variáveis para receber dados de entrada, a ferramenta Developer cria uma porta de parâmetro de serviço para cada variável.

### Output\_Filename

Quando você configura a porta de saída padrão para retornar um nome de arquivo em vez de dados de linha, a ferramenta Developer cria uma porta Output\_Filename. É possível transmitir um nome de arquivo para a porta Output\_Filename a partir de um mapeamento.

Ao definir uma porta de entrada, você pode definir a localização do arquivo de entrada de exemplo para essa porta. Um arquivo de entrada de exemplo é uma pequena amostra do arquivo de entrada. Faça referência a um arquivo de entrada de exemplo quando você criar Scripts. Também é possível usar o arquivo de entrada de exemplo ao testar a transformação na exibição **Visualizador de Dados**. Defina o arquivo de entrada de exemplo no campo **Localização de Entrada**.

## Portas de Parâmetro de Serviço

Você pode criar portas de entrada que recebem os valores de variáveis. As variáveis podem conter qualquer tipo de dados, como uma string, uma data ou um número. Uma variável também pode conter uma localização para um documento de origem. Você pode fazer referência às variáveis em um componente do Processador de Dados.

Quando você cria uma porta de entrada para uma variável, a ferramenta Desenvolvedor exibe uma lista de variáveis para escolha.

### Criando Portas de Parâmetro de Serviço

Você pode criar portas de entrada que recebem os valores de variáveis. Você também pode remover as portas que cria a partir de variáveis.

1. Abra a exibição **Visão Geral** da transformação do Processador de Dados.
2. Clique em **Escolher**.  
A ferramenta Developer exibe uma lista de variáveis e indica quais variáveis já têm várias portas.
3. Selecione uma ou mais variáveis.  
A ferramenta Developer cria uma porta de entrada de buffer para cada variável que você seleciona. Não é possível modificar a porta.
4. Para remover uma porta que você cria a partir de uma variável, desmarque a seleção de lista de variável. Quando você desmarca a seleção, a ferramenta Developer remove a porta de entrada.

## Portas de Saída de Transformação do Processador de Dados

A transformação do Processador de Dados tem uma porta de saída por padrão. Se você definir portas de saída adicionais a um Script, a ferramenta Developer adicionará as portas à transformação do Processador de Dados. Você pode criar grupos de portas se configurar a transformação para retornar dados relacionais. Você também pode criar portas de parâmetros de serviço e portas de passagem.

### Porta de Saída Padrão

A transformação do Processador de Dados tem uma porta de saída por padrão. Ao criar uma saída relacional, você pode definir grupos de portas de saída relacionadas em vez da porta de saída padrão.



Quando você define uma porta de saída adicional em um componente de Script, a ferramenta Developer acrescenta essa porta à transformação.

Configure um dos seguintes tipos de saída para uma porta de saída padrão:

#### **Buffer**

A transformação do Processador de Dados retorna o XML através da porta de Saída. Escolha o tipo de arquivo de Buffer ao analisar documentos ou ao mapear um XML para outros documentos XML na transformação do Processador de Dados.

#### **Arquivo**

O Serviço de Integração de Dados retorna um nome de arquivo de saída na porta de Saída para cada linha ou instância de origem. O componente da transformação do Processador de Dados grava o arquivo de saída em vez de retornar dados através das portas de saída da transformação do Processador de Dados.

Quando você seleciona uma porta de saída de Arquivo, a ferramenta Developer cria uma porta de entrada Output\_Filename. É possível transmitir um nome de arquivo para a porta de nome de arquivo de Saída. A transformação do Processador de Dados cria o arquivo de saída com um nome que ela recebe nessa porta.

Se o nome do arquivo de saída estiver em branco, o Serviço de Integração de Dados retornará um erro de linha. Quando um erro ocorre, o Serviço de Integração de Dados grava um valor nulo na porta de Saída e retorna um erro de linha.

Escolha o tipo de saída de Arquivo ao transformar XML em um arquivo de dados binários, como um arquivo PDF ou um arquivo do Microsoft Excel.

## Portas de Passagem

Você pode configurar portas de passagem para qualquer transformação do Processador de Dados. As portas de passagem são portas de entrada e saída que recebem dados de entrada e retornam os mesmos dados para um mapeamento sem alterá-los.

Você pode configurar portas de passagem em uma instância de transformação do Processador de Dados que está em um mapeamento.

Para adicionar uma porta de passagem, arraste uma porta de outra transformação no mapeamento. Você também pode adicionar portas na guia **Portas** da exibição **Propriedades**. Clique em **Novo** para adicionar uma porta de passagem.

**Nota:** Quando você adicionar portas de passagem em uma transformação do Processador de Dados com entrada relacional e saída hierárquica, adicione as portas ao grupo raiz da estrutura relacional.

As transformações do Processador de Dados podem incluir portas de passagem com tipos de dados personalizados.

## Componente de Inicialização

Um componente de inicialização define o componente que inicia o processamento na transformação do Processador de Dados. Configure o componente de inicialização na exibição **Visão Geral**.

Uma transformação do Processador de Dados pode conter vários componentes para processar dados. Cada componente pode conter outros componentes. É necessário identificar qual componente é o ponto de entrada para a transformação.

Ao configurar o componente de inicialização em uma transformação do Processador de Dados, você pode escolher um componente de XMap, Biblioteca ou Script como o componente de inicialização. Em termos de Scripts, você pode selecionar um dos seguintes tipos de componentes:

- Analisador. Converte documentos de origem em XML. A entrada pode ter qualquer formato, como texto, HTML, Word, PDF ou HL7.
- Mapeador. Converte um documento de origem XML em outra estrutura ou esquema XML.
- Serializador. Converte um arquivo XML em um documento de saída de qualquer formato.
- Streamer. Divide grandes documentos de entrada, como fluxos de dados multigigabytes, em segmentos.
- Transformador. Modifica os dados em qualquer formato. Adiciona, remove, converte ou altera o texto. Use Transformadores com um Analisador, um Mapeador ou um Serializador. Também você pode executar um Transformador como um componente autônomo.

**Nota:** Se o componente de inicialização não for um XMap ou uma Biblioteca, você também poderá configurá-lo em um Script, em vez de na exibição **Visão Geral**.

## Referências

Você pode definir as referências da transformação, como as referência de esquema ou de mapplet, selecionando um esquema ou um mapplet para servir como uma referência. Algumas transformações de Processador de Dados exigem um esquema hierárquico para definir a hierarquia de entrada ou de saída de componentes relevantes na transformação. Para usar o esquema na transformação, defina uma referência de esquema para a transformação. Você também pode usar uma ação especializada denominada ação RunMapplet para chamar um mapplet de uma transformação do Processador de Dados. Para chamar um mapplet, primeiro é necessário definir uma referência de mapplet para a transformação.

Você pode definir as referências da transformação, como as referências de esquema ou de mapplet, na exibição **Referências** da transformação.

### Referências de Esquema

A transformação do Processador de Dados faz referência aos objetos de esquema no repositório do Modelo. Os objetos de esquema podem existir no repositório antes de você criar a transformação. Você também pode importar os esquemas da exibição **Referências** da transformação.

A codificação do esquema deve corresponder à codificação de entrada dos objetos Serializador ou Mapeador. A codificação do esquema deve corresponder à codificação de saída dos objetos Analisador ou Mapeador. Configure a codificação de trabalho na exibição **Configurações** da transformação.

Um esquema pode fazer referência a esquemas adicionais. A ferramenta Developer mostra o espaço de nome e o prefixo de cada esquema ao qual a transformação do Processador de Dados faz referência. Quando você faz referência a vários esquemas com espaços de nome vazios, a transformação não é válida.

### Referências de Mapplet

Você pode chamar um mapplet de uma transformação do Processador de Dados com a ação RunMapplet. Antes de adicionar a ação RunMapplet a um componente da transformação do Processador de Dados, primeiro defina uma referência para o mapplet que você deseja chamar.

# Configurações da Transformação do Processador de Dados

Configure as páginas de código, as opções de processamento de XML e as configurações de log na exibição **Configurações** da transformação do Processador de Dados.

## Codificação de Caracteres

Uma codificação de caracteres é um mapeamento de caracteres de um idioma ou grupo de idiomas para código hexadecimal.

Ao criar um Script, você define a codificação dos documentos de entrada e a codificação dos documentos de saída. Defina a codificação de trabalho para especificar como o editor do IntelliScript exibe caracteres e como a transformação do Processador de Dados processa esses caracteres.

### Codificação de Trabalho

A codificação de trabalho é a página de código para os dados na memória e a página de código para os dados que aparecem na interface do usuário e em arquivos de trabalho. É necessário selecionar uma codificação de trabalho que seja compatível com a codificação dos esquemas aos quais você faz referência na transformação do Processador de Dados.

A seguinte tabela mostra as configurações da codificação de trabalho:

Configuração	Descrição
Usar a Página de Código Padrão do Processador de Dados	Usa a codificação padrão da transformação do Processador de Dados.
Outros	Selecione a codificação na lista.
Codificação de Caracteres Especiais XML	<p>Determina a representação de caracteres especiais XML. Você pode selecionar <b>Nenhuma</b> ou <b>XML</b>.</p> <ul style="list-style-type: none"><li>- <b>Nenhuma.</b> Deixar como <code>&amp;amp;</code>, <code>&amp;lt;</code>, <code>&amp;gt;</code>, <code>&amp;quot;</code>, <code>&amp;apos;</code>; Referências de entidade para caracteres especiais XML são interpretadas como texto. Por exemplo, o caractere <code>&gt;</code> é exibido como <code>&amp;gt;</code>; O padrão é nenhuma.</li><li>- <b>XML.</b> Converter em <code>&amp;</code> e <code>&lt;</code> e <code>&gt;</code> e <code>"</code> e <code>'</code> Referências de entidade para caracteres especiais XML são interpretadas como caracteres regulares. Por exemplo, <code>&amp;gt;</code> aparece como o caractere a seguir: &gt;</li></ul>

### Codificação de Entrada

A codificação de entrada determina como os dados de caracteres são codificados em documentos de entrada. É possível configurar a codificação para portas de entrada adicionais em um Script.

A seguinte tabela descreve as configurações de codificação na área **Entrada**:

Configuração	Descrição
Usar a Codificação Especificada no Documento de Entrada	Use a página de código definida pelo documento de origem, como o atributo de codificação de um documento XML. Se o documento de origem não tiver uma especificação de codificação, a transformação do Processador de Dados usará as configurações de codificação da exibição <b>Configurações</b> .
Usar Codificação de Trabalho	Use a mesma codificação que a codificação de trabalho.
Outros	Selecione a codificação de entrada em uma lista suspensa.
Codificação de Caracteres Especiais XML	Determina a representação de caracteres especiais XML. É possível selecionar <b>Nenhuma</b> ou <b>XML</b> . - <b>Nenhuma.</b> Deixar como & &lt; &gt; &quot; &apos; Referências de entidade para caracteres especiais XML são interpretadas como texto, por exemplo, o caractere > aparece como &gt; O padrão é Nenhuma. - <b>XML.</b> Converter em & < > " ' Referências de entidade para caracteres especiais XML são interpretadas como caracteres regulares. Por exemplo, &gt; aparece como o caractere a seguir: >
Ordem de Bytes	Descreve como os caracteres de vários bytes aparecem no documento de entrada. Você pode selecionar as seguintes opções: - Little-endian. O byte menos significativo aparece em primeiro lugar. Padrão. - Big-endian. O byte mais significativo aparece em primeiro lugar. - Nenhuma conversão binária.

## Codificação de Saída

A codificação de saída determina como os dados de caracteres são codificados no documento de saída principal.

A seguinte tabela descreve as configurações de codificação na área **Saída**:

Configuração	Descrição
Usar Codificação de Trabalho	A codificação de saída é igual à codificação de trabalho.
Outros	O usuário seleciona a codificação de saída na lista.

Configuração	Descrição
Codificação de Caracteres Especiais XML	<p>Determina a representação de caracteres especiais XML. Você pode selecionar <b>Nenhuma</b> ou <b>XML</b>.</p> <ul style="list-style-type: none"> <li>- <b>Nenhuma.</b> Deixar como &amp;amp; &amp;lt; &amp;gt; &amp;quot; &amp;apos; Referências de entidade para caracteres especiais XML são interpretadas como texto, por exemplo, o caractere &gt; aparece como &amp;gt;; Padrão.</li> <li>- <b>XML.</b> Converter em &amp; &lt; &gt; " ' " Referências de entidade para caracteres especiais XML são interpretadas como caracteres regulares. Por exemplo, &amp;gt; aparece como o caractere a seguir: &gt;</li> </ul>
Igual à Codificação de Entrada	A codificação de saída é igual à codificação de entrada.
Ordem de bytes	<p>Descreve como os caracteres de vários bytes aparecem no documento de entrada. Você pode selecionar as seguintes opções:</p> <ul style="list-style-type: none"> <li>- Little-endian. O byte menos significativo aparece em primeiro lugar. Padrão.</li> <li>- Big-endian. O byte mais significativo aparece em primeiro lugar.</li> <li>- Nenhuma conversão binária.</li> </ul>

## Regras e Diretrizes da Codificação de Caracteres

Use as seguintes regras e diretrizes ao configurar codificações:

- Para aumentar o desempenho, defina a codificação de trabalho para ser igual à codificação do documento de saída.
- Defina a codificação de entrada para a mesma codificação do documento de entrada.
- Defina a codificação de saída para a mesma codificação do documento de saída.
- Para idiomas que têm caracteres de bytes múltiplos, defina a codificação de trabalho para UTF-8. Para a codificação de entrada e saída, você pode usar uma codificação Unicode, como UTF-8, ou uma página de código de byte duplo, como Big5 ou Shift\_JIS.

## Configurações de Saída

Defina configurações de controle de saída para controlar se a transformação do Processador de Dados cria logs de eventos e salva documentos de saída.

É possível controlar os tipos de mensagens que a transformação do Processador de Dados grava no log de eventos em tempo de design. Se você salvar os documentos de entrada analisados com os logs de eventos, poderá visualizar o contexto em que o erro ocorreu na exibição **Evento**.

A seguinte tabela descreve as configurações na área **Eventos de Tempo de Design**:

Configuração	Descrição
Registrar eventos de tempo de design	Determina se um log de eventos em tempo de design deve ou não ser criado. Por padrão, a transformação do Processador de Dados registra notificações, avisos e falhas no log de eventos em tempo de design. É possível excluir os seguintes tipos de eventos: <ul style="list-style-type: none"> <li>- Notificações</li> <li>- Avisos</li> <li>- Falhas</li> </ul>
Salvar documentos analisados	Determina quando a transformação do Processador de Dados salva um documento de entrada analisado. Você pode selecionar as seguintes opções: <ul style="list-style-type: none"> <li>- Sempre.</li> <li>- Nunca</li> <li>- Em caso de falha</li> </ul> O padrão é sempre.

A seguinte tabela descreve as configurações na área **Eventos de Tempo de Execução**:

Configuração	Descrição
Registrar eventos de tempo de execução	Determina se um log de eventos é criado quando você executa a transformação por meio de um mapeamento. <ul style="list-style-type: none"> <li>- Nunca.</li> <li>- Em caso de falha</li> </ul> O padrão é Nunca.

A seguinte tabela descreve as configurações na área **Saída**:

Configuração	Descrição
Desativar saída automática	Determina se a transformação do Processador de Dados grava a saída no arquivo de saída padrão. Desabilite a saída padrão nas seguintes situações: <ul style="list-style-type: none"> <li>- Você transmite a saída de um Analisador para a entrada de outro componente antes de a transformação criar um arquivo de saída.</li> <li>- Você usa uma ação <b>WriteValue</b> para gravar dados diretamente na saída de um Script, em vez de transmitir dados através das portas de saída.</li> </ul>
Desabilitar compactação de valor	Determina se a transformação do Processador de Dados usa a compactação de valores para otimizar o uso da memória. Importante: Não desabilite a compactação de valores, exceto quando o Suporte Global a Clientes da Informatica o aconselhar a fazer isso.

A tabela a seguir descreve as configurações na área **Modo de coleta de porta de saída binária**. Você pode selecionar uma dessas opções de saída binária para uma transformação de relacional para hierárquica com

saída XML, Avro ou Parquet ou para um Analisador de transformação de Processador de Dados com saída Avro ou Parquet.

Configuração	Descrição
Coletar linhas de entrada para uma única saída	Determina se a transformação de Processador de Dados acumula a entrada relacional em uma única porta de saída binária.
Dividir saída quando o tamanho exceder	Determina se a transformação de Processador de Dados divide a saída em partes com base em um tamanho máximo de saída declarado.
Linha de saída para cada linha (não coletar)	Determina se a transformação de Processador de Dados transmite a saída em linhas separadas.

## Configurações de Processamento

As configurações de processamento definem como a transformação do Processador de Dados processa um elemento sem um tipo de dados definido. Essas configurações afetam Scripts. Elas não afetam os elementos que são processados por um XMap.

A seguinte tabela descreve as configurações de processamento que afetam o processamento do XML em Scripts:

Configuração	Descrição
Tratar como xs:string	A transformação do Processador de Dados trata um elemento sem tipo como uma cadeia. Na caixa de diálogo <b>Escolher XPath</b> , o elemento ou o atributo aparece como um único nó.
Tratar como xs:anyType	A transformação do Processador de Dados trata um elemento sem tipo como anyType. Na caixa de diálogo <b>Escolher XPath</b> , o elemento ou atributo aparece como uma árvore de nós. Um nó é do tipo xs:string, e todos os tipos de dados complexos nomeados aparecem como nós de árvore.

A seguinte tabela descreve uma configuração de processamento que afeta o processamento de Streamer:

Configuração	Descrição
Tamanho do bloco do streamer	Essa configuração define a quantidade de dados que o Streamer lê a cada vez de um fluxo de arquivo de entrada. A transformação do Processador de Dados aplica essa configuração a um Streamer com um arquivo de entrada.

A seguinte tabela descreve uma configuração de processamento que afeta o processamento da transformação de hierárquico para relacional:

Configuração	Descrição
Impor validação estrita	Essa configuração determina se a transformação de Processador de Dados realiza uma validação estrita para a entrada hierárquica. Quando a validação estrita é aplicada, o arquivo de entrada hierárquica deve estar em conformidade estrita com seu esquema. Essa opção pode ser aplicada quando o modo do Processador de Dados é definido como <b>Mapeamento de Saída</b> , o que cria portas de saída para saída relacional.  Esta opção não é aplicada a mapeamentos com entrada JSON de versões anteriores à versão 10.2.1.
Normalizar entrada XML	Essa configuração determina se a transformação do Processador de Dados normaliza a entrada XML. Por padrão, a transformação realiza a normalização da entrada XML. Em alguns casos, você pode optar por ignorar a normalização automática para melhorar o desempenho.

## Configurações de XMap

A configuração de XMap define como os processos de transformação do Processador de Dados processam os elementos de entrada XMap que não são transformados em elementos de saída. Os elementos não lidos são transmitidos para um porta denominada **XMap\_Unread\_Input\_Values**. A configuração entra em vigor somente quando o XMap é selecionado como o componente de inicialização. A configuração não afeta os elementos que o XMap processa.

Para transmitir elementos XMap não lidos para uma porta dedicada, ative a configuração **Gravar elementos não lidos em uma porta de saída adicional**.

## Configuração de Saída XML

As configurações de geração de XML definem as características de documentos de saída XML.

A seguinte tabela descreve as configurações de geração de XML na área **Título do Esquema**:

Configuração	Descrição
Localização do esquema	Define a schemaLocation para o elemento raiz do principal documento de saída.
Nenhuma localização de esquema de espaço de nome	Define o atributo xsi:noNamespaceSchemaLocation do elemento raiz do principal documento de saída.



Defina as configurações do Modo de Saída XML para determinar como a transformação do Processador de Dados lida com elementos ou atributos ausentes no documento XML de entrada. A seguinte tabela descreve as configurações de geração de XML na área **Modo de Saída XML**:

Configuração	Descrição
No estado em que se encontra	Não adicione ou remova elementos vazios. O padrão é ativada.
Completo	Todos os elementos obrigatórios e opcionais definidos no esquema de saída são gravados na saída. Os elementos que não têm conteúdo são gravados como elementos vazios.
Compacto	Remove os elementos vazios da saída. Se <b>Adicionar a Elementos</b> estiver ativada, a transformação do Processador de Dados removerá apenas os elementos opcionais. Se <b>Adicionar a Elementos</b> estiver desativada, a transformação do Processador de Dados removerá todos os elementos vazios. A saída XML pode não ser válida.

A seguinte tabela descreve as configurações de geração de XML na área **Valores Padrão para Nós Obrigatórios**:

Configuração	Descrição
Adicionar a elementos	Quando o esquema de saída define um valor padrão para um elemento necessário, a saída inclui o elemento com um valor padrão. O padrão é ativada.
Adicionar a atributos	Quando o esquema de saída define um valor padrão para um atributo necessário, a saída inclui o atributo com seu valor padrão. O padrão é ativada.
Validar valores adicionados	Determina se a transformação do Processador de Dados valida os elementos vazios que são adicionados pela saída do modo Completo. O padrão é desativada. Se a opção <b>Validar valores adicionados</b> estiver ativada e o esquema não permitir elementos vazios, a saída XML poderá não ser válida.

A seguinte tabela descreve as configurações de geração de XML na área **Instruções de Processamento**:

Configuração	Descrição
Adicionar instruções de processamento XML	Define a codificação de caractere e versão XML do documento de saída. O padrão é marcado.
Versão XML	Define a versão XML. A definição de versão XML tem as seguintes opções: - 1.0 - 1.1 O padrão é 1.0.

Configuração	Descrição
Codificação	Define a codificação de caracteres especificada na instrução de processamento. A configuração de Codificação tem as seguintes opções: <ul style="list-style-type: none"> <li>- Igual à codificação de saída. A codificação de saída na instrução de processamento é igual à codificação de saída definida nas configurações de transformação do Processador de Dados.</li> <li>- Personalizar. Define a codificação de saída na instrução de processamento. O usuário digita o valor no campo.</li> </ul>
Adicionar instruções personalizadas de processamento	Adiciona outras instruções de processamento ao documento de saída. Digite a instrução de processamento exatamente conforme ela aparece no documento de saída. O padrão é desativada.

A seguinte tabela descreve as configurações de geração de XML na área **Raiz XML**:

Configuração	Descrição
Adicionar o elemento raiz XML	Adiciona um elemento raiz ao documento de saída. Use esta opção quando o documento de saída tiver mais de uma ocorrência do elemento raiz definido no esquema de saída. O padrão é desativada.
Nome do elemento raiz	Define um nome para o elemento raiz para adicionar ao documento de saída.

## Eventos

Um evento é um registro de uma etapa de processamento a partir de um componente na transformação do Processador de Dados. Em um Script ou Biblioteca, cada âncora, ação ou transformador gera um evento. Em um XMap, cada instrução de mapeamento gera um evento.

É possível mostrar eventos na exibição **Eventos do Processador de Dados**.

## Tipos de Eventos

A transformação do Processador de Dados grava os eventos em arquivos de log. Cada evento tem um tipo de evento que indica se o evento foi bem-sucedido, se o evento falhou ou se o evento foi executado com erros.

Um componente pode gerar um ou mais eventos. O componente pode passar ou falhar, dependendo se os eventos obtiveram êxito ou falharam. Se um evento falhar, um componente falhará.

A tabela a seguir descreve os tipos de eventos que a transformação do Processador de Dados gera:

Tipo de Evento	Descrição
Notificação	Operação normal.
Aviso	A transformação do Processador de Dados foi executada, mas ocorreu uma condição inesperada. Por exemplo, a transformação do Processador de Dados gravou dados várias vezes no mesmo elemento. Cada vez que o elemento é substituído, a transformação do Processador de Dados gera um aviso.

Tipo de Evento	Descrição
Falha	A transformação do Processador de Dados foi executada, mas um componente falhou. Por exemplo, um elemento de entrada obrigatório estava vazio.
Falha Opcional	A transformação do Processador de Dados foi executada, mas um componente opcional falhou. Por exemplo, uma âncora opcional não foi encontrada no documento de origem.
Erro Fatal	A transformação do Processador de Dados falhou devido a um erro grave. Por exemplo, o documento de entrada não existia.

## Exibição de Eventos do Processador de Dados

A exibição **Eventos do Processador de Dados** mostra eventos quando você executa uma transformação do Processador de Dados a partir da ferramenta Developer.

A exibição **Eventos do Processador de Dados** possui um painel **Navegação** e um painel **Detalhes**. O painel **Navegação** contém uma árvore de navegação. A árvore de navegação lista os componentes que a transformação executou, em ordem cronológica. Cada nó tem um ícone que representa o evento mais grave abaixo dele na árvore. Quando você seleciona um nó no painel **Navegação**, os eventos aparecem no painel **Detalhes**.

A árvore de navegação contém os seguintes nós de nível superior:

- Inicialização de Serviço. Descreve os arquivos e as variáveis que a transformação do Processador de Dados inicializa.
- Execução. Lista os componentes que o Script, a Biblioteca ou o XMap executou.
- Resumo. Mostra estatísticas sobre o processamento.

Quando você executa um XMap, cada nome de nó no painel de navegação tem um número entre colchetes, como [5]. Para identificar a instrução que gerou os eventos do nó, clique com o botão direito na grade de instruções e selecione Ir para Número de Linha. Insira o número do nó.

Quando você executa um Script e clica duas vezes em um evento no painel **Navegação** ou no painel **Detalhes**, o editor de Script realça o componente de Script que gerou esse evento. O painel **Entrada** da exibição **Visualizador de Dados** realça a parte do exemplo de documento de origem que gerou o evento.

## Logs

Um log contém um registro da transformação do Processador de Dados. A transformação do Processador de Dados grava eventos em logs.

A transformação do Processador de Dados cria os seguintes tipos de logs:

### Log de eventos de tempo de design

O log de eventos de tempo de design contém eventos que ocorrem quando você executa a transformação do Processador de Dados na exibição **Visualizador de Dados**. Visualize o log de tempo de design na exibição **Eventos**.

### Log de eventos de tempo de execução

O log de eventos de tempo de execução contém eventos que ocorrem quando você executa a transformação do Processador de Dados em um mapeamento. É possível visualizar esse log em um editor de texto ou arrastá-lo até a exibição **Eventos** da transformação do Processador de Dados.

### Log do usuário

O log do usuário contém eventos que você configura para componentes em um Script. A transformação do Processador de Dados grava no log do usuário ao ser executada a partir da exibição **Visualizador de Dados** e quando você a executa em um mapeamento. É possível exibir o log do usuário em um editor de texto.

## Log de Eventos de Tempo de Design

O log de eventos de tempo de design contém eventos que ocorrem quando você executa a transformação do Processador de Dados na exibição do **Visualizador de Dados** da ferramenta Developer.

Quando você executar uma transformação do Processador de Dados na exibição do **Visualizador de Dados** o log de eventos de tempo de design será mostrado na exibição de **Eventos do Processador**. Por padrão, o log de eventos de tempo de design contém notificações, avisos e falhas. Nas configurações de transformação, você pode configurar a transformação do Processador de Dados para excluir um ou mais tipos de eventos do log.

Ao salvar os documentos de entrada com o log, você poderá clicar em um evento na exibição de **Eventos do Processador de Dados** para achar a localização no documento de entrada que gerou o evento. Ao configurar a transformação do Processador de Dados, você poderá optar por salvar os arquivos de entrada de cada execução ou somente em falha.

O log de eventos de tempo de design é denominado `events.cme`. Você pode localizar o log de eventos de tempo de design da última execução da transformação do Processador de Dados no seguinte diretório:

```
C:\<Installation_directory>\clients\DT\CMReports\Init\events.cme
```

A transformação do Processador de Dados substituirá o log de eventos de tempo de design sempre que executar a transformação no **Visualizador de Dados**. Renomeie log de eventos de tempo de design se desejar exibi-lo após a execução mais recente da transformação ou se desejar comparar os logs de diferentes execuções. Quando você fechar a ferramenta Developer, ela não salvará os arquivos em

## Log de Eventos de Tempo de Execução

O log de evento de tempo de execução grava os eventos que ocorrem quando você executa a transformação do Processador de Dados em um mapeamento.

Se a transformação do Processador de Dados conclui a execução sem falhas, ele não grava um log de eventos. Se houver falhas na execução, a transformação do Processador de Dados executa uma segunda vez e grava um log de eventos durante a segunda execução. O log de eventos de tempo de execução é denominado `events.cme`.

Em uma máquina Windows, o log de eventos no tempo de execução está no seguinte diretório:

```
C:\<Installation_Directory>\clients\DT\CMReports\Tmp\
```

Em uma máquina Linux ou UNIX, o log de eventos de tempo de execução para um usuário raiz no seguinte diretório:

```
/root/<Installation_Dirctory>/clients/DT/CMReports/Tmp
```

Em uma máquina Linux ou UNIX, você pode localizar o log de eventos de tempo de execução para um usuário que não é raiz no seguinte diretório:

```
/home/[UserName]/<Installation_Directory>/DT/CMReports/Tmp
```

Use o editor de configuração para alterar a localização do log de eventos de tempo de execução.

## Exibindo um Log de Eventos na Exibição de Eventos do Processador de Dados

Use a exibição **Eventos do Processador de Dados** para exibir um log de eventos de tempo de criação ou um log de eventos de tempo de execução.

Abra o Windows Explorer e navegue para o arquivo de log de eventos que você deseja exibir. Arraste o log da janela do Windows Explorer para a exibição **Eventos do Processador de Dados**. Clique com o botão direito do mouse na exibição **Eventos do Processador de Dados** e selecione **Localizar** para pesquisar no log.

**Nota:** Para recarregar o log de eventos de tempo de design mais recente, clique com o botão direito do mouse na exibição **Eventos do Processador de Dados** e selecione **Recarregar Eventos do Projeto**.

## Log do Usuário

O log do usuário contém mensagens personalizadas que você configura sobre falhas de componentes em um Script.

A transformação do Processador de Dados grava mensagens no log do usuário quando um Script é executado a partir da exibição **Visualizador de Dados** e quando você a executa em um mapeamento.

Quando um componente de Script tem a propriedade **on\_fail**, você poderá configurá-lo para gravar uma mensagem no log do usuário se ele falhar. No Script, defina a propriedade **on\_fail** como um dos seguintes valores:

- LogInfo
- LogWarning
- LogError

Cada execução do Script gera um novo log do usuário. O nome de arquivo do log do usuário contém o nome da transformação com um GUID exclusivo:

```
<Transformation_Name>_<GUID>.log
```

Por exemplo, CalculateValue\_Aa93a9d14-a01f-442a-b9cb-c9ba5541b538.log

Em uma máquina Windows, o log do usuário está localizado no seguinte diretório:

```
c:\Users\[UserName]\AppData\Roaming\Informatica\DataTransformation\UserLogs
```

Em uma máquina Linux ou UNIX, o log para o usuário root está localizado no seguinte diretório:

```
/<Installation_Directory>/DataTransformation/UserLogs
```

Em uma máquina Linux ou UNIX, o log para o usuário não root está localizado no seguinte diretório:

```
home/<Installation_Directory>/DataTransformation/UserLogs
```

# Desenvolvimento da Transformação do Processador de Dados

Use o assistente de Nova Transformação para gerar automaticamente uma transformação do Processador de Dados ou criar uma transformação do Processador de Dados em branco e para configuração posterior. Se você criar uma transformação do Processador de Dados em branco, deverá selecionar para criar um objeto Script, XMap, Biblioteca ou Regras de Validação na transformação. Um Script pode analisar documentos de origem em formato hierárquico, converter do formato hierárquico em outros formatos de arquivo ou mapear um documento hierárquico para outro formato hierárquico. Um XMap converte um arquivo hierárquico de entrada em um arquivo hierárquico de saída de outra estrutura. Uma Biblioteca transforma um tipo de mensagem de indústria em um documento XML com uma estrutura hierárquica, ou de XML em um formato padrão da indústria. Escolha os esquemas que definam as hierarquias de entrada ou saída.

1. Crie a transformação na Developer tool.
2. Para uma transformação do Processador de Dados em branco, execute as seguintes etapas adicionais:
  - a. Adicione referências de esquema que definem as hierarquias XML de entrada ou saída.
  - b. Crie um objeto Script, XMap, Biblioteca ou Regras de Validação.
3. Configure as portas de entrada e de saída.
4. Teste a transformação.

## Criar a Transformação do Processador de Dados

Crie uma transformação do Processador de Dados na Developer tool. Se você criar uma transformação do Processador de Dados em branco, deverá criar um objeto Script, XMap, Biblioteca ou Regras de Validação na transformação. Como alternativa, você pode usar o assistente de Nova Transformação para gerar automaticamente uma transformação do Processador de Dados.

1. Na Developer tool, clique em **Arquivo > Nova > Transformação**.
2. Selecione a transformação do Processador de Dados e clique em **Avançar**.
3. Insira um nome para a transformação e procure uma localização do repositório do Modelo para inserir a transformação.
4. Selecione se você deseja criar a transformação do Processador de Dados usando um assistente ou criar uma transformação do Processador de Dados em branco.
5. Se você preferir criar uma transformação do Processador de Dados em branco, clique em **Concluir**.  
A Developer tool cria a transformação em branco no repositório. A exibição **Visão Geral** aparece na Developer tool.
6. Se você preferir criar uma transformação do Processador de Dados usando um assistente, realize as seguintes etapas:
  - a. Clique em **Avançar**.
  - b. Selecione um formato de entrada.
  - c. Navegue para selecionar um esquema, copybook, arquivo de exemplo, ou arquivo de especificação, se necessário para determinados formatos de entrada, como COBOL, JSON ou ASN.1.
  - d. Selecione um formato de saída.
  - e. Navegue para selecionar um esquema, copybook, arquivo de exemplo, ou arquivo de especificação, se necessário para o formato de saída.

f. Clique em **Concluir**. O assistente cria a transformação no repositório.

A transformação pode conter um Analisador, um Serializador, um Mapeador ou um objeto com componentes comuns. Se você selecionar um esquema, um copybook, um arquivo de exemplo ou um arquivo de especificação, o assistente também criará um esquema no repositório que seja equivalente à hierarquia no arquivo.

## Selecionar os Objetos de Esquema

Escolha os objetos de esquema que definem as hierarquias de entrada ou de saída para cada componente de XMap ou de Script que você planeja criar.

É possível adicionar referências de esquema na exibição **Referências** ou adicionar essas referências ao criar objetos de Script ou XMap. Um objeto de esquema deve existir no repositório do Modelo para que você possa fazer referência a ele em um Script ou XMap.

1. Na exibição **Referências** da transformação do Processador de Dados, clique em **Adicionar**.
2. Se o objeto de esquema existir no repositório do Modelo, navegue até ele e selecione o esquema.
3. Se o esquema não existir no repositório do modelo, clique em **Criar um novo objeto de esquema** e importe um objeto de esquema de um arquivo de esquema hierárquico.
4. Clique em **Concluir** para adicionar a referência de esquema à transformação do Processador de Dados.

## Criar Objetos em uma Transformação do Processador de Dados em Branco

Crie um objeto Script, Biblioteca, XMap ou Regras de Validação na exibição **Objetos** da transformação do Processador de Dados. Após criar o objeto, você pode abrir o objeto na exibição **Objetos** para configurá-lo.

### Criando um Script

Crie um objeto de Script e defina o tipo de componente de Script a ser criado. Opcionalmente, é possível definir uma referência de esquema e um arquivo de origem de exemplo.

1. Na exibição **Objetos** da transformação do Processador de Dados, clique em **Novo**.
2. Insira um nome para o Script e clique em **Avançar**.
3. Escolha a criação de um Analisador ou de um Serializador. Selecione Outros para criar um componente Mapeador, Transformador ou Streamer.
4. Insira um nome para o componente.
5. Se o componente for o primeiro a processar dados na transformação, habilite **Definir como componente de inicialização**.
6. Clique em **Avançar** se quiser inserir uma referência de esquema para esse Script. Clique em **Concluir** se não quiser inserir a referência de esquema.
7. Se você optar por criar uma referência de esquema, selecione **Adicionar referência a um objeto de esquema** e procure o objeto de Esquema no repositório do Modelo. Clique em **Criar um novo objeto de esquema** para criar o objeto de Esquema no repositório do Modelo.
8. Clique em **Avançar** para inserir uma referência de origem de exemplo ou para inserir um texto de exemplo. Clique em **Concluir** se não quiser definir uma origem de exemplo.  
Use uma origem de exemplo para definir dados de amostra e para testar o Script.
9. Se você optar por selecionar uma origem de exemplo, selecione **Arquivo** e procure o arquivo de amostra.

Também é possível inserir um texto de amostra na área **Texto**. A Developer tool usa esse texto para testar um Script.

10. Clique em **Concluir**.

A exibição **Script** aparece no editor da Developer tool.

## Criando um XMap

Crie um XMap na exibição de **Objetos** da Data Transformation. Ao criar um XMap, você deve ter um esquema que descreve a entrada e a saída dos documentos hierárquicos. Selecione o elemento no esquema que é o elemento raiz da hierarquia de entrada.

1. Na exibição **Objetos** da transformação do Processador de Dados, clique em **Novo**.
2. Selecione XMap e clique em **Avançar**.
3. Insira um nome para o XMap.
4. Se o componente XMap for o primeiro componente a processar dados na transformação, ative **Definir como componente de inicialização**.

Clique em **Avançar**.

5. Se você optar por criar uma referência de esquema, selecione **Adicionar referência a um Objeto de Esquema** e procure o objeto de Esquema no repositório do Modelo.

Para importar um novo objeto de Esquema, clique em **Criar um novo objeto de esquema**.

6. Se você tiver um arquivo hierárquico de amostra que possa usar para testar o XMap, procure e selecione o arquivo no sistema de arquivos.

Você pode alterar o arquivo hierárquico de amostra.

7. Escolha a raiz da hierarquia de entrada.

Na caixa de diálogo **Seleção de Elemento Raiz**, selecione o elemento no esquema que é o elemento raiz do arquivo hierárquico de entrada. Você pode pesquisar um elemento no esquema. Você pode usar pesquisa padrão. Digite `*<string>` para fazer a correspondência de qualquer número de caracteres na string. Digite `?<character>` para fazer a correspondência de um único caractere.

8. Clique em **Concluir**.

A ferramenta Developer cria uma exibição para cada XMap que você criar. Clique em **exibir** para configurar o mapeamento.

## Criando uma Biblioteca

Crie um objeto Biblioteca na exibição **Objetos** da Data Transformation. Selecione o tipo de mensagem, o componente e o nome. Opcionalmente, você pode definir um arquivo de origem de tipo de mensagem de exemplo que pode usar para testar o objeto Biblioteca.

Antes de criar uma Biblioteca na transformação do Processador de Dados, instale o pacote de software da biblioteca no seu computador.

1. Na exibição **Objetos** da transformação do Processador de Dados, clique em **Novo**.
2. Selecione a Biblioteca e clique em **Avançar**.
3. Procure e selecione o tipo de mensagem.
4. Escolha a criação de um Analisador ou de um Serializador.

Crie um Analisador se a entrada do objeto de Biblioteca for um tipo de mensagem e a saída for XML.

Crie um Serializador se a entrada do objeto de Biblioteca for XML e a saída for um tipo de mensagem.



5. Se a Biblioteca for o primeiro componente a processar dados na transformação do Processador de Dados, habilite **Definir como componente de inicialização**.  
Clique em **Avançar**.
6. Se você tiver um exemplo de tipo de mensagem do arquivo de origem, pode usá-lo para testar a Biblioteca, procurar e selecionar o arquivo do sistema de arquivos.  
Você pode alterar o arquivo de exemplo.
7. Clique em **Concluir**.  
A Developer tool cria uma exibição para cada tipo de mensagem que você criar. Clique na exibição para acessar o mapeamento.

## Criando Regras de Validação

Crie um objeto Regras de Validação na exibição **Objetos** da transformação do Processador de Dados.

1. Na exibição **Objetos** da transformação do Processador de Dados, clique em **Novo**.
2. Selecione as Regras de Validação e clique em **Avançar**.
3. Insira um nome para as Regras de Validação.
4. Se você tiver um arquivo XML de amostra que possa usar para testar as Regras de Validação, procure e selecione o arquivo no sistema de arquivos.  
Você pode alterar o arquivo XML de exemplo.
5. Clique em **Concluir**.  
A ferramenta Developer cria um objeto Regras de Validação e abre-o no editor de Regras de Validação.

## Adicionando um Exemplo de Origem

Escolha a origem de exemplo para testar o Script, o XMap, a Biblioteca ou as Regras de Validação que você planeja criar.

Você pode adicionar uma origem de exemplo ao criar um Script, um XMap, uma Biblioteca ou Regras de Validação. Depois de ser selecionada, a origem de exemplo será adicionada ao repositório do Modelo. Devido às limitações do repositório do Modelo, o tamanho do arquivo de origem de exemplo é limitado a 5 MB.

Você pode alterar a origem de exemplo.

## Criar as Portas

Configure as portas de entrada e saída na exibição **Visão Geral**.

Quando você configura portas de entrada ou saída adicionais em um Script, a Developer tool acrescenta portas de entrada e de saída adicionais à transformação por padrão. Você não adiciona portas de entrada na exibição **Visão Geral**.

1. Se quiser retornar linhas de dados de saída em vez de XML, habilite a **Saída Relacional**.  
Quando você habilita a saída relacional, a Developer tool remove a porta de saída padrão.
2. Selecione o tipo de dados da porta de entrada, o tipo de porta, a precisão e a escala.
3. Se você não estiver definindo portas de saída relacionais, defina o tipo de dados da porta de saída, o tipo de porta, a precisão e a escala.

4. Se um Script tiver portas de entrada adicionais, será possível definir a localização do exemplo de arquivo de entrada para essas portas. Clique no botão **Abrir** no campo **Localização de Entrada** para procurar o arquivo.
5. Se você tiver habilitado a saída relacional, clique em **Mapeamento de Saída** para criar as portas de saída.
6. Na exibição Portas, mapeie nós da área **Saída Hierárquica** para campos na área **Portas Relacionais**.

## Testando a Transformação

Testar a transformação do Processador de Dados na exibição **Visualizador de Dados**.

Antes de testar a transformação, verifique se você definiu o componente de inicialização. É possível definir o componente de inicialização em um Script ou selecioná-lo na guia **Visão Geral**. Você também precisa ter escolhido um arquivo de entrada de exemplo para realizar testes.

1. Abra a exibição **Visualizador de Dados**.
  2. Clique em **Executar**.

A Developer tool valida a transformação. Se não houver erros, a Developer tool mostrará o arquivo de exemplo na área **Entrada**. Os resultados da saída aparecem no painel Saída.
  3. Clique em **Mostrar Eventos** para mostrar a exibição **Eventos do Processador de Dados**.
  4. Clique duas vezes em um evento na exibição **Eventos do Processador de Dados** para depurar esse evento no editor de Script.
  5. Clique em **Sincronizar com Editor** para alterar o arquivo de entrada quando você estiver testando vários componentes, cada uma com um arquivo de entrada de exemplo diferente.
- Se você modificar o conteúdo do arquivo de exemplo no sistema de arquivos, as alterações aparecerão na área **Entrada**.

## Importação e Exportação da Transformação do Processador de Dados

Você pode exportar uma transformação do Processador de Dados como serviço e executá-la a partir do repositório da Data Transformation. Você também pode importar um serviço da Data Transformation para a ferramenta Developer. Quando você importa um serviço de Data Transformation, a ferramenta Developer cria uma transformação do Processador de Dados a partir do serviço.

**Nota:** Quando você importa um serviço de Data Transformation para o repositório do Modelo, a ferramenta Developer importa os esquemas associados para o repositório. Se você modificar o esquema no repositório, as alterações poderão não ser exibidas nas referências do esquema da transformação. Você pode fechar e abrir a conexão do repositório do Modelo, ou feche e abra a ferramenta Developer para fazer com que as alterações do esquema apareçam na transformação.

## Exportando a Transformação do Processador de Dados como um Serviço

Você pode exportar a transformação do Processador de Dados como um serviço de Data Transformation. Exporte o serviço para o arquivo de repositório do sistema da máquina onde você deseja executar o serviço.

Você pode executar o serviço com o PowerCenter, com aplicativos definidos pelo usuário ou com o comando CM\_console da Data Transformation.

1. Na exibição **Object Explorer**, clique com o botão direito na transformação do Processador de Dados que você deseja exportar e selecione **Exportar**.  
A caixa de diálogo **Exportar** será exibida.
2. Selecione **Informatica > Exportar Transformação do Processador de Dados** e clique em **Avançar**.  
A página **Selecionar** será exibida.
3. Clique em **Avançar**.  
A página **Selecione o Nome do Serviço e a Pasta de Destino** será exibida.
4. Escolha uma pasta de destino:
  - Para exportar o serviço na máquina que hospeda a Developer tool, clique em **Pasta de Serviço**.
  - Para implantar o serviço em outra máquina, clique em **Pasta**. Navegue até o diretório \ServiceDB na máquina onde você deseja implantar o serviço.
5. Clique em **Concluir**.

## Importando vários serviços do Data Transformation

É possível importar um diretório de serviços do Data Transformation da máquina em que diretório foi salvo. Quando você importa serviços do Data Transformation para o repositório do Modelo do Developer, a Developer tool importa as transformações, os esquemas e os dados de exemplo com os arquivos .cmw. Se for necessário importar muitos serviços, importe um diretório de serviços em vez de um serviço de cada vez.

1. Clique em **Arquivo > Importar**,  
A caixa de diálogo **Importar** é exibida.
2. Selecione **Informatica Importar Serviços do Data Transformation (Pasta)** e clique em **Avançar**.  
A página **Importar Serviço de Data Transformation** será exibida.
3. Navegue até o diretório que você deseja importar.
4. Procure uma localização no Repositório na qual você queira salvar as transformações e clique em **Concluir**.  
A Developer tool importa as transformações, os esquemas e os dados de exemplo com o arquivo .cmw.

## Importando um Serviço de Data Transformation

É possível importar um arquivo .cmw do serviço do Data Transformation para o repositório do Modelo para criar uma transformação de Processador de Dados. A ferramenta Developer importa a transformação, os esquemas e os dados do exemplo com o arquivo .cmw.

1. Clique em **Arquivo > Importar**,  
A caixa de diálogo **Importar** é exibida.
2. Selecione **Informatica Importar Serviço do Data Transformation (Único)** e clique em **Avançar**.  
A página **Importar Serviço de Data Transformation** será exibida.
3. Navegue até o arquivo .cmw de serviço que você deseja importar.  
A ferramenta Developer nomeia a transformação de acordo com o nome de arquivo do serviço. É possível alterar o nome.

4. Procure uma localização no Repositório onde você deseja salvar a transformação e clique em **Concluir**.  
A ferramenta Developer importa a transformação, os esquemas e os dados de exemplo com o arquivo .cmw.
5. Para editar a transformação, clique duas vezes na transformação no **Object Explorer**.

## Exportando um Mapeamento com uma Transformação do Processador de Dados para o PowerCenter

Quando você exporta um mapeamento com uma transformação do Processador de Dados para o PowerCenter, pode exportar os objetos para um arquivo local e importar o mapeamento para o PowerCenter. Como alternativa, você pode exportar o mapeamento diretamente para o repositório do PowerCenter.

1. Na exibição **Object Explorer**, selecione o mapeamento a ser exportado. Clique duas vezes e selecione **Exportar**.  
A caixa de diálogo **Exportar** é exibida.
2. Selecione **Informatica > PowerCenter**.
3. Clique em **Avançar**.  
A caixa de diálogo **Exportar para o PowerCenter** é exibida.
4. Selecione o projeto.
5. Selecione a versão do PowerCenter.
6. Escolha a localização de exportação, um arquivo XML de importação do PowerCenter ou um repositório do PowerCenter.
7. Especifique as opções de exportação.
8. Clique em **Avançar**.  
A ferramenta Developer solicita que você selecione os objetos a serem exportados.
9. Selecione os objetos a serem exportados e clique em **Concluir**.  
A ferramenta Developer exporta os objetos para a localização selecionada. Se você exportou o mapeamento para uma localização, a ferramenta Developer também exporta as transformações de Processador de Dados no mapeamento, como serviços, para uma pasta na localização especificada.
10. Se você exportou o mapeamento para um repositório do PowerCenter, os serviços são exportados para o seguinte caminho de diretório: %temp%\DTServiceExport2PC\  
A função de exportação cria uma pasta separada para cada serviço com o seguinte nome:  
<date><serviceFullName>  
Se a transformação incluir o mapeamento relacional, uma pasta será criada para o mapeamento relacional para hierárquico e uma pasta separada para o mapeamento hierárquico para relacional.
11. Copie a pasta ou as pastas com serviços de transformação do Processador de Dados da localização local onde você exportou os arquivos para a pasta ServiceDB do PowerCenter.
12. Se você exportou o mapeamento para um arquivo XML de importação do PowerCenter, importe o mapeamento para o PowerCenter. Para obter mais informações sobre como importar um objeto para o PowerCenter, consulte *Guia do Repositório do PowerCenter 9.6.0*.

# Transformação de Processador de Dados em um ambiente não nativo

O processamento da transformação de Processador de dados em um ambiente não nativo depende do mecanismo que executa a transformação.

Considere o suporte para os seguintes mecanismos de tempo de execução não nativos:

- Mecanismo Blaze. Suportado sem restrições.
- Mecanismo Spark. Sem suporte em mapeamentos de fluxo ou em lote.
- Mecanismo Databricks Spark. Sem suporte.

## CAPÍTULO 14

# Transformação de Decisão

Este capítulo inclui os seguintes tópicos:

- [Transformação de Decisão - Visão Geral, 258](#)
- [Transformação de Decisão - Funções, 259](#)
- [Transformação de Decisão - Instruções Condicionais , 261](#)
- [Transformação de Decisão - Operadores, 262](#)
- [Tratamento NULL da Transformação de Decisão, 263](#)
- [Configurando uma Estratégia de Decisão , 263](#)
- [Propriedades Avançadas da Transformação de Decisão, 264](#)
- [Transformação de Decisão em um ambiente não nativo, 264](#)

## Transformação de Decisão - Visão Geral

A transformação de Decisão é uma transformação passiva que avalia condições nos dados de entrada e gera uma saída com base nos resultados dessas condições.

Configure uma transformação de Decisão para gerar valores diferentes com base nos valores encontrados nos campos de entrada. Por exemplo, se a receita do cliente for maior que um valor específico, você poderá adicionar a string "Prioridade" ao nome do cliente.

Você pode adicionar várias estratégias de decisão à transformação de Decisão. Cada estratégia avalia uma instrução condicional `IF-THEN-ELSE`. Nessa instrução, você pode usar condições `ELSEIF` ou aninhar instruções `IF-THEN-ELSE` adicionais.

A transformação de Decisão é semelhante à transformação de Expressão, no sentido de que permite que você use instruções condicionais e funções para testar os dados de origem. No entanto, a transformação de Decisão é diferente da transformação de Expressão nos seguintes aspectos:

- A transformação de Decisão usa instruções `IF-THEN-ELSE` para avaliar condições. A transformação de Expressão usa instruções `IIF`.
- A transformação de Decisão contém funções que não estão disponíveis na transformação de Expressão.
- Cada estratégia de decisão pode gerar várias saídas.

# Transformação de Decisão - Funções

A transformação de Decisão fornece acesso às funções predefinidas que você usa para definir estratégias de decisão.

O editor de expressão da transformação de Decisão contém uma pasta Decisão. Essa pasta contém funções que são específicas para a transformação de Decisão. O Editor também contém outras pastas que fornecem acesso às funções de transformação de Expressão.

Quando você clica em uma função no editor de expressão, a transformação exibe o uso e os tipos de dados para a função, além de uma descrição do que a função faz.

**Nota:** Nem todas as funções da transformação de Expressão são compatíveis com a transformação de Decisão. A transformação de Decisão apenas fornece acesso às funções da transformação de Expressão compatíveis.

## Lista de Funções da Transformação de Decisão

- ABS
- ADD\_TO\_DATE
- ASCII
- CEIL
- CHOOSE
- CHR
- CHRCODE
- CONCAT
- CONTAINS
- CONVERT\_BASE
- COS
- COSH
- CRC32
- CUME
- CURDATE
- CURTIME
- DATE\_COMPARE
- DATE\_DIFF
- DATECONVERT
- EXP
- FLOOR
- FV
- GET\_DATE\_PART
- GREATEST
- IN
- INDEXOF
- INITCAP
- INSTR

- IS\_DATE
- IS\_NUMBER
- ISNULL
- LAST\_DAY
- LEAST
- LEFTSTR
- LENGTH
- LN
- LOG
- LOWER
- LPAD
- LTRIM
- MAKE\_DATE\_TIME
- MAX
- MD5
- METAPHONE
- MIN
- MOD
- MONTHCOMPARE
- MOVINGAVG
- MOVINGSUM
- NPER
- PMT
- POWER
- PV
- RAND
- RATE
- REG\_EXTRACT
- REG\_MATCH
- REG\_REPLACE
- REPLACECHR
- REPLACESTR
- REVERSE
- RIGHTSTR
- ROUND
- RPAD
- RTRIM
- SET\_DATE\_PART
- SIGN



- SIN
- SINH
- SOUNDEX
- SQRT
- SUBSTR
- TAN
- TANH
- TIMECOMPARE
- TO\_CHAR
- TO\_DATE
- TO\_FLOAT
- TO\_INTEGER
- TRUNC
- UPPER
- XOR

**Nota:** Use um valor constante para definir o formato de data nas funções CURDATE, DATE\_COMPARE, DATECONVERT, e MONTHCOMPARE na transformação de Decisão.

## Transformação de Decisão - Instruções Condicionais

A transformação de Decisão usa instruções condicionais IF-THEN-ELSE para avaliar os dados de entrada.

Dentro dessas instruções condicionais, você pode usar condições ELSEIF ou aninhar outras instruções IF-THEN-ELSE. As instruções condicionais da transformação de Decisão usam o seguinte formato:

```
// Primary condition
IF <Boolean expression>
THEN <Rule Block>
// Optional - Multiple ELSEIF conditions
ELSEIF <Boolean expression>
THEN <Rule Block>
// Optional ELSE condition
ELSE <Rule Block>
ENDIF
```

Você pode aninhar outras instruções condicionais dentro de um bloco de regras.

# Transformação de Decisão - Operadores

Use os operadores de transformação de Decisão para definir estratégias de decisão.

A seguinte tabela descreve os operadores de transformação de Decisão:

Tipo de Operador	Operador	Descrição
Atribuição	:=	Atribui um valor a uma porta.
Booleano	AND	Adiciona uma condição lógica necessária. Para a expressão Booleana pai ser verdadeira, todas as condições lógicas vinculadas por este operador devem ser verdadeiras.
Booleano	OR	Adiciona uma condição lógica. Para a expressão Booleana pai ser verdadeira, pelo menos uma condição lógica vinculada por este operador deve ser verdadeira.
Booleano	NOT	Especifica uma condição lógica negativa. Para a expressão Booleana pai ser verdadeira, a condição negativa especificada por este operador deve ser verdadeira.
Decisão	=	Testa se os itens comparados são iguais. Use com uma string ou tipos de dados numéricos.
Decisão	<>	Testa se os itens comparados são diferentes. Use com uma string ou tipos de dados numéricos.
Decisão	<	Testa se um valor é menor que outro valor. Use com tipos de dados numéricos.
Decisão	<=	Testa se um valor é menor ou igual a outro valor. Use com tipos de dados numéricos.
Decisão	>	Testa se um valor é maior que outro valor. Use com tipos de dados numéricos.
Decisão	>=	Testa se um valor é maior ou igual a outro valor. Use com tipos de dados numéricos.
Numérico	-	Subtração
Numérico	NEG	Negação
Numérico	+	Adição
Numérico	*	Multiplicação
Numérico	/	Divisão
Numérico	%	Módulo. Retorna o resto após a divisão de um número por outro.
String		Concatena strings.

# Tratamento NULL da Transformação de Decisão

Tratamento NULL determina como o Serviço de Integração de Dados processa o valor NULL em uma transformação de Decisão.

Quando você habilita tratamento NULL, a transformação de Decisão retém o formato original dos dados de entrada NULL. A transformação avalia funções usando o valor de entrada NULL.

Quando você desabilita tratamento NULL, a transformação de Decisão atribui um valor padrão aos dados de entrada NULL. A transformação avalia funções usando o valor padrão. Por exemplo, se um campo de entrada de tipo de número inteiro tiver um valor NULL, a transformação de Decisão atribuirá um valor de 0 para a entrada e avaliará funções usando um valor de entrada de 0.

Por padrão, o tratamento NULL não está habilitado na transformação de Decisão. Habilite o tratamento NULL na guia **Estratégias**. Você pode habilitar tratamento NULL depois que configurar uma estratégia para a transformação.

## Configurando uma Estratégia de Decisão

Para configurar uma estratégia de decisão, conecte os dados de origem à transformação de Decisão e edite as propriedades nas exibições da transformação.

1. Abra uma transformação de Decisão.
2. Verifique se a transformação contém portas de entrada e saída.
3. Selecione a exibição **Decisão**.
4. Clique em **Adicionar**.
5. Insira um nome para a estratégia.
6. Na área **Expressão**, insira uma instrução condicional `IF-THEN-ELSE`.
7. Para adicionar uma função à expressão, procure as funções na guia **Funções** e clique duas vezes no nome de uma função.  
**Sugestão:** Para inserir rapidamente uma função, digite as primeiras letras do nome de função e selecione `CTRL-Space`.
8. Para adicionar uma porta à expressão, procure as portas na guia **Portas**. Clique duas vezes no nome de uma porta para adicioná-lo à expressão. Opcionalmente, clique em **Editar Portas de Saída** para editar as configurações da porta de saída ou adicionar portas de saída.
9. Opcionalmente, adicione linhas de comentário digitando `"/"` seguido de seus comentários.
10. Clique em **Validar** para determinar se a expressão de decisão é válida.
11. Clique em **OK** para salvar a estratégia.
12. Opcionalmente, adicione outras estratégias. Cada estratégia deve usar portas de saída exclusivas. As estratégias não podem compartilhar portas de saída.

# Propriedades Avançadas da Transformação de Decisão

Configure propriedades que ajudam a determinar como o Serviço de Integração de Dados processa dados para a transformação de Decisão.

Configure as seguintes propriedades avançadas para uma transformação de Decisão:

## Nível de Rastreamento

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.

## Particionável

A transformação pode ser processada com vários segmentos. Desmarque essa opção se quiser que o Serviço de Integração de Dados use um segmento para processar a transformação. O Serviço de Integração de Dados pode usar vários segmentos para processar as fases de pipeline de mapeamento restantes.

Talvez você deseje desativar o particionamento de uma transformação de Decisão quando a transformação usar uma das funções numéricas CUME, MOVINGSUM ou MOVINGAVG. Essas funções calculam totais e médias de execução linha a linha. Uma transformação particionada que usa as funções CUME, MOVINGSUM ou MOVINGAVG pode não retornar o mesmo resultado calculado com cada execução de mapeamento.

Se a transformação não use as funções CUME, MOVINGSUM ou MOVINGAVG, ative o particionamento da transformação para otimizar o desempenho.

# Transformação de Decisão em um ambiente não nativo

O processamento da transformação de Decisão em um ambiente não nativo depende do mecanismo que executa a transformação.

Considere o suporte para os seguintes mecanismos de tempo de execução não nativos:

- Mecanismo Blaze. Suportado sem restrições.
- Mecanismo Spark. Com suporte com restrições em mapeamentos em lote. Sem suporte em mapeamentos de fluxo.
- Mecanismo Databricks Spark. Sem suporte.

## CAPÍTULO 15

# Transformação de Exceção de Registro Duplicado

Este capítulo inclui os seguintes tópicos:

- [Visão Geral da Transformação de Exceção de Registro Duplicado, 265](#)
- [Fluxo do Processo de Exceção de Registro Duplicado, 266](#)
- [Exceções de Registro Duplicado, 266](#)
- [Exibição de Configuração de Exceção de Registro Duplicado, 267](#)
- [Portas, 269](#)
- [Propriedades Avançadas da Transformação de Exceção de Registro Duplicado, 271](#)
- [Exemplo de Mapeamento de Exceção de Registro Duplicado, 272](#)
- [Criando uma Transformação de Exceção de Registro Duplicado, 278](#)

## Visão Geral da Transformação de Exceção de Registro Duplicado

A transformação de Exceção de Registro Duplicado é uma transformação ativa que lê a saída de um processo de qualidade de dados e identifica registros duplicados que exigem revisão manual. A transformação de Exceção de Registro Duplicado é uma transformação de vários grupos.

A transformação de Exceção de Registro Duplicado recebe entrada de outra transformação ou de um objeto de dados em outro mapeamento. A entrada para a transformação de Exceção deve conter uma sequência numérica de pontuação de correspondência que a transformação pode usar para determinar se o registro é uma duplicação. Defina um limite de pontuação de correspondência superior e inferior na transformação de Exceção de Registro Duplicado.

A transformação de Exceção de Registro Duplicado executa uma das ações a seguir:

- Se a pontuação de correspondência for maior ou igual ao limite superior, a transformação tratará o registro como uma duplicação e gravará em um destino de banco de dados.
- Se a pontuação de correspondência for menor do que o limite superior e maior do que o limite inferior, a transformação tratará o registro como uma possível duplicação e gravará em um registro para outro destino para revisão manual. Se o registro pertencer a um cluster, a transformação gravará todos os registros do cluster no destino.

- Quando um cluster tem uma pontuação de correspondência menor que o limite inferior, todos os registros no cluster vão para o grupo de saída de registros exclusivos. Os clusters de tamanho 1 estão roteados para o grupo exclusivo, independentemente da pontuação de correspondência. Por padrão, a transformação de Exceção não grava registros exclusivos em um destino. Você pode configurar a transformação para retornar os registros exclusivos.
- Se alguma pontuação de correspondência em um cluster não estiver no intervalo entre 0 - 100, a transformação de Exceção ignorará todas as linhas no cluster. O Serviço de Integração de Dados registrará uma mensagem que inclui o clusterID.

## Fluxo do Processo de Exceção de Registro Duplicado

A transformação de Exceção de Registro Duplicado analisa a saída de outras transformações de qualidade de dados e cria as tabelas que contêm registros com níveis diferentes de qualidade de dados.

Também pode configurar transformações de qualidade de dados em um único mapeamento, ou pode criar mapeamentos para diferentes fases no processo.

Você pode usar a ferramenta Analyst para analisar e atualizar os registros duplicados que exigem revisão manual.

Use a ferramenta Developer para executar as seguintes tarefas:

1. Crie um mapeamento que gere valores de pontuação para problemas de qualidade de dados.
2. Use uma transformação de Correspondência no modo de cluster para gerar valores de pontuação para exceções de registro duplicado.
3. Configure a transformação de Exceção de Registro Duplicado para ler a saída da transformação de Correspondência. Configure a transformação para gravar registros nas tabelas de banco de dados com base em valores de pontuação de correspondência nos registros.
4. Configure objetos de dados de destino para registros de consolidação automática.
5. Clique na opção **Gerar tabela de registro duplicado** para criar a tabela de registro duplicado e adicioná-la à tela de mapeamento.
6. Adicione o mapeamento para um fluxo de trabalho.
7. Configure uma tarefa Humana para atribuir revisão manual de possíveis registros duplicados aos usuários. Os usuários podem analisar e atualizar os registros na ferramenta Analyst.

## Exceções de Registro Duplicado

Você pode usar uma transformação de Exceção de Registro Duplicado para identificar os clusters de duplicação de dados que precisam de revisão manual. As pontuações de correspondência de registros em clusters determinam as duplicações em potencial. Você pode configurar limites superiores e inferiores para pontuações de correspondência na transformação. Os limites superior e inferior definem o grau de similaridade.

Um cluster contém registros relacionados que uma operação de correspondência agrupa. A transformação de Correspondência cria clusters usando a operação de análise de duplicação e a operação de resolução de identidade. Cada registro em um cluster tem o mesmo ID de cluster. Quando a pontuação de correspondência mais baixa em um cluster estiver entre os limites superior e inferior, a transformação de

Exceção de Registro Duplicado identificará o cluster como um cluster de Exceção de Registro Duplicado. A transformação de Correspondência adiciona uma coluna de valor de ID de cluster a todos os registros. Os registros duplicados recebem o mesmo ID de cluster.

A pontuação mais baixa do registro em um cluster determina o tipo de cluster. Um cluster pode ter 11 registros com uma pontuação de correspondência de 0,95 e um registro com a pontuação de correspondência de 0,79. Se o limite superior for 0,9 e o limite inferior for 0,8, a transformação de Exceção gravará os registros na tabela de registros exclusivos.

## Exibição de Configuração de Exceção de Registro Duplicado

Defina os limites de pontuação de correspondência e configure onde a transformação de Exceção de Registro Duplicado grava os diferentes tipos de dados de saída.

A figura a seguir mostra as propriedades que você pode configurar:

**Limites de Revisão Manual**

Limite Inferior: 0,80

Limite Superior: 0,95

**Opções de Roteamento de Dados**

Tipo	Saída Padrão	Tabela de Registros Duplicados
Consolidação Automática (Acima do limite superior)	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Consolidação Manual (Entre limites)	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Registros Exclusivos (abaixo do limite mínimo)	<input type="checkbox"/>	<input type="checkbox"/>

☐ Criar grupo de saída separado para registros exclusivos

Gerar tabela de registros duplicados

Você pode configurar as seguintes propriedades:

### Limite Inferior

O limite inferior para o intervalo de pontuação de registros duplicados. A transformação processa registros com pontuações de correspondência menor do que esse valor como registros exclusivos. O valor do limite inferior é um número de 0 para 1.

### Limite Superior

O limite superior para o intervalo de pontuação de registros duplicados. A transformação processa registros com pontuações de correspondência maior ou igual ao limite superior como registros duplicados. O valor do limite superior é um número maior que o número do limite inferior.

### Consolidação Automática

Clusters nos quais todos os registros têm pontuações de correspondência maior que o limite superior. Os clusters de consolidação automática não exigem revisão. Os registros são duplicados. Você pode usar a transformação de Consolidação para combinar os registros. Por padrão, a transformação de Exceção de Registro Duplicado grava clusters de consolidação automáticos em portas de saída padrão.

### Consolidação Manual

Clusters em que todos os registros têm pontuações de correspondência maior ou igual ao limite inferior e pelo menos um registro tem uma pontuação de correspondência menor do que o limite superior. Você deve realizar uma revisão manual dos clusters para determinar se eles contêm registros duplicados. Por padrão, a transformação de Exceção de Registro Duplicado grava registros de consolidação manual na tabela de registro duplicado.

### Consolidação Exclusiva

Os clusters com um tamanho igual a um ou clusters nos quais qualquer registro tem uma pontuação de correspondência menor que o limite inferior. Clusters de registro exclusivo não são duplicatas. Por padrão, a transformação de Exceção de Registro Duplicado não grava registros exclusivos para uma tabela de saída.

### Saída Padrão

Os tipos de registro que a transformação grava nas portas de saída padrão.

O padrão é registros de consolidação automática.

### Tabela de Registros Duplicados

Os tipos de registro que a transformação grava nas portas de saída de registros duplicados. O padrão é registros de consolidação manual.

### Criar grupo de saída separado para registros exclusivos

Cria um grupo de saída separado para registros exclusivos. Se você não criar uma tabela separada para os registros exclusivos, poderá configurar a transformação para gravar os registros exclusivos em um dos outros grupos. Ou, você poderá ignorar a gravação de registros exclusivos em uma tabela de saída. Por padrão, essa opção fica desativada.

### Gerar tabela de registros duplicados

Cria um objeto de banco de dados para conter os dados do cluster de registros duplicados. Quando você selecionar essa opção, a ferramenta Developer criará o objeto de banco de dados. A ferramenta Developer adiciona o objeto ao repositório do Modelo, adiciona uma instância do objeto à tela de mapeamento e vincula as portas ao objeto.

## Gerando uma Tabela de Registros Duplicados

Você pode gerar uma tabela de Registros Duplicados em uma instância de transformação de Exceção de Registro Duplicado em um mapeamento.

1. Na exibição de **Configuração**, clique na **tabela Gerar Registros Duplicados** para gerar a tabela.  
A caixa de diálogo **Criar Objeto de Dados Relacionais** é exibida.
2. Procure e selecione uma conexão ao banco de dados para conter a tabela.
3. Insira um nome para a tabela de Registros Duplicados no banco de dados.
4. Insira um nome para o objeto da tabela de Registros Duplicados no repositório do Modelo.
5. Clique em **Concluir**.  
A ferramenta Developer adiciona a nova tabela à tela de mapeamento.



# Portas

A transformação de Exceção de Registro Duplicado tem vários grupos de portas de entrada e saída.

A figura a seguir mostra um exemplo das portas de entrada e saída:

Portas								
	Nome	Tipo	Precisão	Escala	Entrada	Saída	Padrão	Descrição
	Entradas							
	Dados (3)							
1	Employee	decimal	10	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
2	Name	string	10	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
3	Addr1	string	10	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
	Controle (3)							
1	Score	double	15	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
2	Row_Identif...	string	25	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
3	Cluster_ID	integer	10	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
	Saídas							
	Saída Padrão (6)							
1	Score	double	15	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>		
2	Row_Identif...	string	25	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>		
3	Cluster_ID	integer	10	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>		
4	Employee	decimal	10	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>		
5	Name	string	10	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>		
6	Addr1	string	10	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>		
	Dados de Cluster (9)							
1	Row_Identif...	bigint	19	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>		
2	Sequential_...	bigint	19	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>		
3	Cluster_ID	integer	10	0	<input type="checkbox"/>	<input checked="" type="checkbox"/>		

## Portas de Entrada da Transformação de Exceção de Registro Duplicado

Uma transformação de Exceção de Registro Duplicado tem um grupo de Dados e um grupo de Controle, e portas de entrada.

O grupo **Dados** contém portas definidas pelo usuário que recebem os dados de origem.

As portas **Controle** recebem metadados que a transformação de Correspondência adiciona aos dados de origem. A tabela a seguir descreve as portas **Controle**:

Porta	Descrição
Pontuação	Valor decimal entre 0 e 1. Identifica o grau de similaridade com o registro que vinculou o registro ao cluster.
Row_Identifier	Um identificador exclusivo para o registro.
Cluster_ID	A ID do cluster de correspondência ao qual o registro pertence.

## Portas de Saída da Transformação de Exceção de Registro Duplicado

A transformação de Exceção de Registro Duplicado tem vários grupos de saída. Por padrão, a transformação grava registros duplicados no grupo **Saída Padrão**. A transformação grava correspondências possíveis no grupo **Dados de Cluster**. Você pode adicionar um grupo de saída para registros exclusivos.

Você pode alterar os tipos de registro que a transformação grava em portas de saída ao alterar as configurações padrão na exibição **Configuração**.

A tabela a seguir descreve as portas de saída para o grupo **Saída Padrão**:

Porta	Descrição
Pontuação	Valor decimal entre 0 e 1. Identifica o grau de similaridade entre um registro e outro em um cluster.
Row_Identifier	Um identificador exclusivo para o registro.
Cluster_ID	ID do cluster que a transformação de Correspondência atribuiu o registro.
Portas definidas pelo usuário	Os campos de dados de origem.

A tabela a seguir descreve as portas de saída no grupo **Dados de Cluster**:

Porta	Descrição
Row_Identifier	O identificador exclusivo para o registro.
Sequential_Cluster_ID	Identifica o cluster em uma tarefa Humana. Um fluxo de trabalho usa a ID de cluster sequencial para atribuir um cluster a uma instância de uma tarefa humana.
Cluster_ID	Identifica o cluster ao qual o registro pertence. A transformação de Correspondência atribui um ID de cluster para todos os registros.
Pontuação	Valor decimal de 0 para 1. Identifica o grau de similaridade com o registro que vinculou o registro ao cluster.
Is_Master	Valor de string que indica se o registro é o registro preferencial do cluster. Por padrão, a primeira linha no cluster é o registro preferencial. O valor é Y ou N.
Workflow_ID	ID que identifica o fluxo de trabalho do registro em uma tarefa. Quando você executa o mapeamento fora de um fluxo de trabalho, o ID do fluxo de trabalho é DummyWorkflowID.
Portas definidas pelo usuário	As portas de dados de origem.

## Criando portas

Adicione cada porta de entrada ao grupo de Dados. Quando você adiciona a porta de entrada, a ferramenta Developer adiciona uma porta de saída com o mesmo nome ao grupo de Saída Padrão, ao grupo de Dados de Cluster e ao grupo de Registros Exclusivos.

1. Selecione o grupo de entrada de Dados.  
O grupo é destacado.
2. Clique em **Novo (Inserir)**.  
A ferramenta Developer adiciona um campo ao grupo de Dados, ao grupo de Saída Padrão, ao grupo de Dados de Cluster e ao grupo de Registros Exclusivos.
3. Altere o nome do campo conforme necessário.  
A ferramenta Developer altera o nome do campo em outros grupos.
4. Insira as demais portas que você precisa adicionar à fonte de dados.

## Propriedades Avançadas da Transformação de Exceção de Registro Duplicado

A transformação de Exceção de Registro Duplicado contém propriedades avançadas que determinam o comportamento de classificação, o comportamento da memória cache e o nível de rastreamento.

Você pode configurar as seguintes propriedades avançadas:

### Classificação

Determina se a transformação classifica as linhas de entrada nos dados da porta **ID de Cluster**. A propriedade é ativada por padrão.

Selecione a propriedade se as linhas de entrada não estiverem pré-classificadas.

### Diretório de Arquivo de Cache

Especifica o diretório no qual o Serviço de Integração de Dados grava os dados temporários da transformação atual. O Serviço de Integração de Dados grava os arquivos temporários no diretório quando o volume de dados de entrada é maior do que a memória disponível do sistema. O Serviço de Integração de Dados exclui os arquivos temporários depois que o mapeamento é executado.

Você pode inserir um caminho de diretório na propriedade ou usar um parâmetro para identificar o diretório. Especifique um caminho local na máquina do Serviço de Integração de Dados. O Serviço de Integração de Dados deve ser capaz de gravar no diretório. O valor padrão é o parâmetro do sistema CacheDir.

### Tamanho do Arquivo de Cache

Determina a quantidade de memória do sistema que o Serviço de Integração de Dados usa para classificar os dados de entrada na transformação. O valor padrão é 400.000 bytes.

Antes de classificar os dados, o Serviço de Integração de Dados aloca a quantidade de memória especificada por você. Se a operação de classificação gerar uma quantidade maior de dados, o Serviço de Integração de Dados gravará o excesso de dados no diretório do arquivo de cache. Se a operação de classificação exigir mais memória do que a memória do sistema e o armazenamento de arquivos podem fornecer, o mapeamento falhará.

**Nota:** Se você inserir um valor igual ou maior que 65536, a transformação lerá o valor em bytes. Se você inserir um valor menor, a transformação lerá o valor em megabytes.

#### Nível de Rastreamento

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.

## Exemplo de Mapeamento de Exceção de Registro Duplicado

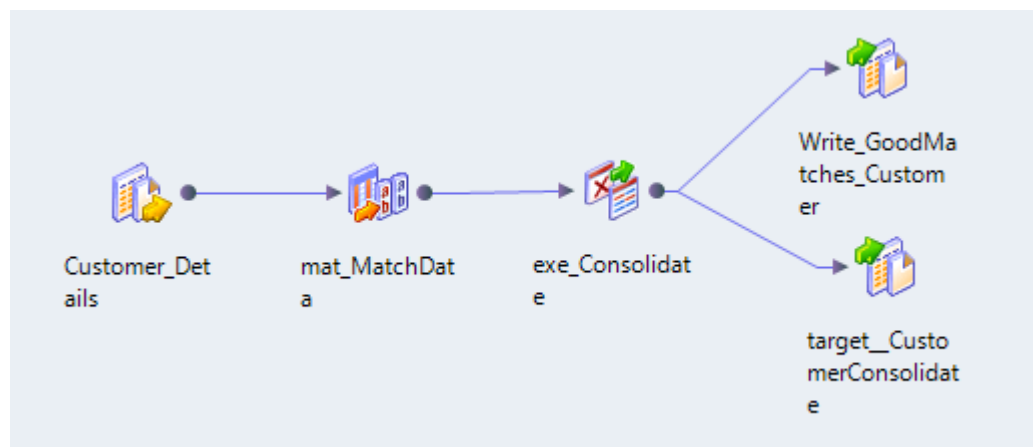
Uma organização executa um projeto de dados para analisar dados de clientes. A organização determina que os dados dos clientes contêm vários registros duplicados. A organização precisa revisar manualmente alguns dos registros que podem ser duplicações.

Crie um mapeamento de qualidade de dados para identificar os registros de clientes duplicados. O mapeamento inclui a transformação de Correspondência. A transformação de Exceção de Registro Duplicado receberá os resultados da transformação de Correspondência. A transformação de Exceção de Registro Duplicado grava cada cluster de registros com um status em dúvida em uma tabela de banco de dados. Um analista de dados analisa os dados na ferramenta Analyst e determina quais registros são duplicados.

### Mapeamento de Exceção de Registro Duplicado

Configure um Mapeamento de Exceção de Registro Duplicado que examina os registros de clientes e localiza registros duplicados.

A figura a seguir mostra o mapeamento de Exceção de Registro Duplicado:



O mapeamento contém os seguintes objetos:

#### Customer\_Details

A fonte de dados que pode conter registros duplicados.

#### mat\_MatchData

Uma transformação de Correspondência que examina os dados de clientes para determinar se os registros são correspondentes. A transformação de Correspondência cria uma pontuação numérica que

representa o grau de similaridade entre dois valores de coluna. Um algoritmo calcula uma pontuação de correspondência como um valor decimal no intervalo de 0 a 1. Um algoritmo atribui uma pontuação de um quando dois valores de coluna são idênticos.

#### exc\_Consolidate

Uma transformação de Exceção de Registro Duplicado que determina quais registros duplicados são possíveis clientes duplicados, clientes duplicados conhecidos ou registros exclusivos.

#### Write\_GoodMatches\_Customer table

A tabela que recebe todos os registros que não precisam de revisão manual. A transformação de Exceção de Registro Duplicado grava registros duplicados e registros exclusivos nessa tabela.

#### A tabela Target\_CustomerConsolidate

A transformação de Exceção grava as possíveis registros duplicados na tabela Target\_CustomerConsolidate. Os registros nessa tabela exigem revisão manual na ferramenta Analyst.

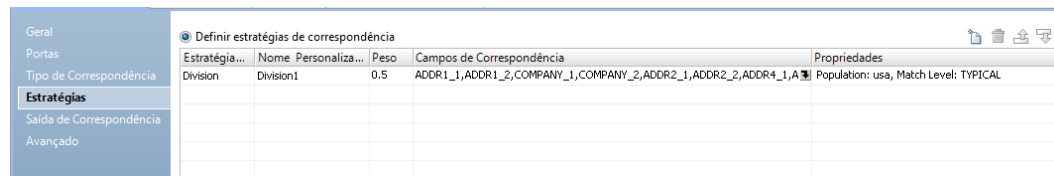
## Transformação de Correspondência

A transformação de Correspondência recebe os dados de clientes e realiza uma correspondência de identidade.

Configure a transformação de Correspondência para o tipo de saída Clusters-Correspondência Todos. A transformação de Correspondência retorna a correspondência de registros em clusters. Cada registro em um cluster deve corresponder a pelo menos um outro registro no cluster com uma pontuação maior ou igual ao limite de correspondência. O limite de correspondência é 0,75.

Selecione a estratégia de correspondência de Divisão na guia **Estratégias** da transformação de Correspondência. A estratégia de Divisão é uma estratégia de correspondência predefinida que identifica uma organização com base nos campos de endereço. Na guia **Estratégias** da transformação de Correspondência, escolha as portas de entrada para examinar em uma correspondência. Configure a estratégia de peso como 0,5.

A figura a seguir mostra a configuração da estratégia de Divisão para a transformação de Correspondência:



A transformação de Correspondência adiciona as informações de cluster para cada registro de saída. A transformação também adiciona um RowID exclusivo para cada registro.

## Grupos de Entrada de Exceção de Registro Duplicado

A transformação de Exceção de Registro Duplicado tem dois grupos de entrada. A transformação tem um grupo de Dados que recebe os dados de clientes. A transformação também tem o grupo de Controle que contém a pontuação de correspondência para a linha, o identificador de linha e o ID de cluster.

A figura a seguir mostra os grupos de entrada na transformação de Exceção:

	Nome	Tipo	Precisão	Escala	Entrada	Saída	Padrão	Descrição
	Entradas							
	Dados (11)							
1	CUST_ID	decimal	20	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
2	COMPANY	string	200	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
3	CONTACT	string	200	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
4	TITLE	string	200	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
5	ADDR1	string	200	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
6	ADDR2	string	100	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
7	ADDR3	string	100	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
8	ADDR4	string	50	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
9	COUNTRY	string	50	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
10	PHONE	string	100	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
11	EMAIL	string	100	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
	Controle (3)							
1	Score	double	15	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
2	Row_Identif...	string	25	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
3	Cluster_ID	integer	10	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		

O grupo de Dados contém os dados de clientes. Os dados de clientes incluem um ID de cliente, o contato, o título e os campos de endereço. O grupo de Controle tem os metadados adicionais que a transformação de Correspondência adicionou a cada registro de cliente. O grupo de Controle contém a pontuação de correspondência, o rowID e o ID de cluster.

## Exemplo de Exibição de Configuração de Exceção de Registro Duplicado

Defina os limites superior e inferior na exibição **Configuração**. Identifique onde a transformação grava os registros duplicados, os possíveis registros duplicados e os registros exclusivos do exclusivos.

A figura a seguir mostra a exibição de **Configuração** da a transformação de Exceção de Registro Duplicado:

**Limites de Revisão Manual**

Limite Inferior: 0,80

Limite Superior: 0,90

**Opções de Roteamento de Dados**

Tipo	Saída Padrão	Tabela de Registros Duplicados
Consolidação Automática...	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Consolidação Manual (En...	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Registros Exclusivos (abai...	<input checked="" type="checkbox"/>	<input type="checkbox"/>

☐ Criar grupo de saída separado para registros exclusivos

**Gerar tabela de registros duplicados**

A tabela a seguir descreve as definições de configuração:

Opção	Definição
Limite inferior	.80
Limite superior	.95
Consolidação Automática	Tabela de Saída Padrão
Consolidação Manual	Tabela de Registros Duplicados
Registros Exclusivos	Tabela de Saída Padrão

Clique em **Gerar tabela de registro duplicado** para criar a tabela de registro duplicado. Não crie uma tabela separada para os registros exclusivos. A transformação grava os registros exclusivos na tabela de Saída Padrão.

## Tabela de Registros de Saída Padrão

A tabela de destino Write\_GoodMatches\_Customer recebe linhas do grupo de Saída Padrão. A tabela recebe registros exclusivos e recebe registros que são duplicados. Esses registros não exigem revisão manual.

A figura a seguir mostra os registros de Saída Padrão que a transformação de Exceção retorna:

Saida										
Nome:	<a href="#">exc_Consoli.Good_Records</a>									
	Score	Row_Identifier	Cluster_ID	CUST_ID	COMPANY	CONTACT	TITLE	ADDR1	ADDR2	ADDR3
1	1	1 - 4	3	1001622	INVESTEX	NEW YORK	United Parcel...	"50 BROAD S...	20TH FL"	310 BRIER RO...
2	1	1 - 5	4	7121564	"ARTHUR VA...	INC."	WHITEHOUSE...	Federal Expre...	310 BRIER RO...	767 FIFTH AV...
3	0,9047619047...	1 - 6	4	7121565	OSTERREICHL...	NEW YORK	United Parcel...	767 FIFTH AV...	570 LEXINGT...	6803 S. TUCS...
4	0,9047619047...	1 - 7	4	7121566	K2 ADVISORS	NEW YORK	Courier	570 LEXINGT...	51 JFK PARK...	6803 S. TUCS...
5	0,9047619047...	1 - 8	4	7121567	MFP INVESTO...	SHORT HILLS	Federal Expre...	51 JFK PARK...	ONE PARKER...	6803 S. TUCS...
6	0,9047619047...	1 - 9	4	7121568	EASTON	FORT LEE	US Postal Inte...	ONE PARKER...	3840 HIGHW...	6803 S. TUCS...
7	0,9047619047...	1 - 10	4	7121569	RICE VOELKER	MANDEVILLE	Federal Expre...	3840 HIGHW...	"2ND & D ST...	SW HOUSE...
8	1	1 - 1	16	1001590	E-AGENCY	OAKLAND	Federal Expre...	291 3RD STRE...	2500 WESTER...	6803 S. TUCS...
9	0,9047619047...	1 - 2	16	1001599	BANK ONE	ELGIN	US Postal Inte...	2500 WESTER...	530 CHESNUT...	6803 S. TUCS...
10	1	1 - 3	16	1001604	KPMG PEAT...	WOODCLIFF L...	US Postal 2-day	530 CHESNUT...	"50 BROAD S...	20TH FL"
11	1	1 - 11	16	1001658	HOUSE INFO...	WASHINGTON	Federal Expre...	"2ND & D ST...	SW HOUSE...	<null>
12	0,9047619047...	1 - 12	16	1001660	OPPENHEIME...	ENGLEWOOD	US Postal Inte...	6803 S. TUCS...	6803 S. TUCS...	6803 S. TUCS...
13	0,9047619047...	1 - 13	16	1001659	THOMSON LE...	FT. WORTH	Courier	301 COMMER...	301 COMMER...	6803 S. TUCS...
14	1	1 - 14	16	1001658	LYNCH JONE...	NEW YORK	US Postal Ove...	875 3RD AVE...	875 3RD AVE...	6803 S. TUCS...
15	0,9047619047...	1 - 15	16	1001691	EPICENTRIC	SAN FRANCIS...	US Postal 2-day	333 BRYANT...	333 BRYANT...	6803 S. TUCS...
16	0,9047619047...	1 - 16	16	1001664	CHARLES SCH...	SAN ANTONIO	United Parcel...	1100 N.E. LO...	1100 N.E. LO...	6803 S. TUCS...
17	0,9047619047...	1 - 17	16	1001694	BANK OF MO...	BOSTON	US Postal 2-day	"125 BROAD...	38TH FLOOR"	"125 BROAD...
18	1	1 - 18	16	1001729	JOSEPH H. GL...	HIGHLAND P...	United Parcel...	1212 LINCOL...	1212 LINCOL...	6803 S. TUCS...
19	0,9047619047...	1 - 19	16	1001724	D E SHAW &...	NEW YORK	US Postal Ove...	120 WEST 45...	120 WEST 45...	6803 S. TUCS...
20	0,9047619047...	1 - 20	16	1001732	BANG NETW...	SAN FRANCIS...	Courier	808 BRANNA...	808 BRANNA...	6803 S. TUCS...
21	0,9047619047...	1 - 21	16	1001736	FOREX CAPIT...	NEW YORK	Federal Expre...	11 BROADWA...	11 BROADWA...	6803 S. TUCS...
22	0,9047619047...	1 - 22	16	1001738	"INTERFACE S...	INC."	ANN ARBOR	United Parcel...	5855 INTERFA...	5855 INTERFA...
23	0,9047619047...	1 - 23	16	1001758	BANK OF NE...	LAKE MARY	United Parcel...	100 COLONIA...	100 COLONIA...	6803 S. TUCS...

O registro contém os seguintes campos:

#### Pontuação

Pontuação de correspondência que indica o grau de semelhança entre um registro e outro registro no cluster. Os registros com uma pontuação de correspondência de um são registros duplicados que não precisam de revisão. Um cluster no qual qualquer registro tem uma pontuação de correspondência abaixo do limite inferior não é uma duplicata de cluster.

#### Row\_Identifier

Um número de linha que identifica exclusivamente cada linha na tabela. Por exemplo, o identificador de linha contém o ID de cliente.

#### ID de Cluster

Um identificador exclusivo para um cluster. Cada registro em um cluster recebe o mesmo ID de cluster. Os quatro primeiros registros do exemplo de dados de saída são exclusivos. Cada registro tem uma ID de cluster exclusiva. As linhas de cinco a nove pertencem ao cluster cinco. Cada registro no cluster é um registro duplicado devido às similaridades dos campos de endereço.

#### Campos de Dados de Origem

O grupo de tabela Saída Padrão também recebe todos os campos de dados de origem.

## Saída de Cluster

A tabela Target\_CustomerConsolidate recebe registros do grupo de Saída de Cluster. O grupo de Saída de Cluster retorna os registros que podem ser registros duplicados. Os registros na tabela Target\_CustomerConsolidate precisam de revisão manual na ferramenta Analyst.

A seguinte imagem mostra alguns dos registros e campos da tabela Target\_CustomerConsolidate:



Nome: [exc\\_ConsoLiCluster\\_Data](#)

	Row_Identifier	Sequential_Clust...	Cluster_ID	Score	Is_Master	Workflow_ID	CUST_ID	COMPANY	CONTACT	TITLE	ADDR1	ADDR2
1	0	0	3	1	Y	DummyWorkf...	1001622	INVESTEX	NEW YORK	United Parcel...	"50 BROAD S...	20TH FL"
2	1	0	3	1	N	DummyWorkf...	1001622	INVESTEX	NEW YORK	United Parcel...	"50 BROAD S...	20TH FL"
3	2	1	4	1	Y	DummyWorkf...	7121564	"ARTHUR VA...	INC."	WHITEHOUSE...	Federal Expre...	310 BRIER RC
4	3	1	4	1	N	DummyWorkf...	7121564	"ARTHUR VA...	INC."	WHITEHOUSE...	Federal Expre...	310 BRIER RC
5	4	1	4	0,904761...	N	DummyWorkf...	7121565	OSTERREICHL...	NEW YORK	United Parcel...	767 FIFTH AV...	570 LEXINGT
6	5	1	4	0,904761...	N	DummyWorkf...	7121566	K2 ADVISORS	NEW YORK	Courier	570 LEXINGT...	51 JFK PARK
7	6	1	4	0,904761...	N	DummyWorkf...	7121567	MFP INVESTO...	SHORT HILLS	Federal Expre...	51 JFK PARK...	ONE PARKER
8	7	1	4	0,904761...	N	DummyWorkf...	7121568	EASTON	FORT LEE	US Postal Inte...	ONE PARKER...	3840 HIGHW
9	8	1	4	0,904761...	N	DummyWorkf...	7121569	RICE VOELKER	MANDEVILLE	Federal Expre...	3840 HIGHW...	"2ND & D ST
10	9	2	16	1	Y	DummyWorkf...	1001590	E-AGENCY	OAKLAND	Federal Expre...	291 3RD STRE...	2500 WESTE
11	10	2	16	1	N	DummyWorkf...	1001590	E-AGENCY	OAKLAND	Federal Expre...	291 3RD STRE...	2500 WESTE
12	11	2	16	0,904761...	N	DummyWorkf...	1001599	BANK ONE	ELGIN	US Postal Inte...	2500 WESTER...	530 CHESNU
13	12	2	16	1	N	DummyWorkf...	1001604	KPMG PEAT...	WOODCLIFF...	US Postal 2-day	530 CHESNUT...	"50 BROAD S
14	13	2	16	1	N	DummyWorkf...	1001658	HOUSE INFO...	WASHINGTON	Federal Expre...	"2ND & D ST...	SW HOUSE.
15	14	2	16	0,904761...	N	DummyWorkf...	1001660	OPPENHEIME...	ENGLEWOOD	US Postal Inte...	6803 S. TUCS...	6803 S. TUCS
16	15	2	16	0,904761...	N	DummyWorkf...	1001659	THOMSON LE...	FT. WORTH	Courier	301 COMMER...	301 COMME
17	16	2	16	1	N	DummyWorkf...	1001658	LYNCH JONE...	NEW YORK	US Postal Ove...	875 3RD AVE.	875 3RD AVE
18	17	2	16	0,904761...	N	DummyWorkf...	1001691	EPICENTRIC	SAN FRANCI...	US Postal 2-day	333 BRYANT...	333 BRYANT.
19	18	2	16	0,904761...	N	DummyWorkf...	1001664	CHARLES SCH...	SAN ANTONIO	United Parcel...	1100 N.E. LO...	1100 N.E. LO
20	19	2	16	0,904761...	N	DummyWorkf...	1001694	BANK OF MO...	BOSTON	US Postal 2-day	"125 BROAD...	38TH FLOOR

O registro contém os seguintes campos:

#### Row\_Identifier

Um número que identifica exclusivamente cada linha na tabela.

#### ID de Cluster Sequencial

Um identificador sequencial para cada cluster para analisar em uma tarefa Humana. A transformação de Exceção de Registro Duplicado adiciona o ID de cluster sequencial aos registros do grupo de saída de Dados do Cluster.

#### ID de Cluster

Um identificador exclusivo para um cluster. A transformação de Correspondência atribui um ID de cluster a todos os registros de saída. Os registros duplicados e os possíveis registros duplicados compartilham o mesmo ID de cluster. Um registro exclusivo recebe um ID de cluster, mas o registro não compartilha o número de ID com qualquer outro registro.

#### Pontuação

Pontuação de correspondência que indica o grau de semelhança entre um registro e outro registro no cluster. Os registros que exigem revisão manual têm pontuações de correspondência menor que 0,95 e maior que 0.80.

#### É Mestre

Indica se o registro é o registro preferencial do cluster.

#### WorkflowID

O WorkflowID é um DummyWorkflowID porque a transformação não está em um fluxo de trabalho.

#### Campos de Registro

Os outros campos no registro contém os dados de origem do cliente.

# Criando uma Transformação de Exceção de Registro Duplicado

Quando você configurar uma transformação de Exceção de Registro Duplicado, configure as portas de entrada. Defina os limites superior e inferior para determinar correspondências. Configure onde gravar os registros duplicados e os registros exclusivos.

1. Crie uma transformação de Exceção de Registro Duplicado reutilizável ou não reutilizável.
  - Para criar uma transformação reutilizável, selecione **Arquivo > Novo > Transformação** e selecione uma transformação de Exceção de Registro Duplicado.
  - Para criar uma transformação não reutilizável, abra um mapeamento e adicione essa transformação à tela de mapeamento. Selecione uma transformação de Exceção de Registro Duplicado no assistente.
2. Clique em **Avançar** ou em **Concluir**.

Se você clicar em **Concluir**, poderá atualizar os valores de limite padrão e as opções de roteamento de dados antes de criar a transformação.
3. Na exibição Configuração, configure os limites de pontuação de correspondência superior e inferior.
4. Na seção **Opções de Roteamento de Dados**, configure a saída padrão e as propriedades da tabela de exceção para definir onde a transformação grava cada tipo de registro.

Opcionalmente, modifique onde gravar os registros duplicados, os registros duplicados em revisão e os registros exclusivos.
5. Opcionalmente, gere uma tabela de registros exclusivos. Insira a conexão de banco de dados e as informações de nome de tabela na nova tabela. Quando você gerar uma tabela de registros exclusivos, a transformação criará um objeto de banco de dados no repositório do Modelo.
6. Configure as portas de entrada. Quando você adiciona uma porta de entrada, a Developer tool adiciona o mesmo nome da porta aos grupos de saída.
  - Se você criar uma transformação reutilizável, selecione a guia **Portas** e adicione portas aos dados que você deseja conectar à transformação.
  - Se você criar uma transformação não reutilizável, adicione outros objetos à tela de mapeamento e arraste as portas de entrada para a transformação.
7. Conecte as portas de saída da transformação a um ou mais destinos de dados. Conecte as portas de saída aos objetos de dados que correspondem às opções de saída definidas na exibição **Configuração**.
  - Se criar uma transformação reutilizável, adicione a transformação a um mapeamento e conecte as portas de saída.
  - Se você criar uma transformação não reutilizável, a transformação conectará as portas à tabela de Dados de Cluster. Você deve conectar portas de saída a outros destinos de dados.

## CAPÍTULO 16

# Transformação de Expressão

Este capítulo inclui os seguintes tópicos:

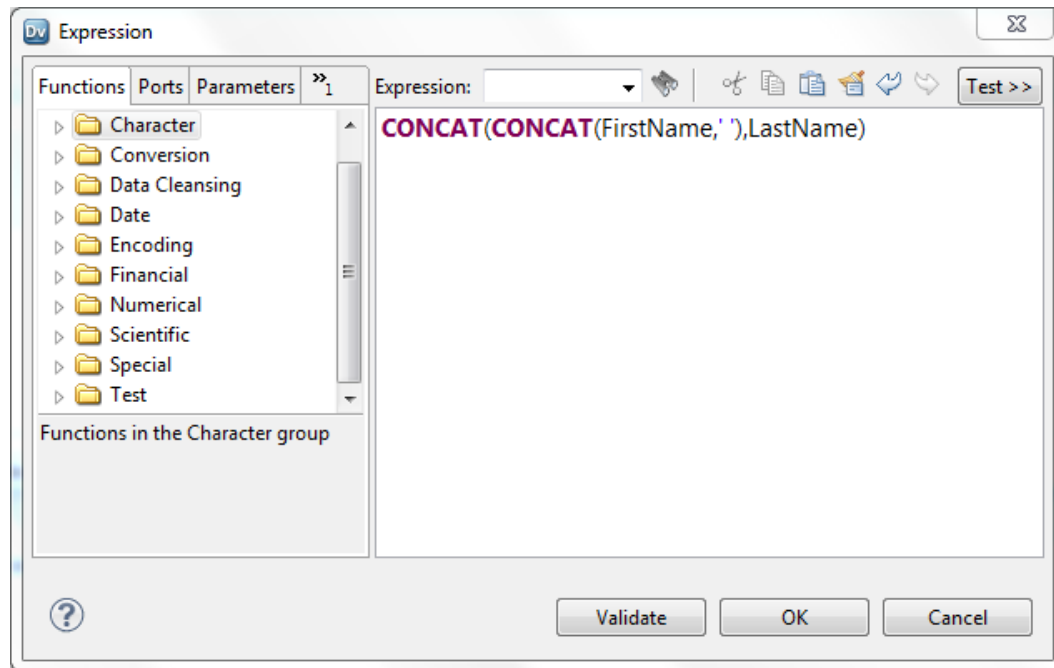
- [Visão Geral da Transformação de Expressão, 279](#)
- [Portas de Transformação de Expressão, 280](#)
- [Testar Expressões, 281](#)
- [Seletores de Portas, 283](#)
- [Definição de Janelas, 286](#)
- [Expressões Dinâmicas, 290](#)
- [Transformação de Expressão - Propriedades Avançadas, 295](#)
- [Transformação de Expressão em um ambiente não nativo, 295](#)

## Visão Geral da Transformação de Expressão

A transformação de Expressão é uma transformação passiva que você pode usar para realizar cálculos ou testar instruções condicionais em uma linha. Em transformação de Expressão não reutilizáveis, é possível definir uma expressão de saída de mapeamento a ser agregada quando você define saídas de mapeamento.

Em uma única linha, talvez seja necessário criar uma expressão para ajustar salários de funcionários, concatenar nomes e sobrenomes ou converter cadeias em números.

A seguinte figura mostra uma expressão em uma transformação de Expressão que concatena o nome, um espaço e o sobrenome:



É possível inserir várias expressões em uma transformação de Expressão, criando uma expressão para cada porta de saída. Por exemplo, talvez você queira calcular diferentes tipos de impostos do salário de cada funcionário, como os impostos de renda local e federal. Ambos os cálculos de impostos exigem o salário do funcionário e uma alíquota de imposto. Defina uma porta de saída separada para cada cálculo. Defina uma expressão diferente para cada porta de saída. Você pode definir portas de passagem para o salário e a alíquota de imposto, pois os valores das portas não mudam.

## Portas de Transformação de Expressão

Uma transformação de Expressão tem diferentes tipos de portas que você pode referenciar ao definir expressões.

Uma transformação de Expressão tem os seguintes tipos de porta:

### Entrada

Recebe dados de transformações upstream. Se a transformação de Expressão não alterar o valor da porta, você poderá definir uma porta de passagem em vez de uma porta de entrada.

### Saída

Contém o valor de retorno da expressão. Você insere a expressão como uma opção de configuração para a porta de saída. Também é possível configurar um valor padrão para cada porta.

**Nota:** Se uma expressão resultar em erros numéricos, como divisão por zero ou SQRT de um número negativo, ela retornará um valor infinito ou NaN.

**Passagem**

Defina uma porta de passagem para transmitir dados através da transformação sem alterar o valor. Você pode fazer referência a uma porta de passagem em um cálculo, mas não pode alterar o valor dos dados na porta de passagem.

**Variável**

Armazena temporariamente os dados a serem usados em expressões. É possível armazenar dados através de várias linhas. Uma expressão pode ser definida para retornar um valor para uma porta variável.

**Porta Dinâmica**

Recebe ou retorna portas em um mapeamento dinâmico. Uma porta dinâmica pode receber uma ou mais colunas de uma transformação em upstream e criar uma porta gerada para cada coluna. Uma porta de saída dinâmica pode retornar uma ou mais portas geradas. Você pode definir regras de entrada para determinar quais colunas são recebidas por uma porta dinâmica. Uma porta de saída dinâmica pode conter uma expressão que gera várias portas de saída.

**Porta Gerada**

Uma porta que representa uma única coluna em uma porta dinâmica. As portas geradas na transformação de Expressão podem mudar de acordo com as colunas que essa transformação recebe de uma transformação em upstream.

## Testar Expressões

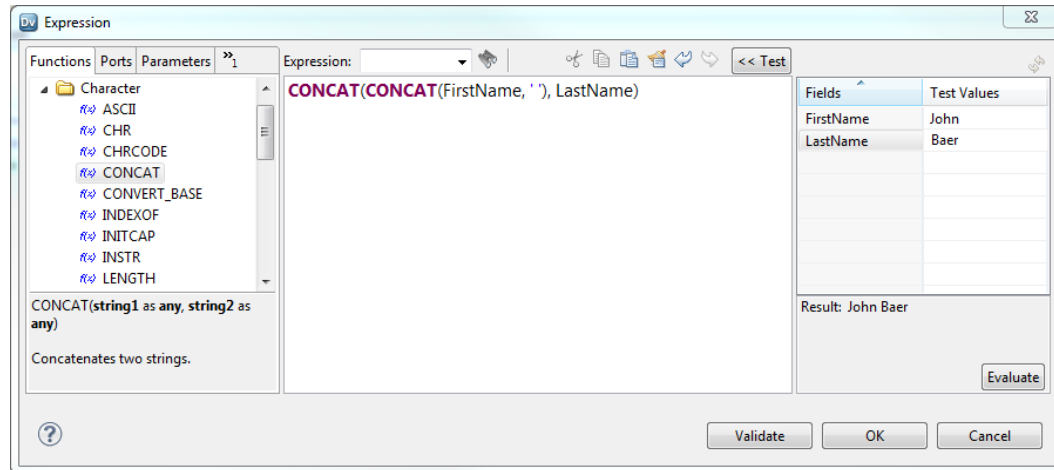
Você pode testar expressões configuradas no Editor de Expressões. Ao testar uma expressão, você insere dados de amostra e, em seguida, avalia a expressão.

Você pode testar expressões ao configurar expressões das seguintes maneiras:

- Em uma porta de saída ou variável na transformação de Expressão
- Na exibição Saídas de Mapeamento de uma transformação de Expressão, após a adição da transformação a um mapeamento

Por exemplo, depois de configurar uma expressão que concatena o nome, um espaço e o sobrenome, você pode inserir dados de amostra para as portas e, em seguida, avaliar a expressão para verificar o resultado.

A seguinte imagem mostra os resultados de uma expressão que concatena uma amostra de nome e sobrenome:



## Cadeias de Formato de Data para Dados de Amostra

Ao testar uma expressão que utiliza uma porta com o tipo de dados Data/Hora ou Registro de Data/Hora com Fuso Horário, você deve inserir os dados de amostra da porta usando a cadeia de formato de data necessária.

Para inserir dados de amostra para uma porta com o tipo de dados Data/Hora, use o formato MM/DD/YYYY HH24:MI:SS. Quando a expressão for avaliada, o Editor de Expressão exibirá o resultado utilizando o formato especificado na expressão. Se você omitir a cadeia de formato na expressão, o Editor de Expressão exibirá o resultado usando o mesmo formato MM/DD/YYYY HH24:MI:SS.

Para inserir dados de amostra para uma porta com o tipo de dados Registro de Data/Hora com Fuso Horário, use o formato MM/DD/YYYY HH24:MI:SS TZR. Quando a expressão for avaliada, o Editor de Expressão exibirá o resultado utilizando o formato YYYY-MM-DD HH24:MI:SS.NS TZR.

## Testando uma Expressão

Teste uma expressão no Editor de Expressão para avaliar essa expressão e verificar seu resultado.

- Abra o Editor de Expressão de uma das seguintes maneiras:
  - Em uma transformação de Expressão, clique o botão **Abrir** (🔗) na coluna **Expressão** de uma porta de saída ou de uma porta variável.
  - Selecione uma transformação de Expressão incluída em um mapeamento. Na exibição **Saídas de Mapeamento**, clique no botão **Abrir** (🔗) na coluna **Expressão** de uma saída.
- Configure a expressão.
- Clique em **Testar >>** para abrir o painel de teste.
- Insira dados de amostra para cada campo na coluna **Valores de Teste**.  
É possível inserir valores de teste para cada porta ou parâmetro incluído na expressão.
- Clique em **Avaliar**.  
O resultado da expressão é exibido na parte inferior do painel de teste.

# Seletores de Portas

Quando uma transformação gerou portas, você precisa configurá-la para ser executada com êxito quando essas portas geradas mudarem. É possível usar um seletor de portas para determinar quais portas devem ser usadas em uma expressão dinâmica, uma condição de pesquisa ou uma condição de associador.

Um seletor de portas é uma lista ordenada de portas que você pode referenciar em uma expressão. Quando as portas geradas mudam em um mapeamento dinâmico, o seletor de portas pode conter diferentes portas.

Por exemplo, a seguinte expressão faz referência a uma porta gerada em um mapeamento dinâmico:

```
Salary * 12
```

Configure o mapeamento para usar as origens dinâmicas, mas a coluna que contém informações salariais em cada arquivo de origem tem um nome diferente. Os nomes de coluna são `Salary`, `Monthly_Salary` e `Base_Salary`.

Realize as seguintes tarefas para acomodar os nomes de coluna diferentes:

1. Crie um seletor de portas denominado "Salary\_PortSelector".
2. Crie uma regra de seleção para aceitar qualquer nome de porta com o sufixo "Salary".
3. Configure a expressão para incluir o nome do seletor de portas em vez do nome da coluna Salary. A expressão tem a seguinte sintaxe:

```
Salary_PortSelector * 12
```

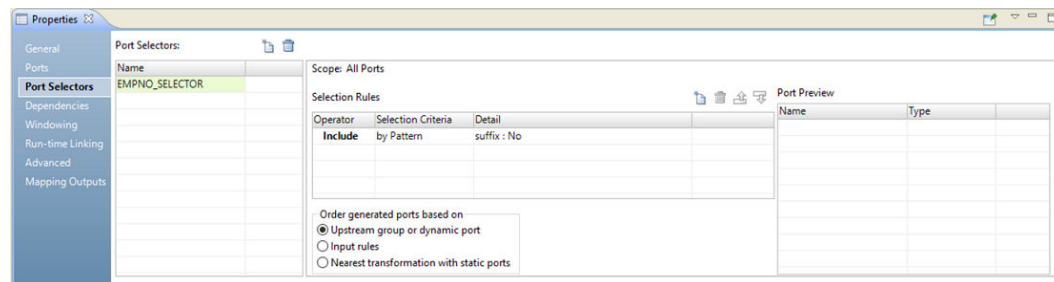
A expressão é executada com êxito com qualquer um dos nomes de portas de salário.

## Configuração do seletor de portas

Ao configurar um seletor de portas, você define regras de seleção para determinar quais portas geradas devem ser incluídas. As regras de seleção são semelhantes às regras de entrada que você pode configurar para portas dinâmicas.

Um seletor de portas pode incluir portas estáticas ou portas geradas. Configure um seletor de portas na guia **Seletor de Portas**.

A seguinte imagem mostra a guia **Seletor de Portas**:



Configure as seguintes propriedades para um seletor de portas:

### Nome

Identifica o seletor de portas. Você pode criar vários seletores de portas em uma transformação e fazer referência a eles em expressões.

### Escopo

Identifica um grupo de portas ao qual o seletor de portas se aplica. Você deve escolher o escopo ao criar um seletor de portas para uma transformação de Associador ou Pesquisa. Essas transformações

têm vários grupos de entrada. A transformação de Associador tem um escopo Mestre ou Detalhado. A transformação de Pesquisa tem um escopo de Importação ou Pesquisa. A transformação de Expressão tem um grupo de entrada. O escopo é sempre Todas as Portas.

#### Regras de Seleção

Determina as portas a serem incluídas no seletor de portas. Quando você cria as regras de seleção, o painel **Visualização de Portas** mostra as portas qualificadas a partir das portas de entrada atuais. Essas portas podem mudar. Configure as regras de seleção para acomodar portas de diferentes origens.

## Regras de Seleção

As regras de seleção associadas a um seletor de portas determinam as portas a serem incluídas no seletor de portas.

Quando você cria as regras de seleção, o painel **Visualização de Portas** mostra as portas qualificadas a partir das portas de entrada atuais. Essas portas podem mudar. Configure as regras de seleção para acomodar portas de diferentes origens.

Crie regras de seleção com base nos critérios a seguir:

#### Operador

Inclui ou exclui as portas que são retornadas pelas portas de seleção. O padrão é incluir. Você deve incluir portas antes de poder excluí-las.

#### CrITÉRIOS de Seleção

O tipo de regra de seleção que você deseja criar. Você pode criar uma regra com base no nome da coluna, tipo de porta, padrão ou definição de tipo de dados complex. Para incluir portas com base no nome da coluna, procure nomes específicos ou procure um padrão de caracteres no nome.

#### Detalhe

Os valores a serem aplicados aos critérios de seleção. Se os critérios de seleção forem por nome de coluna, configure a cadeia ou o nome a ser pesquisado. Se os critérios de seleção forem por tipo de porta, selecione os tipos de porta a serem incluídos.

A seguinte tabela descreve os critérios de seleção e como especificar os detalhes dos critérios:

CrITÉRIOS de Seleção	Descrição	Detalhe
Tudo	Inclui todas as portas.	Nenhum detalhe é necessário.
Nome	Filtra portas com base no nome da porta.	Selecione os nomes de portas de uma lista de valores ou use um parâmetro do tipo Porta ou Lista de Portas.
Tipo	Filtra portas com base no tipo de dados de cada uma.	Selecione tipos de dados de uma lista.
Padrão	Filtra portas por uma cadeia de caracteres no nome ou por uma expressão regular.	Escolha o prefixo, o sufixo ou a expressão regular como o tipo de padrão para o nome da porta. Em seguida, insira um valor para o padrão ou use um parâmetro do tipo Cadeia.
Definição de Tipo de Dados Complex	Filtra portas por uma definição de tipo de dados complex.	Escolha prefixo, sufixo ou expressão regular como o tipo de padrão para a definição de tipo de dados complex. Em seguida, insira um valor para o padrão ou use um parâmetro do tipo Cadeia.

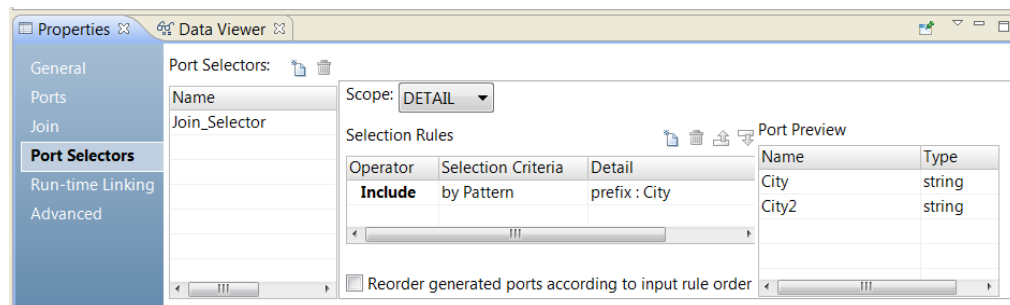


## Criando um Seletor de Portas

Crie um seletor de portas para determinar quais portas devem ser usadas em uma expressão dinâmica, uma condição de pesquisa ou uma condição de associador.

1. Clique na guia **Seletores de Portas**.
2. Na área **Seletores de Portas**, clique em **Novo**.  
A Developer tool cria um seletor de portas com uma regra de seleção padrão que inclui todas as portas.
3. Na área **Seletores de Portas**, mude o nome do seletor de portas para um nome exclusivo.
4. Se estiver trabalhando na transformação de Associador ou na transformação de Pesquisa, escolha o escopo.  
As portas disponíveis mudam de acordo com o grupo de portas que você escolhe.
5. Na área **Regras de Seleção**, selecione um **Operador**.
  - Incluir. Crie uma regra que inclua portas para o seletor de portas. Você deve incluir portas antes de poder excluí-las.
  - Excluir. Crie uma regra que exclua portas específicas do seletor de portas.
6. Escolha os **Critérios de Seleção**.
  - Por Nome. Selecione portas específicas por nome. Você pode selecionar os nomes das portas em uma lista de portas no escopo.
  - Por Tipo. Selecione portas por tipo. Você pode selecionar um ou mais tipos de dados.
  - Por Padrão. Selecione portas por um padrão de caracteres no nome da porta. É possível pesquisar com caracteres específicos ou criar uma expressão regular.

A seguinte imagem mostra a guia Seletor de Portas:



7. Clique na coluna **Detalhes**.  
A caixa de diálogo **Detalhes da Regra de Entrada** é exibida.
8. Selecione os valores pelos quais filtrar as portas.
  - Por Nome. Opte por criar uma lista de portas por valor ou de acordo com um parâmetro. Clique em **Escolher** para selecionar as portas na lista.
  - Por Tipo. Selecione um ou mais tipos de dados de uma lista. A área **Visualização de Portas** mostra portas dos tipos que você seleciona.
  - Por Padrão. Opte por pesquisar o prefixo ou sufixo do nome da porta em busca de um padrão específico de caracteres. Ou, opte por criar uma expressão regular com a qual pesquisar. Configure um parâmetro ou configure o padrão com base no qual pesquisar.A área **Visualização de Portas** mostra as portas no seletor de portas à medida que você configura as regras.
9. Para reordenar as portas no seletor de portas, selecione **Reordenar portas geradas de acordo com a ordem das regras de entrada**.

# Definição de Janelas

Quando uma transformação contém uma função de janela, você precisa configurar as propriedades de definição de janelas. A definição de janelas está disponível apenas para transformações no mecanismo Spark.

Funções de janela operam em um grupo de linhas e calculam um valor de retorno para cada linha de entrada.

Antes de definir uma função de janela em uma transformação de Expressão, você precisa descrever a janela, configurando as propriedades de definição de janelas. Essas propriedades de definição de janelas incluem uma especificação de quadro, as chaves de partição e as chaves de ordem. A especificação de quadro indica quais linhas estão incluídas no cálculo geral da linha atual. As chaves de partição determinam quais linhas estão na mesma partição. As chaves de ordem determinam como as linhas em uma partição são ordenadas.

Depois de configurar propriedades de definição de janelas, você define uma função de janela na transformação de Expressão. A Informatica oferece suporte às funções de janela LEAD e LAG. Você também pode usar funções agregadas como funções de janela em uma transformação de Expressão.

## Configuração da definição de janelas

Ao incluir uma função de janela em uma transformação de Expressão, você configura as propriedades de definição de janelas associadas a essa função. Propriedades de definição de janelas definem os limites de particionamento, ordenação e quadro associados a uma linha de entrada específica.

Configure uma transformação para definição de janelas na guia Definição de Janelas.

A seguinte imagem mostra a guia Definição de Janelas:

The screenshot shows the 'Properties' dialog box with the 'Windowing' tab selected. The 'Frame' section includes 'Start Offset' (0), 'End Offset' (0), and checkboxes for 'All Rows Preceding' and 'All Rows Following'. The 'Order Keys' section has a 'Specify by' dropdown set to 'Value' and a list of ports with '\_event\_id' selected and 'Ascending (A)' as the order. The 'Partition Keys' section also has a 'Specify by' dropdown set to 'Value' and a list of ports with 'trip\_id' selected. Buttons for 'Add', 'Choose...', 'Delete', 'Move Up', and 'Move Down' are present for both key lists.

Você configura os seguintes grupos de propriedades na guia Definição de Janelas:

### Quadro

Define as linhas que estão incluídas no quadro para a linha de entrada atual, com base em deslocamentos físicos a partir da posição da linha de entrada atual.

Configure um quadro se você usar uma função de agregação como uma função de janela. As funções de janela LEAD e LAG fazem referência a linhas individuais e ignoram a especificação de quadro.

### Chaves de Partição

Separe as linhas de entrada em diferentes partições. Se você não definir chaves de partição, todas as linhas pertencerão a uma única partição.

### Chaves de Ordem

Defina como as linhas em uma partição são ordenadas. As portas que você escolher determinam a posição de uma linha dentro de uma partição. A chave de ordem pode ser crescente ou decrescente. Se você não definir chaves de ordem, as linhas não terão uma ordem específica.

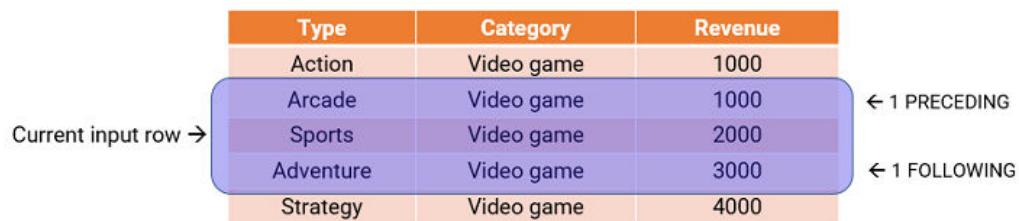
## Quadro

O quadro determina quais linhas estão incluídas no cálculo para a linha de entrada atual, com base na sua posição relativa à linha atual.

Se você usar uma função agregada em vez de LEAD ou LAG, deverá especificar um quadro de janela. LEAD e LAG fazem referência a linhas individuais e ignoram a especificação de quadro.

Os deslocamentos inicial e final descrevem o número de linhas que aparecem antes e depois da linha de entrada atual. Um deslocamento de "0" representa a linha de entrada atual. Por exemplo, um deslocamento inicial de -3 e um deslocamento final de 0 descrevem um quadro que inclui a linha de entrada atual e as três linhas antes da linha atual.

A seguinte imagem mostra um quadro com um deslocamento inicial de -1 e um deslocamento final de 1:



Type	Category	Revenue
Action	Video game	1000
Arcade	Video game	1000
Sports	Video game	2000
Adventure	Video game	3000
Strategy	Video game	4000

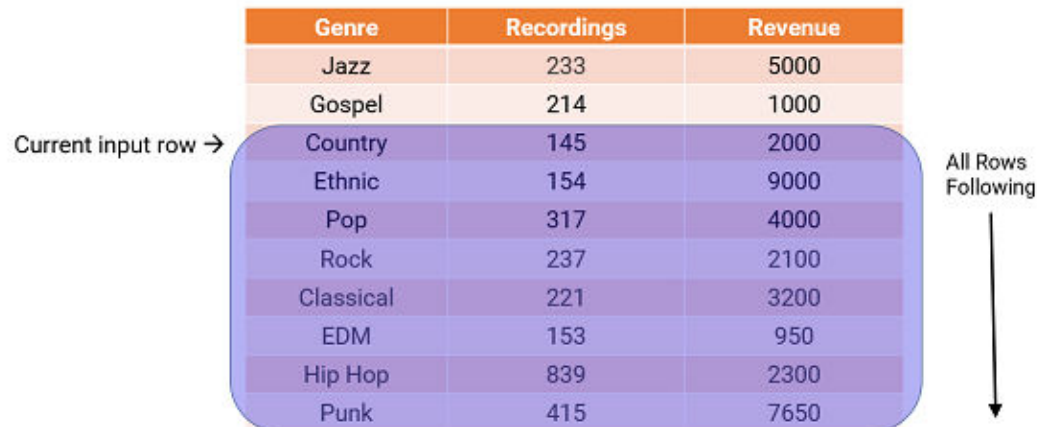
Para cada linha de entrada, a função realiza uma operação de agregação nas linhas dentro do quadro. Se você configurar uma expressão agregada como SUM com o quadro anterior, a expressão calculará a soma dos valores dentro do quadro e retornará um valor de 6000 para a linha de entrada.

Você também pode especificar um quadro que não inclua a linha de entrada atual. Por exemplo, um deslocamento inicial de 10 e um deslocamento final de 15 descrevem um quadro que inclui seis linhas no total, da décima à décima quinta linha após a linha atual.

**Nota:** O deslocamento inicial deve ser menor que ou igual ao deslocamento final.

Deslocamentos de **Todas as Linhas Anteriores** e **Todas as Linhas Seguintes** representam a primeira linha da partição e a última linha da partição. Por exemplo, se o deslocamento inicial for Todas as Linhas Anteriores e o deslocamento final for -1, o quadro incluirá uma linha antes da linha atual e todas as linhas antes dela.

A seguinte imagem ilustra um quadro com um deslocamento inicial de 0 e um deslocamento final de Todas as Linhas Seguintes:



Genre	Recordings	Revenue
Jazz	233	5000
Gospel	214	1000
Country	145	2000
Ethnic	154	9000
Pop	317	4000
Rock	237	2100
Classical	221	3200
EDM	153	950
Hip Hop	839	2300
Punk	415	7650

## Chaves de Partição e Ordem

Configure chaves de partição e ordem para formar grupos de linhas e definir a ordem ou a sequência de linhas em cada partição.

Use as seguintes chaves para especificar como agrupar e ordenar as linhas em uma janela:

### Chaves de partição

Configure chaves de partição para definir limites de partição, em vez de realizar o cálculo em todas as entradas. A função de janela opera entre as linhas que estão na mesma partição que a linha atual.

Você pode especificar as chaves de partição por valor ou parâmetro. Selecione **Valor** para usar nomes de portas. Escolha **Parâmetro** para usar um parâmetro de lista de chaves de classificação. Um parâmetro de lista de chaves de classificação contém uma lista de portas com base nas quais classificar. Se você não especificar chaves de partição, todos os dados serão incluídos na mesma partição.

### Chaves de ordem

Use chaves de ordem para determinar como as linhas em uma partição são ordenadas. As chaves de ordem definem a posição de uma linha específica em uma partição.

Você pode especificar as chaves de ordem por valor ou parâmetro. Selecione **Valor** para usar nomes de portas. Escolha **Parâmetro** para usar um parâmetro de lista de chaves de classificação. Um parâmetro de lista de chaves de classificação contém uma lista de portas com base nas quais classificar. Você também deve optar por organizar os dados em ordem crescente ou decrescente. Se você não especificar chaves de ordem, as linhas em uma partição não serão organizadas em uma ordem específica.

## Exemplo

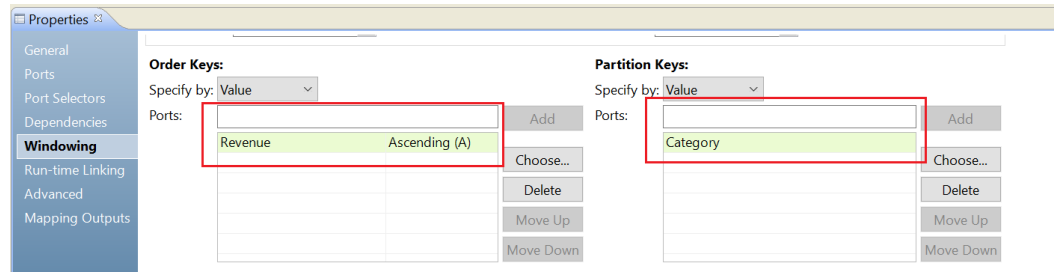
Você é o proprietário de uma cafeteria. Você deseja calcular os produtos de café e chá mais vendidos e os segundos mais vendidos.

A seguinte tabela lista os produtos, as categorias de produto correspondentes e a receita de cada produto:

Product	Category	Revenue
Espresso	Coffee	600
Black	Tea	550
Cappuccino	Coffee	500
Americano	Coffee	600
Oolong	Tea	250
Macchiato	Coffee	300
Green	Tea	450
White	Tea	650

Você particiona os dados por categoria e os ordena por receita em ordem decrescente.

A seguinte imagem mostra as propriedades que você configura na guia Definição de Janelas:



A seguinte tabela mostra os dados agrupados em duas partições de acordo com a categoria. Dentro de cada partição, a receita está organizada em ordem decrescente:

Product	Category	Revenue
Espresso	Coffee	600
Americano	Coffee	600
Cappuccino	Coffee	500
Macchiato	Coffee	300
White	Tea	650
Black	Tea	550
Green	Tea	450
Oolong	Tea	250

Com base nas especificações de particionamento e ordenação, você determina que os dois cafés mais vendidos são o Espresso e o Americano, e que os dois chás mais vendidos são o branco e o preto.

## Regras e diretrizes para a configuração da definição de janelas

Certas diretrizes se aplicam quando você configura uma transformação para definição de janelas.

Considere as seguintes regras e diretrizes ao definir propriedades de definição de janelas para uma função de janela:

- Quando você configura um quadro, o deslocamento inicial deve ser menor que ou igual ao deslocamento final. Caso contrário, o quadro não será válido.
- Configure uma especificação de quadro se você usar uma função de agregação como uma função de janela. LEAD e LAG operam com base no valor de deslocamento e ignoram a especificação do quadro.
- Não é possível usar portas complexas como chaves de partição ou ordem.
- Atribua nomes de porta exclusivos a chaves de partição e ordem para evitar erros em tempo de execução.
- As chaves de partição e ordem não podem usar ao mesmo tempo uma porta dinâmica e uma ou mais portas geradas da mesma porta dinâmica. Você deve selecionar a porta dinâmica ou as portas geradas.

# Expressões Dinâmicas

Ao configurar uma expressão em uma porta de saída dinâmica, essa expressão se torna uma expressão dinâmica. Uma expressão dinâmica pode gerar várias portas de saída.

Você pode fazer referência a um seletor de portas ou a uma porta dinâmica em uma expressão dinâmica. Quando o seletor de portas ou a porta dinâmica contém várias portas, a expressão dinâmica é executada em cada porta.

Quando você configura uma expressão dinâmica, a Developer tool não valida se as portas geradas são tipos válidos para essa expressão. Por exemplo, se você fizer referência a um seletor de portas que contém portas do tipo decimal em uma expressão que requer tipos de cadeia, a expressão aparecerá como válida em tempo de design.

## Exemplo

Uma transformação de Expressão tem as seguintes portas de entrada geradas:

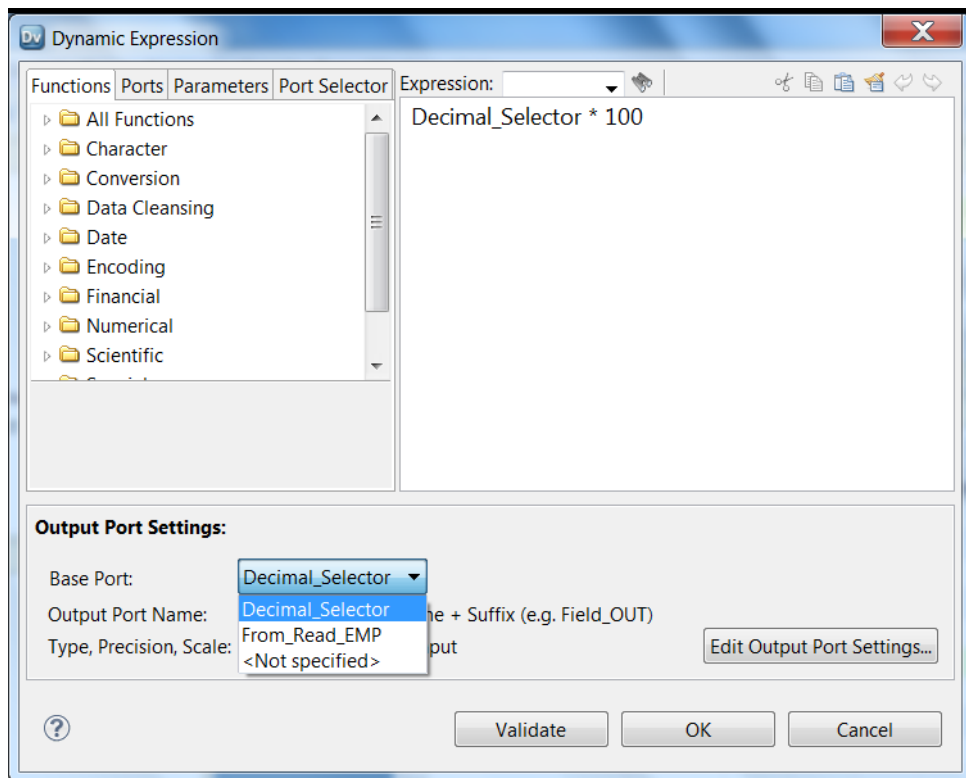
```
EMPNO    Decimal
NAME     String
SALARY    Decimal
DEPTNO    Decimal
```

A transformação contém uma porta de saída dinâmica denominada MyDynamicPort. A porta de saída retorna o resultado de uma expressão dinâmica. A expressão dinâmica multiplica o valor de cada porta em um seletor de portas por 100. A expressão é executada uma vez para cada porta no seletor de portas. Cada instância pode retornar um resultado diferente. A transformação de Expressão gera uma porta de saída separada para cada resultado.

O seletor de portas Decimal\_Selector tem uma regra de seleção que inclui as portas que são do tipo de dados decimal:

```
EMPNO    Decimal
SALARY    Decimal
DEPTNO    Decimal
```

A seguinte imagem mostra uma expressão dinâmica que faz referência ao seletor de portas Decimal\_Selector:



Edite as configurações da porta de saída para alterar nomes e propriedades de portas de saída. Também é possível escolher a porta base.

## Configurações da Porta de Saída

Você pode indicar quais portas deseja usar como entrada para uma expressão dinâmica. Selecione as portas na área **Porta Base**.

Se você escolher o seletor de portas `Decimal_Selector` como a porta base, a expressão dinâmica retornará portas do tipo decimal. A expressão dinâmica não gera uma porta para a porta `NAME`, pois ela é uma cadeia.

A seguinte imagem mostra as portas geradas na transformação:

Properties   Data Viewer   Tags   Notifications								
General								
<b>Ports</b>								
Port Selectors								
Dependencies								
Parameters								
Run-time Linking								
Advanced								
Mapping Outputs								
Name	Type	Precis...	Scale	Input	Output	Varia...	Expression	
1 From_Read_EMP	dynamic		0	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	From_Read_EMP	
1 EMPNO	decimal	3	0	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
2 NAME	string	10	0	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
3 SALARY	decimal	4	0	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
4 DEPTNO	decimal	4	0	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
2 MyDynamicPort	dynamic		0	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Decimal_Selector * 100	
1 EMPNO_OUT	decimal	3	0	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
2 SALARY_OUT	decimal	4	0	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
3 DEPTNO_OUT	decimal	4	0	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		

Embora a porta dinâmica `From_Read_Emp` seja uma porta de entrada/saída, a transformação retorna apenas as portas na porta de saída dinâmica `MyDynamicPort`.

Você pode configurar como deseja nomear as portas de saída. O nome da porta de saída padrão é o nome da porta de entrada e o sufixo `_OUT`.

É possível trocar a porta base por um seletor de portas.

A seguinte imagem mostra as configurações da porta de saída no Editor de Expressão:

**Output Port Settings:**

Base Port: From\_Read\_EMP ▼

Output Port Name: Primary input port name + Suffix (e.g. Field\_O)

Type, Precision, Scale: Inherit from primary input Edit Output Port Settings...

Se você configurar a porta base como From\_Read\_EMP, selecione a porta dinâmica que contém todas as portas de entrada geradas. O Serviço de Integração de Dados executa a expressão dinâmica com base em todas as portas de From\_Read\_EMP.

A seguinte imagem mostra as portas de saída geradas com base na entrada From\_Read\_Emp:

General								
Ports								
	Name	Type	Precis...	Scale	Input	Output	Varia...	Expression
1	From_Read_EMP	dynamic		0	✓	✓		From_Read_EMP
1	EMPNO	decimal	3	0				
2	NAME	string	10	0				
3	SALARY	decimal	4	0				
4	DEPTNO	decimal	4	0				
2	MyDynamicPort	dynamic		0		✓		Decimal_Selector * 100
1	EMPNO_OUT	decimal	3	0				
2	NAME_OUT	string	10	0				
3	SALARY_OUT	decimal	4	0				
4	DEPTNO_OUT	decimal	4	0				

As portas de saída geradas incluem uma porta de saída denominada NAME\_OUT, que é um tipo de cadeia.

O Serviço de Integração de Dados gera portas de saída para cada expressão dinâmica. Se você criar uma expressão dinâmica que gera 15 portas e definir outra expressão dinâmica que gera 5 portas, o Serviço de Integração de Dados gerará 20 portas de saída. Cada porta de saída dinâmica gera um grupo diferente de portas.

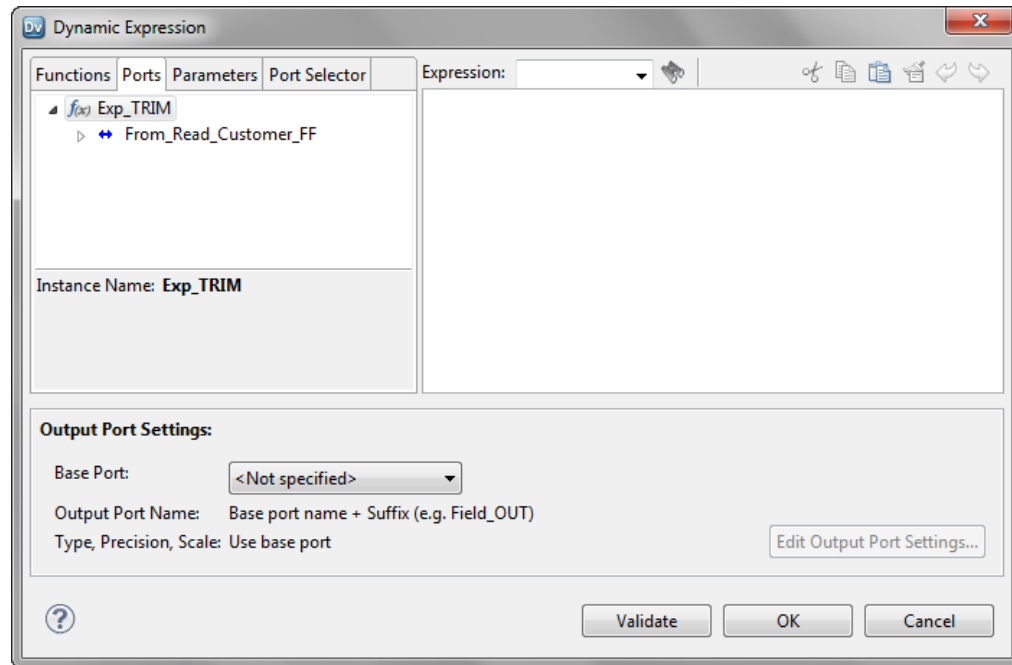
## Criando uma Expressão Dinâmica

Crie uma expressão dinâmica em uma transformação de Expressão para executar essa expressão uma vez para cada porta em uma porta dinâmica ou em um seletor de portas. A expressão dinâmica retorna os resultados a uma porta gerada separada para cada instância.

1. Na transformação de Expressão, acesse a exibição **Propriedades** e clique na guia **Portas**.
2. Clique em **Nova Porta Dinâmica**.  
A Developer tool cria uma porta dinâmica com propriedades padrão.
3. Renomeie a porta dinâmica e desative a opção de entrada.  
A porta dinâmica deve ser uma porta de saída.
4. Na coluna **Expressão** da porta de saída dinâmica, clique no botão **Abrir** (🔽).



A caixa de diálogo **Expressão Dinâmica** é exibida:



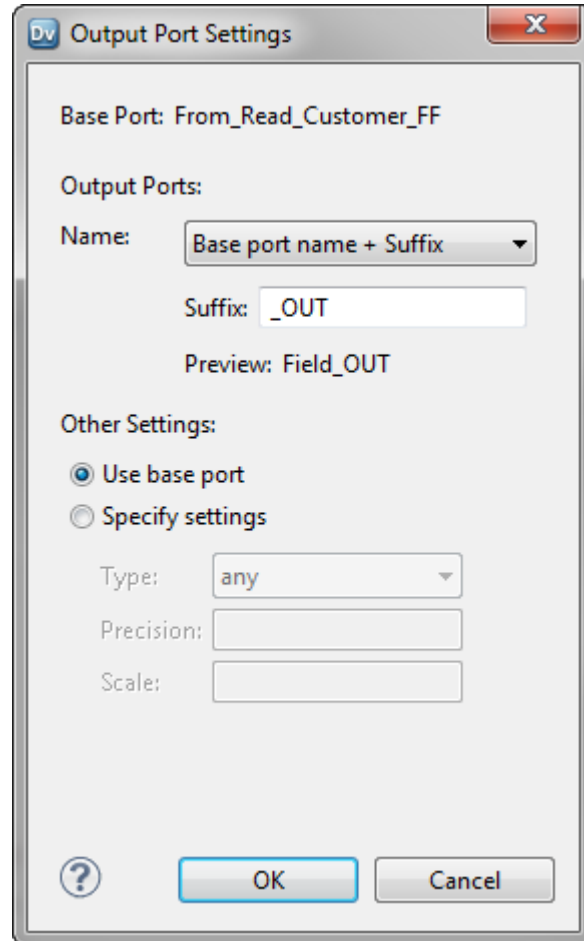
5. No Editor de Expressão, insira uma expressão. Essa expressão pode incluir um seletor de portas ou uma porta dinâmica.  
Por exemplo, `LTRIM(RTRIM(Dynamic_Customer))`, em que `Dynamic_Customer` é uma porta dinâmica.
6. Clique em **Validar** para validar a expressão.
7. Clique em **OK** para sair da caixa de diálogo **Validar Expressão**.
8. Na área **Configurações da Porta de Saída**, selecione a porta de saída dinâmica na lista **Porta Base** ou escolha um seletor de portas que você referenciou na expressão.

A Developer tool gera portas de saída com base na opção selecionada.

9. Use as seguintes etapas para renomear as portas de saída:

- a. Clique em **Editar Configurações da Porta de Saída**.

A caixa de diálogo **Configurações da Porta de Saída** é exibida.



- b. Na lista **Nome**, selecione uma das opções e insira um valor para o prefixo ou sufixo. Se você tiver selecionado **Cadeia fixa + Número automático**, insira o texto para o nome da porta de saída. Por exemplo, se você inserir TRIM para o nome da porta de saída, os nomes das portas de saída aparecerão como TRIM1, TRIM2, TRIM3.
  - c. Opcionalmente, escolha **Especificar configurações** na área **Outras configurações** para alterar o tipo, a precisão e a escala das portas de saída. Por padrão, as portas de saída utilizam as configurações das portas base.
  - d. Clique em **OK**.
10. Clique em **OK** para sair do editor de **Expressão Dinâmica**.

# Transformação de Expressão - Propriedades Avançadas

Configure propriedades para determinar como o Serviço de Integração de Dados processa dados para a transformação de Expressão.

Configure as seguintes propriedades avançadas para uma transformação de Expressão:

## **Nível de Rastreamento**

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.

## **Manter Ordem das Linhas**

Mantenha a ordem de linhas dos dados de entrada na transformação. Selecione essa opção se o Serviço de Integração de Dados não realizar qualquer otimização que possa alterar a ordem de linhas.

Quando o Serviço de Integração de Dados realiza as otimizações, ele pode perder uma ordem estabelecida anteriormente no mapeamento. Você pode estabelecer a ordem em um mapeamento com uma origem de arquivo simples classificada, uma origem relacional classificada ou uma transformação do Classificador. Quando você configura uma transformação para manter a ordem de linhas, o Serviço de Integração de Dados considera essa configuração quando ele executa as otimizações para o mapeamento. O Serviço de Integração de Dados realizará otimizações para a transformação se ela puder manter a ordem. O Serviço de Integração de Dados não realizará otimizações para a transformação se a otimização alterar a ordem de linha.

# Transformação de Expressão em um ambiente não nativo

O processamento da transformação de Expressão em um ambiente não nativo depende do mecanismo que executa a transformação.

Considere o suporte para os seguintes mecanismos de tempo de execução não nativos:

- Mecanismo Blaze. Suportado com restrições.
- Mecanismo Spark. Com suporte com restrições em mapeamentos de fluxo e em lote.
- Mecanismo Databricks Spark. Suportado com restrições.

## Transformação de Expressão no mecanismo Blaze

A validação do mapeamento falha nas seguintes situações:

- A transformação contém portas de variável com estado.
- A transformação contém funções não suportadas em uma expressão.

Uma transformação de Expressão com uma função definida pelo usuário retorna um valor nulo para linhas que possuem um erro de exceção na função.

## Transformação de Expressão no mecanismo Spark

A validação do mapeamento falha nas seguintes situações:

- A transformação contém portas de variável com estado.
- A transformação contém funções não suportadas em uma expressão.

**Nota:** Se uma expressão resultar em erros numéricos, como divisão por zero ou SQRT de um número negativo, ela retornará um valor nulo e as linhas não aparecerão na saída. No ambiente nativo, a expressão retorna um valor infinito ou NaN.

### Transformação de Expressão em um mapeamento de fluxo

Os mapeamentos de fluxo têm as mesmas regras de processamento que os mapeamentos em lote no mecanismo Spark.

## Transformação de Expressão no mecanismo Databricks Spark

A validação do mapeamento falha nas seguintes situações:

- A transformação contém portas de variável com estado.
- A transformação contém funções não suportadas em uma expressão.

**Nota:** Se uma expressão resultar em erros numéricos, como divisão por zero ou SQRT de um número negativo, ela retornará um valor nulo e as linhas não aparecerão na saída. No ambiente nativo, a expressão retorna um valor infinito ou NaN.

## CAPÍTULO 17

# Transformação de Filtro

Este capítulo inclui os seguintes tópicos:

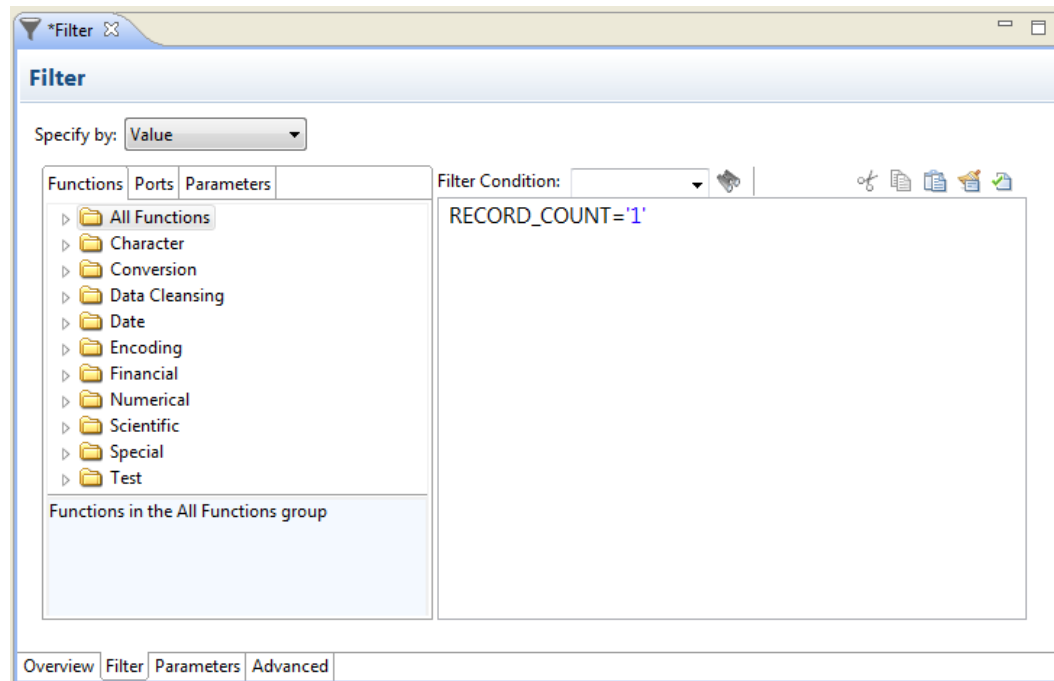
- [Transformação de Filtro - Visão Geral, 297](#)
- [Transformações de Filtro em Mapeamentos Dinâmicos, 298](#)
- [Condição de Filtro, 299](#)
- [Transformação de Filtro - Propriedades Avançadas, 301](#)
- [Transformação de Filtro - Dicas de Desempenho, 301](#)
- [Transformação de Filtro em um ambiente não nativo, 301](#)

## Transformação de Filtro - Visão Geral

Use a transformação de Filtro para filtrar as linhas em um mapeamento. Como uma transformação ativa, a transformação de Filtro pode alterar o número de linhas que passam por ela.

A transformação Filtro permite a transferência de linhas que cumpram a condição de filtro especificada. Ela descarta linhas que não cumprem a condição. Você pode filtrar os dados com base em uma ou mais condições.

A seguinte imagem mostra uma condição de filtro em uma transformação de Filtro:



Uma condição de filtro retorna TRUE ou FALSE para cada linha que o Serviço de Integração de Dados avalia, com base no fato de a linha cumprir ou não a condição especificada. O Serviço de Integração de Dados passa pela transformação cada linha que retorna TRUE. O Serviço de Integração de Dados descarta cada linha que retorna FALSE e grava uma mensagem no log.

Você não pode concatenar portas de mais de uma transformação na transformação Filtro. As portas de entrada para o filtro devem vir de uma única transformação.

## Transformações de Filtro em Mapeamentos Dinâmicos

É possível usar uma transformação de Filtro em um mapeamento dinâmico. Você pode configurar portas dinâmicas na transformação e fazer referência às portas geradas na condição de filtro.

É possível parametrizar a condição de filtro completa. Configure um parâmetro de expressão com um valor padrão que contenha a expressão inteira. A Developer tool não valida uma condição de filtro em um valor padrão de parâmetro.

É possível fazer referência a uma porta dinâmica em uma condição de filtro. A porta dinâmica pode conter várias portas geradas. O Serviço de Integração de Dados expande a condição de filtro de forma que ela contenha cada porta gerada. Cada porta gerada deve ser um tipo válido para inclusão na expressão.

É possível fazer referência a uma porta gerada em uma condição de filtro. No entanto, se a porta gerada não existir em tempo de execução, o mapeamento falhará.

# Condição de Filtro

A condição de filtro é uma expressão que retorna TRUE ou FALSE.

Insira condições no Editor de Expressão. A condição de filtro diferencia maiúsculas de minúsculas.

Você pode usar como filtro qualquer expressão que retorne um valor único. Por exemplo, se você deseja eliminar por filtragem as linhas de funcionários cujos salários sejam inferiores ou iguais a US\$ 30.000, insira a seguinte condição:

```
SALARY > 30000
```

Você pode especificar vários componentes da condição usando os operadores lógicos AND e OR. Se desejar filtrar os funcionários que ganham menos de US\$ 30.000 e mais de US\$ 100.000, insira a seguinte condição:

```
SALARY > 30000 AND SALARY < 100000
```

Você pode usar portas, parâmetros, portas dinâmicas e portas geradas na condição de filtro. Selecione as portas e os parâmetros no Editor de Expressão.

Se você usar uma porta dinâmica na condição de filtro, a condição se expandirá para incluir todas as portas geradas na porta dinâmica. Por exemplo, a porta dinâmica, MyDynamicPort, contém três portas decimais:

```
Salary  
Bonus  
Stock
```

Se você configurar a seguinte condição de filtro:

```
MyDynamicPort > 100
```

A condição de filtro se expandirá para a seguinte expressão:

```
Salary > 100 AND Bonus > 100 AND Stock > 100
```

É possível inserir uma constante para a condição de filtro. O equivalente numérico de FALSE é zero (0). Qualquer valor diferente de zero equivale a TRUE. Por exemplo, a transformação contém uma porta com o nome NUMBER\_OF\_UNITS com um tipo de dados numérico. Você configura uma condição de filtro para retornar FALSE se o valor de NUMBER\_OF\_UNITS for igual a zero. Caso contrário, a condição retorna TRUE.

**Nota:** Não é possível usar um único seletor de portas ou porta dinâmica como valor booliano.

Não é necessário especificar TRUE ou FALSE como valores na expressão. TRUE e FALSE são valores de retorno implícitos de qualquer condição que você definir. Se a condição de filtro for avaliada como NULL, a linha será FALSE.

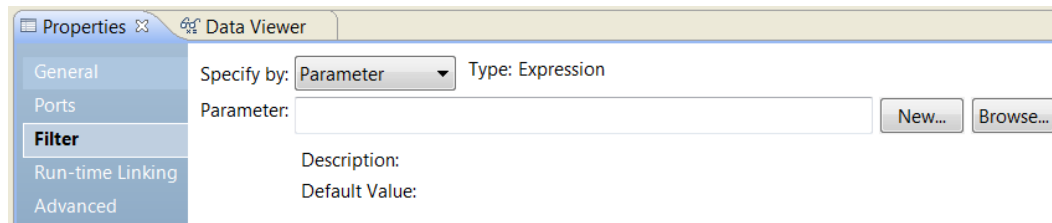
## Parametrizar a Condição de Filtro

É possível configurar um parâmetro de expressão para definir a condição de filtro. Um parâmetro de expressão contém a expressão inteira.

Talvez seja necessário parametrizar a condição de filtro quando a transformação de Filtro está em um mapeamento dinâmico. A condição de filtro pode mudar de acordo com as portas geradas na transformação em tempo de execução.

Para usar um parâmetro de expressão para a condição de filtro, escolha **Especificar por Parâmetro** na guia **Filtro** das propriedades da transformação de Filtro.

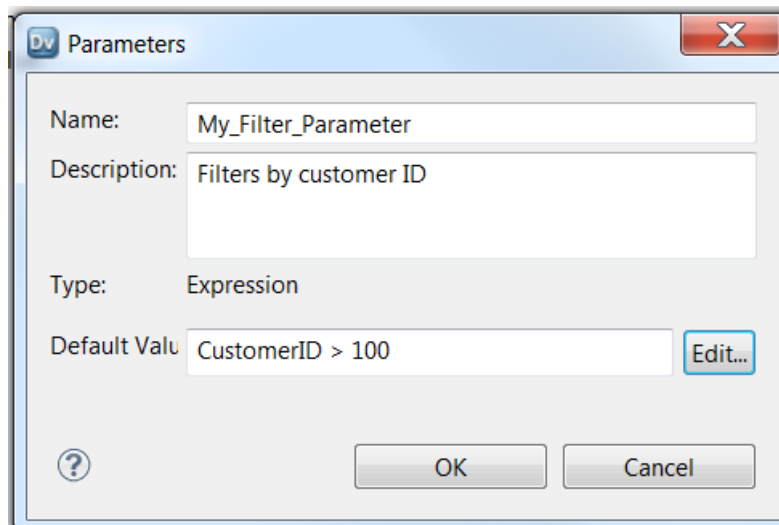
A seguinte imagem mostra a guia **Filtro** quando você especifica a condição de filtro com um parâmetro:



É possível procurar e selecionar um parâmetro de expressão que você já criou. Outra alternativa é criar um parâmetro de expressão.

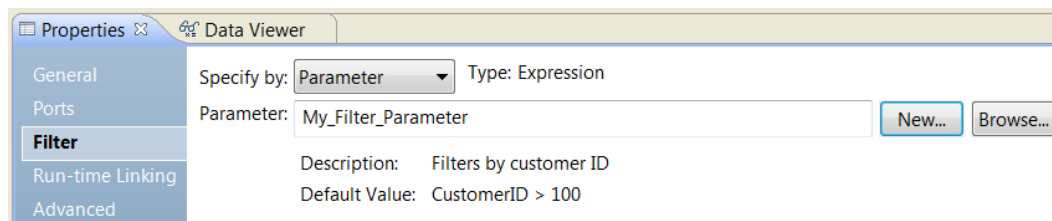
Para criar um parâmetro de expressão, clique em **Novo**. Insira um nome para o parâmetro, uma descrição e o valor da expressão padrão.

A imagem a seguir mostra onde você insere o parâmetro:



É possível inserir uma expressão predefinida na caixa de diálogo Parâmetros. Se quiser usar um editor de Expressão, clique em Editar. Ao usar o editor de Expressão, você pode selecionar funções e portas para uso na expressão. Você pode validar a expressão.

A seguinte imagem mostra a guia Filtro com um parâmetro para a condição de filtro:



O parâmetro de expressão é um parâmetro de mapeamento. É possível substituir o parâmetro em um conjunto de parâmetros ou um arquivo de parâmetros em tempo de execução.



## Filtrando Linhas com Valores Nulos

Para filtrar as linhas que contêm valores nulos ou espaços, use as funções ISNULL e IS\_SPACES para testar o valor da porta.

Por exemplo, se desejar filtrar as linhas que contêm o valor NULL na porta FIRST\_NAME, use a seguinte condição:

```
IIF (ISNULL (FIRST_NAME) , FALSE, TRUE)
```

Esta condição declara que, se a porta FIRST\_NAME é NULL, o valor de retorno é FALSE, e a linha deve ser descartada. Do contrário, a linha é passada para a próxima transformação.

## Transformação de Filtro - Propriedades Avançadas

Configure propriedades que ajudam a determinar como o Serviço de Integração de Dados processa dados para a transformação de Filtro.

Você pode configurar níveis de rastreamento para logs.

Configure a seguinte propriedade na guia **Avançado**:

### Nível de Rastreamento

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.

## Transformação de Filtro - Dicas de Desempenho

Use dicas para melhorar o desempenho da transformação de Filtro.

### Use a transformação de Filtro no início do mapeamento.

Mantenha a transformação de Filtro o mais próximo possível das origens no mapeamento. Em vez de passar pelo mapeamento linhas que você planeja descartar, filtre dados indesejáveis logo no início do fluxo de dados das origens para os destinos.

## Transformação de Filtro em um ambiente não nativo

O processamento da transformação de Filtro em um ambiente não nativo depende do mecanismo que executa a transformação.

Considere o suporte para os seguintes mecanismos de tempo de execução não nativos:

- Mecanismo Blaze. Suportado com restrições.
- Mecanismo Spark. Com suporte sem restrições em mapeamentos de fluxo e em lote.
- Mecanismo Databricks Spark. Suportado sem restrições.

## Transformação de Filtro no mecanismo Blaze

Quando um mapeamento contém uma transformação de Filtro em uma coluna particionada de uma origem de Hive, o mecanismo Blaze pode ler apenas as partições contendo dados que satisfazem a condição de filtro. Para enviar o filtro para a origem de Hive, configure a transformação de Filtro para ser a próxima transformação no mapeamento após a origem.

## CAPÍTULO 18

# Transformação Hierárquica para Relacional

Este capítulo inclui os seguintes tópicos:

- [Visão geral da transformação de Hierárquica para Relacional, 303](#)
- [Exemplo - transformação de Hierárquica para Relacional, 304](#)
- [Portas relacionais de saída e a exibição Visão Geral, 305](#)
- [Portas da Transformação Hierárquica para Relacional, 306](#)
- [Referências de esquema, 307](#)
- [Configuração de portas, 307](#)
- [Desenvolvimento da Transformação Hierárquica para Relacional, 308](#)

## Visão geral da transformação de Hierárquica para Relacional

A transformação de Hierárquica para Relacional processa uma entrada hierárquica XML ou JSON e a transforma em uma saída relacional. Uma transformação de Hierárquica para Relacional lê a entrada hierárquica de portas de entrada e transforma os dados em uma saída relacional nas portas de saída da transformação. Para transformar uma entrada hierárquica em uma saída relacional, use um arquivo de esquema para definir os dados hierárquicos.

É possível usar o assistente de transformação de Hierárquica para Relacional para mapear os dados automaticamente. Você pode configurar o mapeamento para as portas de saída relacionais na exibição **Visão Geral** da transformação.

Depois que o assistente gera a transformação, você pode transmitir os dados das portas de saída relacionais para outra transformação em um mapeamento.

**Nota:** A transformação Hierárquica em Relacional pode processar até 10.000 elementos de esquema em um arquivo .xsd. Para processar mais de 10.000 elementos, divida os dados em vários arquivos.

## Exemplo - transformação de Hierárquica para Relacional

O departamento de Logística da empresa Harrinder Shipping precisa processar dados de remessa. Primeiro, eles precisam transformar dados de inventário e clientes do formato hierárquico para dados relacionais que podem ser armazenados em tabelas de banco de dados.

Eles precisam criar um mapeamento que transforme dados hierárquicos em dados relacionais. O sistema de inventário da organização gera dados de inventário de remessa em formato hierárquico. O mapeamento precisa usar uma transformação de Hierárquica para Relacional que processa a entrada de dados de remessa e gera a saída dos detalhes em um formato relacional utilizável.

A entrada Shipments está em formato hierárquico. O elemento Shipment contém subelementos com dados de clientes e inventário para cada remessa:

```
Shipments
Shipment
  Items
    Item_Name
    Inventory_ID
  Customer
    Customer_Name
    Customer_ID
    Customer_Address
```

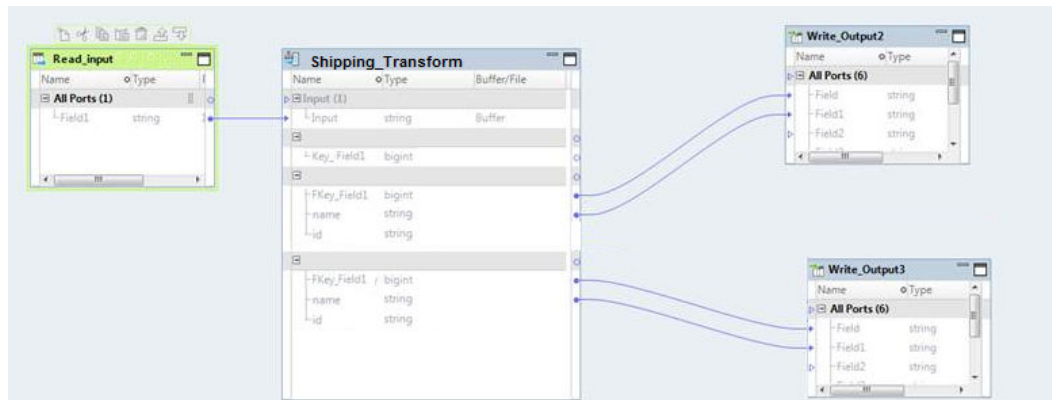
Na saída relacional, o elemento Customer\_ID é uma chave primária na tabela Customer e uma Chave externa na tabela Shipment.

Customer_ID	Customer_Name	Customer_Address
3543766	Tony Birch	6 Moby Drive
6342562	Sujita Man	22 Dan Street
6471862	Dwayne Horace	7 Jafendar Boulevard
7265204	Carmela Perez	23 Dan Street
4559672	Delilah Soraya	28 Jafendar Boulevard

Shipment_ID	Inventory_Item	Customer_ID
9173327437	908274	7265204
9174562342	553439	7265204
8484526471	546584	3543766
7023847265	908274	3543766
9174596725	553439	3543766

A seguinte imagem mostra o mapeamento nesse exemplo:



O mapeamento contém os seguintes objetos:

#### **Read\_input**

A origem que contém o caminho para o arquivo com dados hierárquicos. Lê dados de cobrança de um arquivo XML.

#### **Shipping\_Transform**

Uma transformação de Hierárquica para Relacional que transforma entrada XML em saída relacional.

#### **Write\_Output2**

Um destino que armazena parte dos dados transformados, a tabela Customer, em formato relacional.

#### **Write\_Output3**

Um segundo destino que armazena outra parte dos dados transformados, a tabela Shipment, em formato relacional.

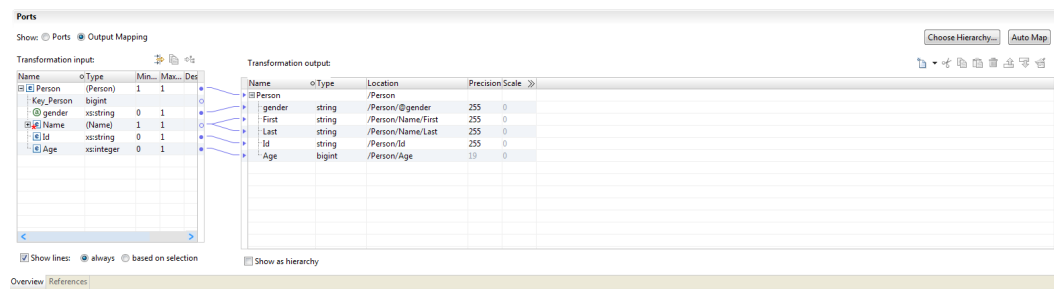
O mapeamento usa o arquivo simples **Read\_input** para inserir o caminho de destino da entrada hierárquica. O mapeamento processa e transforma dados com a transformação **Shipping\_Transform**. Em seguida, o mapeamento armazena a saída nos dois destinos de saída.

## Portas relacionais de saída e a exibição Visão Geral

Para converter dados hierárquicos em uma saída relacional na transformação de Hierárquica para Relacional, o assistente gera links entre nós hierárquicos para portas relacionais. Use a exibição **Visão Geral** para visualizar os links entre portas relacionais para portas hierárquicas. Também é possível criar grupos de portas de saída vinculando nós da saída hierárquica a grupos de portas.

Para visualizar o mapeamento de grupos relacionais, use a exibição **Visão Geral**. Selecione **Mapeamento de Saída**. O painel **Portas** aparece na exibição **Visão Geral**.

A seguinte imagem mostra o painel **Portas**:



A área **Entrada da transformação**, que mostra o esquema hierárquico, está à esquerda. A área **Saída de transformação**, que mostra as portas de saída relacionais, está à direita.

É possível definir portas de saída relacionais na área **Saída de transformação** e vincular nós do esquema a essas portas. Também é possível arrastar o ponteiro de um nó no esquema até um campo vazio na área **Saída de transformação** para criar uma porta. Quando você arrasta um nó do esquema de saída até uma porta, a Developer tool mostra um link entre eles.

## Portas da Transformação Hierárquica para Relacional

As portas da transformação de Hierárquica para Relacional são definidas na exibição **Visão Geral** da transformação.

Uma transformação de Hierárquica para Relacional pode ler a entrada de um arquivo ou buffer. As portas de saída retornam dados relacionais da transformação.

Quando você cria uma transformação de Hierárquica para Relacional, a Developer tool cria uma porta de entrada padrão. O tipo de entrada determina o tipo de dados que o Serviço de Integração de Dados transmite para a transformação de Hierárquica para Relacional. O tipo de entrada determina se ela é um caminho de arquivos de origem ou de dados.

Configure um dos seguintes tipos de entrada:

### Buffer

A transformação de Hierárquica para Relacional recebe linhas de dados de origem na porta de Entrada. Use o tipo de entrada de buffer ao configurar a transformação para receber dados de uma transformação do Informatica.

### Arquivo

A transformação de Hierárquica para Relacional recebe o caminho do arquivo de origem na porta de Entrada. A transformação de Hierárquica para Relacional abre o arquivo de origem. Também é possível usar o tipo de entrada de Arquivo para arquivos extensos que podem exigir muita memória do sistema para processamento com uma porta de entrada de buffer.

Ao criar a transformação com o assistente de Nova Transformação, você pode definir um arquivo de entrada de exemplo. Um arquivo de entrada de exemplo é uma pequena amostra do arquivo de entrada. Faça referência a um arquivo de entrada de exemplo ao criar uma transformação de Hierárquica para Relacional. Também é possível usar o arquivo de entrada de exemplo ao testar a transformação na exibição

### Visualizador de Dados.

A transformação contém um ou mais grupos de portas que retornam dados relacionais.

## Referências de esquema

Uma transformação de Hierárquica para Relacional requer um esquema hierárquico para definir a hierarquia de entrada na transformação. Para usar o esquema na transformação, defina uma referência de esquema.

É possível definir referências de esquema de transformação na exibição **Referências** da transformação.

A transformação de Hierárquica para Relacional faz referência a objetos de esquema no repositório do Modelo. Os objetos de esquema podem existir no repositório antes de você criar a transformação. Você também pode importar os esquemas da exibição **Referências** da transformação.

Um esquema pode fazer referência a esquemas adicionais. A exibição **Referências** mostra o espaço de nome e o prefixo para cada esquema referenciado pela transformação de Hierárquica para Relacional. Quando você faz referência a vários esquemas com espaços de nome vazios, a transformação não é válida.

## Configuração de portas

No painel **Portas**, a transformação mostra o mapeamento entre os nós de esquema hierárquicos e as portas relacionais. A transformação usa um esquema para definir a entrada hierárquica. Se o esquema tiver mais de um elemento que possa ser um elemento raiz, escolha um nó para ser o elemento raiz.

O assistente gera links entre nós de esquema hierárquicos e portas relacionais. Se você deseja alterar os links gerados, pode usar o painel **Portas** para adicionar, excluir ou editar links. É possível vincular nós a portas e criar uma porta.

Quando você vincula nós à área **Saída de transformação**, a Developer tool atualiza o campo de localização com a localização do nó na hierarquia. Se você criar portas manualmente, deverá mapear um nó para a porta. Atualize a coluna **Localização** e selecione um nó da lista.

Quando você vincula um nó de ocorrência múltipla a um grupo que contém o elemento pai, pode configurar o número de ocorrências de elementos filho a serem incluídas. Outra opção é substituir o grupo pai pelo grupo filho de ocorrência múltipla na saída de transformação.

Para criar um grupo, vincule um nó a uma coluna vazia na área **Saída de transformação**. Se você vincular um nó filho de ocorrência múltipla a uma coluna de entrada ou saída vazia, a Developer tool solicitará que o grupo seja relacionado a outros grupos de saída. Quando você seleciona um grupo, a Developer tool cria chaves para relacionar os grupos.

Configure grupos relacionados de portas de saída na área **Saída de transformação**. Quando a Developer tool solicita que você relacione grupos de saída, ela adiciona as chaves aos grupos. Você também pode adicionar portas manualmente para representar chaves.

# Desenvolvimento da Transformação Hierárquica para Relacional

Use o assistente de Nova Transformação para gerar uma transformação de Hierárquica para Relacional automaticamente. Escolha um arquivo de amostra de esquema ou hierárquico para definir a hierarquia de entrada.

1. Crie a transformação na Developer tool.
2. Configure a porta de entrada e o mapeamento.
3. Teste a transformação.

## Criando a transformação de Hierárquica para Relacional

1. Na Developer tool, clique em **Arquivo > Novo > Transformação**.
2. Selecione a transformação de Hierárquica para Relacional e clique em **Avançar**.
3. Insira um nome para a transformação, procure uma localização do repositório do Modelo para inserir a transformação e clique em **Avançar**.
4. Para selecionar um esquema, selecione um dos seguintes métodos:
  - Para usar um esquema do repositório do Modelo para definir a hierarquia de entrada, perto do campo **Objeto de Esquema**, navegue para selecionar o arquivo de esquema do repositório.
  - Para importar um arquivo de esquema, clique em **Criar um novo objeto de esquema**. Na janela **Novo objeto de esquema**, você pode procurar e selecionar um arquivo de esquema ou pode optar por criar um esquema a partir de arquivo hierárquico de amostra.
5. Escolha a raiz da hierarquia de saída. Na caixa de diálogo **Raiz da hierarquia**, selecione o elemento no esquema que é o elemento raiz do arquivo hierárquico de saída. Para ajudar a selecionar o objeto raiz, você pode adicionar um arquivo hierárquico de amostra. Para adicionar um arquivo de amostra, ao lado do campo **Arquivo de amostra**, procure e selecione o arquivo no sistema de arquivos.
6. Clique em **Concluir**.  
O assistente cria a transformação no repositório.

## Configurando as Portas e o Mapeamento

Configure as portas de entrada e saída na exibição **Visão Geral**.

1. Selecione o tipo de dados da porta de entrada, o tipo de porta, a precisão e a escala.
2. Para visualizar o mapeamento, na área **Portas** da exibição **Visão Geral**, selecione **Mapeamento de Saída**.
3. Expanda as árvores na grade **Portas**. À esquerda, o painel **Entrada de transformação** mostra a entrada hierárquica esperada e, à direita, o painel **Saída de transformação** mostra a saída relacional.
4. Para definir um nó como uma raiz, clique em **Escolher Hierarquia**.  
A Developer tool exibe apenas os nós do nível raiz e abaixo dele na área **Entrada de transformação**.
5. Para exibir linhas que conectam as portas aos nós hierárquicos, clique em **Mostrar Linhas**. Selecione para exibir todas as linhas de conexão ou apenas as linhas das portas selecionadas.
6. Para adicionar uma porta ou um grupo de entrada à área **Saída de transformação**, use um dos seguintes métodos:



- Arraste um elemento simples ou complexo na área **Entrada de transformação** até uma coluna vazia na área **Saída de transformação**. Se o nó for um nó de grupo, a Developer tool adicionará um grupo relacional sem portas.
  - Para adicionar um grupo relacional, selecione uma linha e clique com o botão direito do mouse para selecionar **Novo > Grupo**.
  - Para adicionar uma porta relacional, clique com o botão direito do mouse para selecionar **Novo > Campo**.
7. Para limpar as configurações de nós hierárquicos para localizações de portas, use um dos seguintes métodos:
    - Selecione um ou mais nós na área **Entrada de transformação**, clique com o botão direito do mouse e selecione **Limpar**.
    - Selecione uma ou mais linhas que conectam as portas relacionais aos nós hierárquicos, clique com o botão direito do mouse e selecione **Excluir**.
  8. Para exibir as portas de saída em uma hierarquia, clique em **Mostrar como Hierarquia**. Cada grupo filho é exibido embaixo do grupo pai.

## Testando a Transformação

Teste a transformação de Hierárquica para Relacional na exibição **Visualizador de Dados**.

Antes de testar a transformação, verifique se você definiu a localização do arquivo de entrada. A localização de entrada é definida na máquina do Serviço de Integração de Dados, na coluna Localização de Entrada do painel **Portas**, na exibição **Visão Geral**.

1. Abra a exibição **Visualizador de Dados**.
2. Clique em **Executar**.

A Developer tool valida a transformação. Se não houver erros, a Developer tool mostrará o conteúdo do arquivo hierárquico no painel **Saída**.

## CAPÍTULO 19

# Transformação Java

Este capítulo inclui os seguintes tópicos:

- [Visão Geral de Transformações Java, 310](#)
- [Projetando uma Transformação Java, 314](#)
- [Portas da Transformação Java, 314](#)
- [Transformação Java - Propriedades Avançadas, 315](#)
- [Desenvolvendo Código Java, 318](#)
- [Propriedades Java da Transformação Java, 322](#)
- [Otimização de Filtro com a Transformação Java, 325](#)
- [Criando uma Transformação Java, 328](#)
- [Compilando uma transformação Java, 329](#)
- [Solucionando Problemas de uma Transformação Java, 329](#)
- [Convertendo para exemplo de dados struct, 331](#)
- [Transformação Java em um ambiente não nativo, 334](#)

## Visão Geral de Transformações Java

Use a transformação Java para ampliar a funcionalidade da ferramenta Developer.

A transformação Java fornece uma interface de programação nativa simples para definir a funcionalidade de transformação com a linguagem de programação Java. Você pode usar a transformação Java para definir funcionalidade de transformação simples ou moderadamente complexa, sem conhecimento avançado da linguagem de programação Java ou um ambiente de desenvolvimento Java externo. A transformação Java pode ser ativa ou passiva.

A ferramenta Developer usa o Java Development Kit (JDK) para compilar o código Java e gerar o código de bytes para a transformação. A ferramenta Developer armazena o código de bytes no repositório do Modelo.

O Data Integration Service usa o Java Runtime Environment (JRE) para executar o código de bytes gerado em tempo de execução. Quando o Data Integration Service executa um mapeamento com uma transformação Java, ele usa o JRE para executar o código de bytes e as linhas de entrada do processo, bem como gerar linhas de saída.

Crie transformações Java escrevendo snippets de código Java que definem a lógica da transformação. Defina o comportamento para uma transformação Java, com base nos eventos a seguir:

- A transformação recebe uma linha de entrada.

- A transformação processou todas as linhas de entrada.

Nos mapeamentos executados no mecanismo Spark, você pode usar tipos de dados complex em transformações Java para processar dados hierárquicos. Com tipos de dados complex, o mecanismo Spark lê, processa e grava diretamente dados hierárquicos em arquivos complexos Avro, Parquet e JSON.

## Transformações Java Reutilizáveis e Não Reutilizáveis

Você pode criar uma transformação Java reutilizável ou não reutilizável.

Transformações reutilizáveis podem existir em vários mapeamentos. As transformações não reutilizáveis existem dentro de um único mapeamento.

As exibições no editor onde você define as propriedades e cria o código Java são diferentes, dependendo do fato de você estar criando uma transformação Java reutilizável ou não reutilizável.

## Transformações Java Ativas e Passivas

Ao criar uma transformação Java, você define o tipo como ativo ou passivo.

Depois de definir o tipo de transformação, não será possível alterá-la.

Uma transformação Java executa o código Java que você define na guia **Na Entrada** uma vez para cada linha de dados de entrada.

Uma transformação Java trata linhas de saída com base no tipo de transformação da seguinte forma:

- Uma transformação Java passiva gera uma linha de saída para cada linha de entrada na transformação após processar cada linha de entrada.
- Uma transformação Java ativa gera várias linhas de saída para cada linha de entrada na transformação.

Use o método `generateRow` para gerar cada linha de saída. Por exemplo, se a transformação contiver duas portas de entrada que representem uma data de início e uma data de término, você poderá usar o método `generateRow` para gerar uma linha de saída para cada data entre a data de início e a data de término.

## Conversão de Tipo de Dados

Uma transformação Java converte os tipos de dados da Developer tool do em tipos de dados Java com base no tipo da porta da transformação Java.

Quando uma transformação Java lê linhas de entrada, ela converte tipos de dados da porta de entrada em tipos de dados Java.

Quando uma transformação Java grava as linhas de saída, ela converte os tipos de dados Java em tipos de dados da porta de saída.

Por exemplo, o seguinte processamento ocorre para uma porta de entrada com o tipo de dados integer em uma transformação Java:

1. A transformação Java converte o tipo de dados integer da porta de entrada para o tipo de dado int primitivo Java.
2. Na transformação, a transformação trata o valor da porta de entrada como o tipo de dado int primitivo Java.
3. Quando a transformação gera a linha de saída, ele converte o tipo de dado int primitivo Java para o tipo de dados integer.

A tabela a seguir mostra como a transformação Java mapeia os tipos de dados da Developer tool do para tipos de dados primitivos e complexos Java:

Tipo de dados da Developer Tool	Tipos de dados Java
array*	java.util.List
bigint	long
binary	byte[]
date/time	Com o processamento de nanossegundos habilitado, BigDecimal com precisão de nanossegundo Com o processamento de nanossegundos desabilitado, long com precisão de milissegundos (o número de milissegundos desde 1 de janeiro de 1970 00:00:00.000 GMT)
decimal	Com processamento de alta precisão desabilitado, double com precisão 15 Com processamento de alta precisão habilitado, BigDecimal
double	duplo
número inteiro	int
map*	java.util.Map
string	String
struct*	Classe JavaBean personalizada com getters e setters para os elementos do campo struct
text	String
* Com suporte no mecanismo Spark.	

Em Java, os tipos de dados java.util.List, java.util.Map, String, byte [] e BigDecimal são tipos de dados complexos. Os tipos de dados double, int e long são tipos de dados primitivos.

Na Developer tool, os tipos de dados array, struct e map são tipos de dados complexos.

A transformação Java configura valores nulos nos tipos de dados primitivos como zero. Você pode usar os métodos API isNull e setNull na guia **Na Entrada** para definir valores nulos na porta de entrada para valores nulos na porta de saída. Para ver um exemplo, consulte [“setNull” na página 345](#).

**Nota:** O tipo de dados decimal é mapeado para BigDecimal quando a alta precisão está ativada. O BigDecimal não pode ser usado com alguns operadores, como o operador +. Se o código Java contiver uma expressão que use uma porta decimal ou uma porta complexa com um elemento de um tipo de dados decimal e a porta for usada com um dos operadores, o código Java não conseguirá compilar.

## Conversão de tipo de dados complex no mecanismo Spark

Você pode adicionar portas complexas à transformação Java em mapeamentos executados no Mecanismo Spark. Uma porta complexa é uma porta que recebe um tipo de dados complex. A transformação Java converte tipos de dados complex em tipos de dados complex Java comparáveis. Ela também converte

elementos de tipos de dados complex em tipos de dados Java comparáveis, que são uma versão em caixa de tipos de dados primitivos.

### Tipo de Dados Array

Quando uma transformação Java lê linhas de entrada, ela converte o tipo de dados array em tipo de dados List Java. A transformação converte o tipo de dados de elementos de array em uma versão em caixa de tipos de dados Java.

Por exemplo, a transformação converte o tipo de dados complex da Developer tool `array<integer>` para um tipo de dados Java `List<Integer>`.

Quando uma transformação Java grava linhas de saída, ele converte o tipo de dados List Java em tipo de dados array. A transformação converte o tipo de dados de elementos de Lista em uma versão desconvertida dos tipos de dados da Developer tool.

### Tipo de Dados Struct

Quando uma transformação Java lê as linhas de entrada do tipo de dados struct, gera uma classe do bean Java equivalente. O nome da classe do bean Java é o mesmo que o nome do tipo de dados struct. A transformação converte o tipo de dados de elementos de struct em uma versão em caixa de tipos de dados Java.

Se o tipo de dados struct usar caracteres especiais ou palavras-chave reservadas do Java, como `final` ou particular, a transformação Java substituirá os caracteres especiais por um sublinhado (`_`) no nome da classe. Você pode abrir o código completo na visualização Java da guia de propriedades da transformação para ver as classes geradas.

A transformação Java gera nomes de campo do membro da classe prefixados com um sublinhado (`_`). Também gera getters e setters para os campos de membros.

A classe do bean Java gerada possui o namespace do nome da biblioteca de definição de tipos como uma classe externa. O nome da classe externa é o mesmo que o nome da biblioteca de definição de tipos. O nome do tipo de dados Java da porta de struct é do seguinte formato:

```
type_library_name.struct_type_name
```

Por exemplo, o nome da biblioteca de tipos é `m_Type_Definition_Library`. A definição de tipo de dados complex para a porta struct é:

```
Customer {
    name string
    age integer
}
```

A transformação Java gera as classes de beans Java com getters e setters para os campos de membros. O trecho de código a seguir mostra a classe externa e a classe interna:

```
public static final class m_Type_Definition_Library {
    public static final class Customer implements Serializable {
        private String _name;
        private Integer _age;

        public Customer() {}
    }
}
```

O trecho de código a seguir mostra o getter e setter para o nome do elemento struct do tipo string:

```
public String get_name() {
    return _name;
}

public void set_name(String _name) {
    this._name = _name;
}
```

Quando uma transformação Java grava linhas de saída do tipo de dados struct, ele converte a classe do bean Java em struct. A transformação converte os campos de membros da classe em struct. O tipo de

dados dos campos de membros é convertido em uma versão desconvertida dos tipos de dados da Developer tool.

### Tipo de Dados Map

Quando uma transformação Java lê linhas de entrada, ela converte o tipo de dados map em tipo de dados Map Java. A transformação converte o tipo de dados de elementos de mapa em uma versão em caixa de tipos de dados Java.

Por exemplo, o tipo de dados complex da Developer tool `map<string, bigint>` é convertido em um tipo de dados Java `Map<String, long>`.

Quando uma transformação Java grava linhas de saída, ela converte o tipo de dados Map Java em tipo de dados map. A transformação converte o tipo de dados de elementos de Map em uma versão desconvertida dos tipos de dados da Developer tool.

## Projetando uma Transformação Java

Ao projetar uma transformação Java, é necessário considerar fatores como o tipo de transformação que você deseja criar.

Ao projetar uma transformação Java, considere as seguintes questões:

- Você precisa criar uma transformação Java ativa ou passiva?

Uma transformação Java passiva gera uma única linha de saída para cada linha de entrada na transformação.

Uma transformação Java ativa gera várias linhas de saída para cada linha de entrada na transformação.

- Você precisa definir funções na transformação Java? Em caso afirmativo, quais expressões você deseja incluir em cada função?

Por exemplo, você pode definir uma função que invoque uma expressão para procurar os valores de portas de entrada ou saída, ou para procurar os valores de variáveis da transformação Java.

- Deseja criar uma transformação Java reutilizável ou não reutilizável?

Uma transformação reutilizável pode existir em vários mapeamentos.

Uma transformação não reutilizável pode existir dentro de um único mapeamento.

## Portas da Transformação Java

Uma transformação Java pode ter portas de entrada e saída.

Para criar e editar portas para uma transformação Java não reutilizável, use a guia **Portas** no editor. Para criar e editar portas para uma transformação Java reutilizável, use a exibição **Visão Geral** no editor.

Você pode especificar valores padrão para as portas. Após adicionar portas a uma transformação, você pode usar os nomes de porta como variáveis em trechos de código Java.

## Criando portas

Ao criar uma transformação Java, um grupo de entrada e um grupo de saída são incluídos.

Ao criar uma porta, a ferramenta Developer a adiciona abaixo da linha ou grupo selecionado.

## Definindo Valores Padrão para Portas

Você pode definir valores padrão para portas em uma transformação Java.

A transformação Java inicializa variáveis de porta com o valor padrão para a porta dependendo do tipo de dados da porta.

### Portas de Entrada e Saída

A transformação Java inicializa o valor das portas de entrada desconectadas ou das portas de saída que não são atribuídas a um valor nos trechos de código Java.

A transformação Java inicializa as portas com base nos seguintes tipos de dados Java:

#### Tipo de dados primitivo

Se você definir um valor padrão para a porta que não seja igual a nulo, a transformação inicializará o valor da variável de porta com o valor padrão. Caso contrário, ela inicializará o valor da variável de porta como 0.

#### Tipo de dados complexo

Se você definir um valor padrão para a porta, a transformação criará um novo objeto e o inicializará com o valor padrão. Caso contrário, a transformação inicializará a variável de porta para NULL. Por exemplo, se você definir um valor padrão para a porta string, a transformação criará um novo objeto String e o inicializará com o valor padrão.

**Nota:** Se você acessar uma variável de porta de entrada com um valor nulo no código Java, uma exceção `NullPointerException` ocorrerá.

É possível ativar uma porta de entrada como uma chave de partição e uma chave de classificação e também é possível atribuir uma direção de classificação. O Serviço de Integração de Dados particiona e classifica os dados em cada partição com base na chave e na direção de classificação. A Chave de Partição e a Chave de Classificação são válidas quando o escopo de transformação é definido como Todas as Entradas.

Use as seguintes propriedades para o particionamento e a classificação de dados:

#### Chave de Partição

Porta de entrada que determina as linhas de dados a serem agrupadas na mesma partição.

#### Chave de Classificação

Porta de entrada que determina os critérios de classificação em cada partição.

#### Direção

Ordem ascendente ou descendente. O padrão é decrescente.

## Transformação Java - Propriedades Avançadas

A transformação Java inclui propriedades avançadas para o código de transformação e a transformação.

É possível substituir as propriedades da transformação ao usá-la em um mapeamento.

É possível definir as seguintes propriedades avançadas para a transformação de Java na guia **Avançado**:

**Nível de Rastreamento**

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.

**Particionável**

A transformação pode ser processada com vários segmentos. Desmarque essa opção se quiser que o Serviço de Integração de Dados use um segmento para processar a transformação. O Serviço de Integração de Dados pode usar vários segmentos para processar as fases de pipeline de mapeamento restantes.

Desative o particionamento de uma transformação Java quando o código Java exigir que a transformação seja processada com um segmento.

**Ativar alta precisão**

Processa uma porta de tipo de dados decimal com uma precisão menor que ou igual a 38, como uma porta de tipo de dados Java BigDecimal.

Desative a alta precisão para processar uma porta de tipo de dados decimal como uma porta de tipo de dados Double Java.

A seguinte tabela mostra como uma transformação de Java trata um valor em uma porta de entrada de tipo de dados decimal, dependendo de você ter ativado ou desativado a opção de alta precisão:

Exemplo	Processamento de Alta Precisão Ativado	Processamento de Alta Precisão Desativado
Uma porta de entrada de tipo decimal recebe um valor de 40012030304957666903.	A transformação Java deixa o valor como está.	A transformação Java converte o valor para o seguinte: $4.00120303049577 \times 10^{19}$

Se a transformação Java contiver uma porta decimal ou uma porta complexa com um elemento de um tipo de dados decimal, a transformação deverá usar o mesmo modo de precisão que o mapeamento. Por exemplo, se você ativar a alta precisão na transformação Java, deverá ativar a alta precisão no mapeamento.

**Usar nanossegundos em Data/Hora**

Converte portas de tipo de dados date/time em portas de tipo de dados BigDecimal Java com precisão de nanossegundos.

Desative o processamento de nanossegundo para que o código Java gerado converta portas de tipo de dados date/time em portas de tipo de dados Long Java com precisão de milissegundos.

**Classpath**

Define o classpath para diretórios de arquivo de classe ou jar que são associados a pacotes Java não padrão importados na guia **Importações**.

Os diretórios de arquivo de classe ou jar deve estar acessíveis no computador cliente da Developer tool para compilação do código Java.

Com base no sistema operacional, separe as entradas de classpath da seguinte forma:

- No UNIX, use dois pontos para separar as entradas de classpath.
- No Windows, use ponto e vírgula para separar as entradas de classpath.



Por exemplo, se você importar o pacote conversor Java na guia **Importações** e definir o pacote no converter.jar, deverá adicionar a localização do arquivo converter.jar ao classpath antes de compilar o código Java para a transformação Java.

**Nota:** Você não precisa configurar o classpath para os pacotes Java incorporados. Por exemplo, como java.io é um pacote Java interno, você não precisa definir o classpath para o java.io.

#### **Está Ativo**

A transformação pode gerar mais de uma linha de saída para cada linha de entrada.

Não é possível alterar esta propriedade após criar a transformação Java. Se precisar alterar esta propriedade, crie uma nova transformação Java.

#### **Escopo de Transformação**

Define o método usado pelo Serviço de Integração de Dados para aplicar a lógica da transformação aos dados de entrada. Você pode escolher um destes valores:

- **Linha.** Aplica a lógica da transformação a uma linha de dados por vez. Escolha Linha quando os resultados do procedimento dependem de uma única linha de dados.
- **Transação.** Aplica a lógica da transformação a todas as linhas em uma transação. Escolha Transação quando os resultados do procedimento dependerem de todas as linhas na mesma transação, mas não de linhas em outras transações. Ao escolher Transação, você deve conectar todos os grupos de entrada ao mesmo ponto de controle de transação.
- **Todas as Entradas.** Aplica a lógica da transformação a todos os dados de entrada. quando você escolhe Todas as Entradas, o Serviço de Integração de Dados descarta limites de transação. Escolha Todas as Entradas se o resultado do procedimento depender de todas as linhas de dados da origem.

#### **Sem Estado**

Mantenha a ordem de linhas dos dados de entrada na transformação. Selecione essa opção se o Serviço de Integração de Dados não realizar qualquer otimização que possa alterar a ordem de linhas.

Quando o Serviço de Integração de Dados realiza as otimizações, ele pode perder uma ordem estabelecida anteriormente no mapeamento. Você pode estabelecer a ordem em um mapeamento com uma origem de arquivo simples classificada, uma origem relacional classificada ou uma transformação do Classificador. Quando você configura uma transformação para manter a ordem de linhas, o Serviço de Integração de Dados considera essa configuração quando ele executa as otimizações para o mapeamento. O Serviço de Integração de Dados realizará otimizações para a transformação se ela puder manter a ordem. O Serviço de Integração de Dados não realizará otimizações para a transformação se a otimização alterar a ordem de linha.

## **Configurando o Classpath para a ferramenta Developer do**

Você pode adicionar arquivos JAR ou diretórios de arquivo de classe ao classpath da ferramenta Developer do .

Para definir o classpath para a máquina em que a ferramenta Developer do é executada, conclua uma das seguintes tarefas:

- Configure a variável do ambiente CLASSPATH. Defina a variável de ambiente CLASSPATH na máquina cliente da ferramenta Developer do . Isto se aplica a todos os processos Java executados na máquina.
- Para uma transformação Java não reutilizável, onfigure o classpath nas das configuraçõesda transformação Java. Isso se aplica a sessões que incluem essa transformação Java. O cliente da Developer tool inclui arquivos dentro do classpath ao compilar o código Java.

Para adicionar arquivos JAR ou diretórios de arquivo de classe ao classpath em uma transformação Java, conclua as seguintes etapas:

1. Na guia **Avançadas**, clique no ícone de seta para baixo na coluna **Valor** ao lado de **Classpath**.  
A caixa de diálogo **Editar Classpath** é exibida.
2. Para adicionar um classpath, conclua as seguintes etapas:
  - a. Clique em **Adicionar**.  
A janela **Salvar Como** é exibida.
  - b. Na janela **Salvar Como**, navegue até o diretório em que o arquivo JAR está localizado.
  - c. Clique em **OK**.  
O classpath é exibido na caixa de diálogo **Editar Classpath**.
3. Para remover um arquivo JAR ou diretório de arquivo de classe, selecione o diretório JAR ou de arquivo de classe e clique **Remover**.  
O diretório desaparece da lista de diretórios.

## Configurando o Classpath para o Serviço de Integração de Dados

Você pode adicionar diretórios de arquivo de classe ou jar necessários no tempo de execução ao classpath no nó do Serviço de Integração de Dados.

Coloque os arquivos jar necessários durante o tempo de execução no seguinte diretório no nó do Serviço de Integração de Dados:

```
$INFA_HOME/services/shared/jars
```

Os arquivos jar nesse local são carregados dinamicamente. Todos os arquivos de classe exigidos por mapeamentos individuais no tempo de execução são encontrados e carregados a partir desse diretório.

**Nota:** A transformação Java adiciona os arquivos jar desse diretório ao classpath no nível do mapeamento.

## Desenvolvendo Código Java

Use as guias de entrada de código na exibição **Java** para gravar e compilar o código Java que define o comportamento da transformação para eventos de transformação específicos.

Você pode desenvolver trechos de código nas guias de entrada de código em qualquer ordem. É possível exibir (mas não editar) o código Java completo na guia **Código Completo**.

Após desenvolver trechos de código, você pode compilá-los ou o código Java completo e exibir os resultados da compilação na janela **Resultados**, nas propriedades da **Compilação** da exibição **Java**.

Cada guia de entrada de código contém componentes que você usa para gravar, exibir e compilar o código Java:

### Propriedades do código

Fornecer controles que permitem exibir e inserir código Java, incluindo os métodos de API da transformação Java. A seguinte tabela descreve os controles que estão disponíveis nas propriedades do **Código**:

Controle	Descrição
Navegador	<p>Mostra as portas de entrada, as portas de saída e os métodos de API da transformação Java chamáveis.</p> <p>Clique em um item no navegador para exibir uma descrição do item.</p> <p>Clique duas vezes em um item para adicioná-lo à janela <b>Código Java</b>. Como alternativa, você pode arrastar um item do navegador para a janela <b>Código Java</b>.</p> <p>O navegador está disponível nas seguintes guias de entrada de código:</p> <ul style="list-style-type: none"><li>- <b>Auxiliares</b></li><li>- <b>Na Entrada</b></li><li>- <b>No Final</b></li></ul>
Janela <b>Código Java</b>	<p>Permite exibir ou inserir código Java para a transformação. A janela <b>Código Java</b> exibe o código Java usando o realce básico de sintaxe Java.</p> <p><b>Nota:</b> Na guia <b>Código Completo</b>, é possível exibir (mas não editar) o código de classe completo para a transformação Java.</p> <p>A janela <b>Código Java</b> está disponível nas seguintes guias de entrada de código:</p> <ul style="list-style-type: none"><li>- <b>Importações</b></li><li>- <b>Auxiliares</b></li><li>- <b>Na Entrada</b></li><li>- <b>No Final</b></li><li>- <b>Funções</b></li><li>- <b>Interfaces do Otimizador</b></li><li>- <b>Código Completo</b></li></ul>
Comando <b>Nova Função</b>	<p>Abre a caixa de diálogo <b>Definir Função</b>, usada para definir funções que invocam expressões Java.</p> <p>O comando <b>Função</b> está disponível na guia <b>Funções</b>.</p>
Barra de ferramentas de edição	<p>Permite que você clique em ícones de ferramenta, como recortar, copiar e colar, para editar o código Java.</p> <p>A barra de ferramentas de edição está disponível nas seguintes guias de entrada de código:</p> <ul style="list-style-type: none"><li>- <b>Importações</b></li><li>- <b>Auxiliares</b></li><li>- <b>Na Entrada</b></li><li>- <b>No Final</b></li><li>- <b>Funções</b></li><li>- <b>Interfaces do Otimizador</b></li></ul>

### Propriedades de compilação

Fornece controles para compilar e depurar o código Java. A seguinte tabela descreve os controles nas propriedades da **Compilação**:

Controle	Descrição
Comando <b>Compilar</b>	Compila o código Java para a transformação.
Janela <b>Resultados</b>	Exibe os resultados da compilação para a classe de transformação Java e permite que você localize a origem de erros no código.  Para localizar um erro no código, clique com o botão direito do mouse em uma mensagem de erro na janela <b>Resultados</b> e selecione para exibir o erro no trecho de código ou no código completo.  Você também pode clicar duas vezes em uma mensagem de erro na janela <b>Resultados</b> para localizar sua origem.

## Criando trechos de código Java

Para criar trechos de código Java para definir o comportamento da transformação, use a janela **Código Java** na entrada na guia da entrada de código.

1. Clique na guia apropriada de entrada de código.

A tabela a seguir descreve as tarefas que você pode executar nas guias de entrada do código no modo de exibição **Java**:

Guia	Descrição
<b>Importações</b>	Importa pacotes Java personalizados e integrados de outras empresas para uma transformação Java passiva ou ativa. Depois de importar os pacotes, você pode usá-los nas outras guias de entrada de código.
<b>Auxiliares</b>	Declara variáveis e métodos definidos pelo usuário para a classe de transformação Java em uma transformação Java passiva ou ativa. Depois de declarar variáveis e métodos, você poderá usá-los em qualquer outra entrada de código, exceto o guia <b>Importações</b> .
<b>Na Entrada</b>	Define como uma transformação Java passiva ou ativa se comporta quando recebe uma linha de entrada. O código Java que você define nesta guia é executado uma vez para cada linha de entrada.  Nessa guia, você também pode acessar e usar os dados, as variáveis e os métodos API da transformação Java da porta de entrada e saída.
<b>No Final</b>	Define como uma transformação Java passiva ou ativa se comporta depois de processar todos os dados de entrada.  Nessa guia, você também pode definir os dados de saída para transformações ativas e chamar métodos API de transformação Java.

Guia	Descrição
<b>Funções</b>	Define as funções que invocam expressões em uma transformação Java com a linguagem de programação Java. Por exemplo, você pode definir uma função que invoca uma expressão que procura os valores das portas de entrada ou saída ou que procura os valores de variáveis da transformação Java. Na guia <b>Funções</b> , você pode definir manualmente as funções ou clicar em <b>Nova Função</b> para invocar a caixa de diálogo <b>Definir Função</b> , que permite definir facilmente uma função.
<b>Interfaces do Otimizador</b>	Define a otimização de filtro de seleção antecipada ou de push-into. Selecione o método de otimização no navegador. Atualize os trechos de código para ativar a otimização. Defina as portas de entrada e as portas de saída associadas para passar a lógica de filtro.
<b>Código Completo</b>	Somente leitura. Nessa guia, você pode exibir e compilar o código de classe completo para a transformação Java.

- Para acessar as variáveis da coluna de entrada ou saída no trecho, expanda a lista **Entrada** ou **Saída** no navegador e clique duas vezes no nome da porta.
- Para chamar uma API de transformação Java no trecho, expanda a lista **APIs que podem ser chamadas** no navegador e clique duas vezes no nome do método. Se necessário, configure os valores de entrada adequados para o método.
- Grave o código Java apropriado com base no tipo de guia da entrada de código do .  
Exiba o código de classe completo para a transformação Java na janela **Código Java** na guia **Código Completo**.

## Importando pacotes Java

Na guia **Importações**, você pode importar pacotes Java para transformações Java ativas ou passivas.

Você pode importar pacotes Java personalizados e integrados de outras empresas. Depois de importar pacotes Java, você poderá usar os pacotes importados nas outras guias de entrada de código.

**Nota:** Na guia **Importações**, você não pode declarar nem usar variáveis estáticas, variáveis de instância nem métodos de usuário.

Na ferramenta Developer do , quando você exporta ou importa metadados que contêm uma transformação Java, os arquivos JAR ou de classe que contêm os pacotes personalizados ou de terceiros exigidos pela transformação com Java não são incluídos na exportação ou importação.

Se você importar metadados que contêm uma transformação Java, será necessário copiar os arquivos JAR ou de classe que contêm os pacotes personalizados ou de terceiros para o nó Data Integration Service e cliente da ferramenta Developer.

Por exemplo, para importar o pacote Java de E/S, digite o seguinte código na guia **Importações**:

```
import java.io.*;
```

Ao importar pacotes Java não padrão, adicione o pacote ou classe ao classpath nas da transformação Java.

## Definindo Passiva

Na guia **Auxiliares**, você pode declarar variáveis e métodos definidos pelo usuário para a classe de transformação Java da transformação Java ativa ou passiva.

Depois de declarar variáveis e métodos na guia **Auxiliares**, você poderá usar as variáveis e os métodos em qualquer guia de entrada de código, exceto as guias **Importações**.

Na guia **Auxiliares**, você pode declarar os seguintes tipos de código, variáveis e métodos:

- Código estático e variáveis estáticas.

Em um bloco estático, você pode declarar variáveis estáticas e código estático. Todas as instâncias de uma transformação Java reutilizável em um mapeamento compartilham o código estático e as variáveis. O código estático é executado antes de qualquer outro código em uma transformação Java.

Por exemplo, o código a seguir declara uma variável estática para armazenar o limite de erro para todas as instâncias de uma transformação de Java em um mapeamento:

```
static int errorThreshold;
```

Use esta variável para armazenar o limite de erros para a transformação e acessá-la a partir de todas as instâncias da transformação Java em um mapeamento

**Nota:** É necessário sincronizar variáveis estáticas em uma transformação Java reutilizável.

- Variáveis de instância.

Várias instâncias de uma transformação Java reutilizável em um mapeamento não compartilham variáveis de instância. Declare as variáveis de instância com um prefixo para evitar conflitos e inicializar variáveis de instância não primitivas.

Por exemplo, o código a seguir utiliza uma variável booleana para declarar se uma linha de saída vai ou não ser gerada:

```
// boolean to decide whether to generate an output row
// based on validity of input
private boolean generateRow;
```

- Métodos de instância ou estáticos definidos pelo usuário.

Estende a funcionalidade da transformação Java. Os métodos Java declarados na guia **Auxiliares** podem usar ou modificar variáveis de saída ou variáveis de instância declaradas localmente. Você não pode acessar as variáveis de entrada a partir dos métodos Java na guia **Auxiliares**.

Por exemplo, use o seguinte código na guia **Auxiliares** para declarar uma função que adiciona dois inteiros:

```
private int myTXAdd (int num1,int num2)
{
    return num1+num2;
}
```

## Propriedades Java da Transformação Java

Use as guias de entrada de código na exibição **Java** para gravar e compilar o código Java que define o comportamento da transformação para eventos de transformação específicos.

As seguintes guias são guias de entrada de código:

- **Importações**
- **Auxiliares**
- **Na Entrada**
- **No Final**
- **Funções**
- **Interfaces do Otimizador**

Exiba o código de classe completo para a transformação Java na guia **Código Completo**.

## Guia Importações

Na guia **Importações**, você pode importar pacotes Java de terceiros, internos ou personalizados para transformações Java ativas ou passivas.

Para importar um pacote Java, insira o código para importar o pacote na janela **Código Java** das propriedades do **Código** na guia **Importações**.

Por exemplo, você pode inserir o seguinte código para importar o pacote java.io:

```
import java.io.*;
```

Para compilar o código que importa pacotes Java, clique em **Compilar** nas propriedades da **Compilação** na guia **Importações**. Os resultados da compilação aparecem na janela **Resultados** da guia **Importações**.

Após importar pacotes Java, você pode usá-los nas outras guias de entrada de código.

## Guia Auxiliares

Na guia **Auxiliares**, você pode declarar métodos e variáveis definidos pelo usuário para a classe de transformação Java em uma transformação Java ativa ou passiva.

Para declarar métodos e variáveis definidos pelo usuário, insira o código na janela **Código Java** nas propriedades do **Código** da guia **Auxiliares**.

Para compilar o código auxiliar para a transformação Java, clique em **Compilar** nas propriedades da **Compilação** da guia **Auxiliares**. Os resultados da compilação aparecem na janela **Resultados** da guia **Auxiliares**.

Após declarar variáveis e métodos, você poderá usá-los em qualquer outra guia de entrada de código, exceto em **Importações**.

## Guia Na Entrada

Na guia **Na Entrada**, você define como uma transformação Java ativa ou passiva se comporta quando recebe uma linha de entrada. Nessa guia, você também pode acessar e usar dados das portas de entrada e saída, variáveis e métodos de API da transformação Java.

O código Java que você define nesta guia é executado uma única vez para cada linha de entrada.

Para definir como uma transformação Java se comporta quando recebe uma linha de entrada, insira o código na janela **Código Java**, nas propriedades do **Código** da guia **Na Entrada**.

A partir do navegador na guia **Na Entrada**, você pode acessar e definir as seguintes variáveis e métodos de API:

- Variáveis de porta de entrada e de saída. Acesse os dados das portas de entrada e de saída como uma variável, usando o nome da porta como nome da variável. Por exemplo, se "in\_int" é uma porta de entrada de Inteiro, é possível acessar os dados dessa porta referindo-se como uma variável "in\_int" com o inteiro de tipo de dados primitivos Java. Não é necessário declarar as portas de entrada e de saída como variáveis.

Não atribua um valor a uma variável de porta de entrada. Se você atribuir um valor a uma variável de entrada na guia **Na Entrada**, não será possível obter os dados de entrada para a porta correspondente na linha atual.

- Variáveis de instância e métodos definidos pelo usuário. Use qualquer instância, variável estática ou método definido pelo usuário que você declarou na guia **Auxiliares**.

Por exemplo, uma transformação com Java tem duas portas de entrada, BASE\_SALARY e BONUSES, com um tipo de dados inteiro, e uma única porta de saída, TOTAL\_COMP, com um tipo de dados inteiro. Você

cria um método definido pelo usuário na guia **Auxiliares**, `myTXAdd`, que adiciona dois inteiros e retorna o resultado. Use o seguinte código Java na guia **Na Entrada** para atribuir os valores totais das portas de entrada à porta de saída e gerar uma linha de saída:

```
TOTAL_COMP = myTXAdd (BASE_SALARY, BONUSSES);  
generateRow();
```

Quando a transformação com Java recebe uma linha de entrada, ela adiciona o valor das portas de entrada `BASE_SALARY` e `BONUSSES`, atribui o valor à porta de saída `TOTAL_COMP` e gera uma linha de saída.

- Métodos de API de transformação Java. Você pode chamar métodos de API fornecidos pela transformação Java.

Para compilar o código para a transformação Java, clique em **Compilar** nas propriedades da **Compilação** da guia **Na Entrada**. Os resultados da compilação aparecem na janela **Resultados** da guia **Na Entrada**.

## Guia No Final

Na guia **No Final**, você define como uma transformação Java ativa ou passiva se comporta depois de processar todos os dados de entrada. Nessa guia, você também pode definir os dados de saída para transformações ativas e chamar métodos de API de transformação Java.

Para definir como uma transformação Java se comporta depois de processar todos os dados de entrada, insira o código na janela **Código Java** das propriedades do **Código** na guia **No Final**.

Você pode acessar e definir as seguintes variáveis e métodos de API na guia **No Final**:

- Variáveis da porta de saída. Você pode usar os nomes das portas de saída que definiu na guia **Portas** como variáveis ou definir dados de saída para transformações Java ativas.
- Variáveis de instância e métodos definidos pelo usuário. Use as variáveis de instância ou métodos definidos pelo usuário que você declarou na guia **Auxiliares**.
- Métodos de API de transformação Java. Chame métodos de API fornecidos pela transformação Java.

Por exemplo, use o seguinte código Java para gravar informações no log quando o final dos dados for atingido:

```
logInfo("Number of null rows for partition is: " + partCountNullRows);
```

Para compilar o código para a transformação Java, clique em **Compilar** nas propriedades da **Compilação** da guia **No Final**. Os resultados da compilação são exibidos na janela **Resultados** da guia **No Final**.

## Guia Funções

Na guia **Funções**, você define as funções que invocam expressões em uma transformação Java com a linguagem de programação Java.

Por exemplo, você pode definir uma função que invoca uma expressão que procura os valores das portas de entrada ou saída ou que procura os valores de variáveis da transformação Java.

Você pode definir manualmente as funções na janela **Código Java**, nas propriedades do **Código** na guia **Funções**, ou pode clicar em **Nova Função** para invocar a caixa de diálogo **Definir Função**, que permite definir facilmente uma função.

Para compilar o código, clique em **Compilar** nas propriedades da **Compilação** da guia **Funções**. Os resultados da compilação aparecem na janela **Resultados** da guia **Funções**.



## Guia Código Completo

Na guia **Código Completo**, é possível exibir (mas não editar) e compilar o código de classe completo para a transformação Java.

Você pode exibir o código de classe completo na janela **Código Java** das propriedades do **Código**.

Para compilar o código completo para a transformação Java, clique em **Compilar** nas propriedades da **Compilação** na guia **Código Completo**. Os resultados da compilação são exibidos na janela **Resultados**, na guia **Código Completo**.

## Otimização de Filtro com a Transformação Java

O Serviço de Integração de Dados pode aplicar otimização de filtro para transformações Java ativas. Ao definir a transformação Java, adicione o código para otimização de filtro na guia **Interfaces do Otimizador** da transformação Java.

## Otimização de Seleção Antecipada com a Transformação Java

Você pode habilitar uma transformação Java ativa ou passiva para otimização com seleção antecipada quando essa transformação não tem efeitos colaterais. O otimizador transmite a lógica de filtro através da transformação Java e modifica a condição de filtro conforme necessário.

Para exibir os trechos de código para otimização com seleção antecipada, escolha PredicatePushOptimization no navegador da guia **Interfaces do Otimizador**.

### allowPredicatePush

Booleano. Permite a seleção antecipada. Altere a função para retornar um resultado "true" e a mensagem para habilitar a seleção antecipada. O padrão é "false", e a função retorna uma mensagem informando que não há suporte para a otimização.

```
public ResultAndMessage allowPredicatePush(boolean ignoreOrderOfOp) {
    // To Enable PredicatePushOptimization, this function should return true
    //return new ResultAndMessage(true, "");
    return new ResultAndMessage(false, "Predicate Push Optimization Is Not
Supported");
}
```

### canGenerateOutputFieldEvalError

Booleano. Indica se a transformação Java pode retornar um erro de campo de saída, como um erro de divisão por zero. Altere a função para retornar "false" se a transformação Java não gerar erros de campos de saída. Quando a transformação Java pode gerar erros de campo, o Data Integration Service não pode usar a otimização com seleção antecipada.

```
public boolean canGenerateOutputFieldEvalError() {
    // If this Java transformation can never generate an output field evaluation error,
    // return false.
    return true;
}
```

## getInputExpr

Retorna uma expressão da Informatica que descreve quais valores de entrada de campos de entrada formam um campo de saída. O otimizador precisa saber quais campos de entrada formam um campo de saída para enviar a lógica de filtro pela transformação.

```
public InfaExpression getInputExpr(TransformationField field,
    TransformationDataInterface group) {
    // This should return an Informatica expression for output fields in terms of input
    fields
    // We will only push predicate that use fields for which input expressions are
    defined.
    // For example, if you have two input fields in0 and in1 and three output fields
    out0, out1, out2
    // out0 is the pass-through of in1, out2 is sum of in1 and in2, and out3 is unknown,
    the code should be:
    //if (field.getName().equals("out0"))
    //    return new InfaExpression("in0", instance);
    //else if (field.getName().equals("out1"))
    //    return new InfaExpression("in0 + in1", instance);
    //else if (field.getName().equals("out2"))
    //    return null;
    return null;
}
```

Por exemplo, um mapeamento contém uma expressão de filtro, "out0 > 8". Out0 é o valor da porta de saída out0 na transformação Java. Você pode definir o valor de out0 como o valor da porta de entrada in0 + 5. O otimizador pode enviar a seguinte expressão "(in0 + 5) > 8" após a transformação Java com otimização de seleção antecipada. Você poderá retornar NULL se um campo de saída não tiver uma expressão de campo. O otimizador não envia expressões de filtro após campos de saída sem uma expressão de entrada. Você pode incluir o seguinte código:

```
if (field.getName().equals("out0"))
    return new InfaExpression("in0 + 5", instance);
else if (field.getName().equals("out2"))
    return null;
```

## inputGroupsPushPredicateTo

Retorna uma lista de grupos que podem receber a lógica de filtro. A transformação Java tem um grupo de entrada. Não modifique essa função para a transformação Java.

```
public List<TransformationDataInterface> inputGroupsPushPredicateTo(
    List<TransformationField> fields) {
    // This functions returns a list of input data interfaces to push predicates to.
    // Since JavaTx only has one input data interface, you should not have to modify
    this function
    AbstractTransformation tx = instance.getTransformation();
    List<DataInterface> dis = tx.getDataInterfaces();
    List<TransformationDataInterface> inputDIs = new
    ArrayList<TransformationDataInterface>();
    for (DataInterface di : dis){
        TransformationDataInterface tdi = (TransformationDataInterface) di;
        if (tdi.isInput())
            inputDIs.add(tdi);
    }
    if(inputDIs.size() == 1)
        return inputDIs;
    else
        return null;
}
```

## Otimização de Envio com a Transformação Java

Você poderá habilitar uma transformação Java ativa para otimização de envio se ela não tiver efeitos colaterais e se a otimização não afetar os resultados do mapeamento.

Ao configurar a otimização de envio para a transformação Java, você define uma forma para essa transformação Java armazenar a condição de filtro que ela recebe do otimizador. Adicionar um código que examine a condição de filtro. Se a transformação Java puder absorver a lógica de filtro, a transformação Java transmitirá uma condição "true" de volta ao otimizador. O otimizador remove a transformação de Filtro do mapeamento otimizado.

Ao configurar a transformação Java, você grava o código que armazena a condição de filtro como metadados de transformação durante a otimização. Você também grava o código para recuperar a condição de filtro em tempo de execução e para descartar as linhas de acordo com a lógica de filtro.

Ao definir a transformação Java, você adiciona o código para otimização de envio na guia **Interfaces do Otimizador** da transformação Java. Para acessar os trechos de código para otimização de envio, escolha FilterPushdownOptimization no navegador da guia **Interfaces do Otimizador** da transformação.

A ferramenta Developer exibe trechos de código para habilitar a otimização de envio e para receber a condição de filtro do otimizador. Atualize os trechos de código para habilitar a otimização e para salvar a lógica de filtro como metadados de transformação.

### isFilterSupported

Retorna "true" para habilitar a otimização de envio. Retorna "false" para desabilitar a otimização de envio. Altere a função para retornar "true" de forma a habilitar a otimização de envio.

```
public ResultAndMessage isFilterSupported() {
    // To enable filter push-into optimization this function should return true
    // return new ResultAndMessage(true, "");
    return new ResultAndMessage(false, "Filter push-into optimization is not supported");
}
```

### pushFilter

Recebe a condição de filtro do otimizador.

Adicione um código para examinar o filtro e determinar se a lógica de filtro pode ser usada na transformação. Se a transformação puder absorver o filtro, use o método a seguir para armazenar a condição de filtro como metadados de transformação:

```
storeMetadata(String key, String data)
```

A chave é um identificador para os metadados. Você pode definir qualquer string como uma chave. Os dados são os dados que você deseja armazenar para determinar quais linhas devem ser descartadas em tempo de execução. Por exemplo, os dados podem ser a condição de filtro que a transformação Java recebe do otimizador.

```
public ResultAndMessage pushFilter(InfaExpression condition) {
    // Add code to absorb the filter
    // If filter is successfully absorbed return new ResultAndMessage(true, ""); and the
    optimizer
    // will remove the filter from the mapping
    // If the filter is not absorbed, return new ResultAndMessage(false, msg);
    return new ResultAndMessage(false, "Filter push-into optimization is not supported");
}
```

# Criando uma Transformação Java

Na ferramenta Developer, você pode criar uma transformação Java reutilizável ou não reutilizável.

## Criando uma Transformação Java Reutilizável

As transformações reutilizáveis podem existir dentro de vários mapeamentos.

Crie uma transformação Java reutilizável na ferramenta Developer.

1. Selecione um projeto ou pasta na exibição **Object Explorer**.
2. Clique em **Arquivo > Novo > Transformação**.  
A caixa de diálogo **Novo** é exibida.
3. Selecione a transformação Java.
4. Clique em **Avançar**.
5. Insira um nome para a transformação.
6. Para criar uma transformação ativa, selecione a opção **Criar como ativo**.
7. Clique em **Concluir**.  
A transformação aparece no editor.
8. Na exibição **Portas**, clique no botão **Novo** para adicionar uma porta à transformação.
9. Edite a porta para definir o nome, o tipo de dados e a precisão.  
Utilize nomes de porta como variáveis nos trechos de código Java.
10. Na exibição **Java**, use as guias de entrada de código para gravar e compilar o código Java para a transformação.
11. Na exibição **Java**, use a guia **Funções** para definir as funções que invocam expressões.
12. Em qualquer guia de entrada de código, clique duas vezes nas mensagens de erro da janela **Resultados** nas propriedades da **Compilação** para localizar e corrigir erros de compilação no código Java para a transformação.
13. Na exibição **Avançado**, edite as propriedades da transformação.

## Criando uma Transformação Java Não Reutilizável

Transformações não reutilizáveis existem dentro de um único mapeamento.

Crie uma transformação Java não reutilizável na ferramenta Developer.

1. Em um mapeamento ou mapplet, arraste uma transformação Java da paleta de Transformação para o editor.
2. Na caixa de diálogo **Nova Transformação Java**, insira um nome para a transformação.
3. Para criar uma transformação ativa, selecione a opção **Criar como ativo**.
4. Clique em **Concluir**.  
A transformação aparece no editor.
5. Na guia **Geral**, edite o nome e a descrição da transformação.
6. Na guia **Portas**, clique no botão **Novo** para adicionar uma porta à transformação.
7. Edite a porta para definir o nome, o tipo de dados e a precisão.  
Utilize nomes de porta como variáveis nos trechos de código Java.

8. Na exibição **Java**, use as guias de entrada de código para gravar e compilar o código Java para a transformação.
9. Na exibição **Java**, use a guia **Funções** para definir as funções que invocam expressões.
10. Em qualquer guia de entrada de código, clique duas vezes nas mensagens de erro da janela **Resultados** nas propriedades da **Compilação** para localizar e corrigir erros de compilação no código Java para a transformação.
11. Na exibição **Avançado**, edite as propriedades da transformação.

## Compilando uma transformação Java

A ferramenta Developer do usa o compilador Java para compilar o código Java e gerar o código de bytes para a transformação.

O compilador Java compila o código Java e exibe os resultados da compilação na janela **Resultados** nas propriedades **Compilação** nas guias de entrada de código. O compilador Java é instalado com a ferramenta Developer do no diretório `java/bin`.

Para compilar o código completo para a transformação Java, clique em **Compilar** nas propriedades **Compilação** da guia **Código Completo**.

Quando você cria uma transformação Java, ela contém uma classe Java que define a funcionalidade básica para uma transformação Java. O código completo da classe Java contém o código da classe de modelo para a transformação, além do código Java que você define nas guias de entrada de código.

Quando você compila uma transformação Java, a ferramenta Developer do adiciona o código das guias de entrada de código à classe do modelo para a transformação a fim de gerar o código de classe completo para a transformação. Em seguida, a ferramenta Developer do chama o compilador Java para compilar o código de classe completo. O compilador Java compila a transformação e gera o código de bytes para ela.

Os resultados da compilação são mostrados na janela **Resultados**. Use os resultados da compilação para identificar e localizar erros de código Java.

## Solucionando Problemas de uma Transformação Java

Na janela **Resultados** das propriedades da **Compilação** em qualquer guia de entrada de código, é possível localizar e corrigir erros de código Java.

Erros em uma transformação Java podem ocorrer devido a um erro no código em uma guia de entrada de código ou no código completo da classe de transformação Java.

Para solucionar problemas de uma transformação Java, conclua as seguintes etapas gerais:

1. Localize a origem do erro no trecho de código Java ou no código de classe completo da transformação.
2. Identifique o tipo de erro. Use os resultados da compilação na janela **Resultados** e a localização do erro para identificar o tipo de erro.
3. Corrija o código Java na guia de entrada de código.
4. Compile a transformação novamente.

## Localizando a Origem de Erros de Compilação

Para localizar a origem de erros de compilação, use os resultados da compilação exibidos na janela **Resultados** das propriedades da **Compilação** em uma guia de entrada de código ou na guia **Código Completo**.

Quando você clica duas vezes em uma mensagem de erro na janela **Resultados**, o código de origem que causou o erro é realçado na janela **Código Java** na guia de entrada de código ou na guia **Código Completo**.

É possível localizar os erros na guia **Código Completo**, mas não é possível editar o código Java na guia **Código Completo**. Para corrigir os erros encontrados na guia **Código Completo**, modifique o código na guia de entrada de código apropriada. Pode ser necessário usar a guia **Código Completo** para exibir os erros provocados pela adição de códigos de usuário ao código de classe completo da transformação.

### Encontrando um Erro em uma Guia de Entrada de Código ou na Guia Código Completo

Você pode encontrar erros de compilação em uma guia de entrada de código ou na guia **Código Completo**.

1. Na janela **Resultados** das propriedades da **Compilação** em qualquer guia de entrada de código ou na guia **Código Completo**, clique com o botão direito do mouse em uma mensagem de erro.
2. Clique em **Mostrar em > Trecho** ou **Mostrar em > Guia Código Completo**.

A ferramenta Developer realça a origem do erro na guia selecionada.

**Nota:** É possível exibir, mas não corrigir os erros na guia **Código Completo**. Para corrigir erros, você deve navegar até a guia de entrada de código apropriada.

## Identificando a origem de erros de compilação

Os erros de compilação podem surgir como resultado de erros no código do usuário.

Erros no código do usuário também podem gerar um erro no código de não usuário da classe. Erros de compilação ocorrem no código do usuário e não usuário da transformação Java.

### Erros de Código do Usuário

Podem ocorrer erros no código do usuário nas guias de entrada de código. Os erros de código do usuário incluem erros de idioma e sintaxe Java padrão.

Também podem ocorrer erros de código do usuário quando a ferramenta Developer do do adiciona o código do usuário das guias de entrada de código ao código de classe completo.

Por exemplo, uma transformação Java tem uma porta de entrada com o nome `int1` e um tipo de dados inteiro. O código completo da classe declara a variável da porta de entrada com este código:

```
int int1;
```

No entanto, se você usar o mesmo nome da variável na guia **Na Entrada**, o compilador Java emitirá um erro de redeclaração de uma variável. Para corrigir o erro, renomeie a variável na guia **Na Entrada**.

## Erros de código de não usuário

O código do usuário nas guias de entrada de código pode causar erros no código de não usuário.

Por exemplo, uma transformação Java tem uma porta de entrada e uma de saída, int1 e out1, com tipos de dados inteiro. Grave o seguinte código na guia de entrada de código **Na Entrada** para calcular os juros da porta de entrada int1 e atribuí-los à porta de saída out1:

```
int interest;
interest = CallInterest(int1); // calculate interest
out1 = int1 + interest;
}
```

Ao compilar a transformação, a do Cliente do PowerCenter adiciona o código da guia de entrada de código **Na Entrada** ao código de classe completo da transformação. Quando o compilador Java compila o código Java, a chave incompatível causa a terminação precoce de um método no código de classe completo e o compilador Java emite um erro.

## Convertendo para exemplo de dados struct

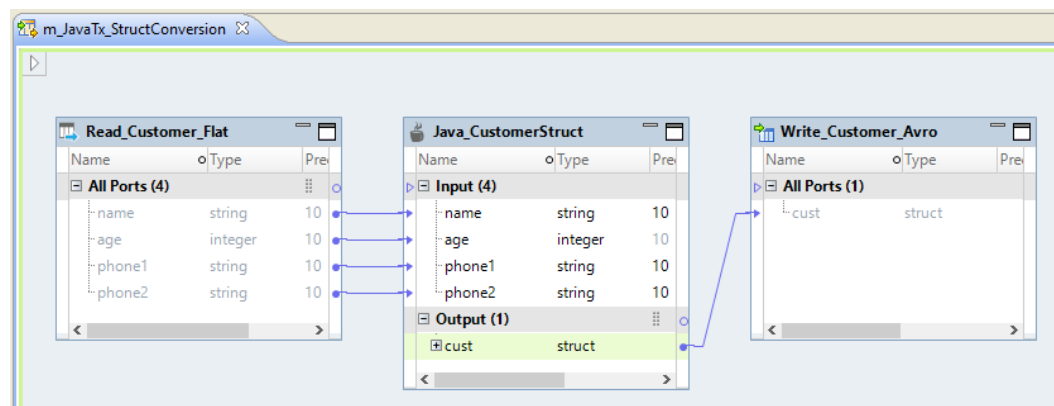
Sua organização precisa converter um grande volume de dados do cliente em um arquivo simples em dados struct e gravá-los em um arquivo Avro. O arquivo de entrada contém detalhes do cliente, como nome, idade e números de telefone. Se o nome do cliente for nulo no arquivo de entrada, você não deseja incluir detalhes do cliente no arquivo de saída.

Você pode desenvolver um mapeamento com uma transformação Java para definir a funcionalidade de transformação. No ambiente do Hadoop, execute o mapeamento no mecanismo Spark para transformar os dados e gravar os dados struct em um arquivo Avro.

Crie um mapeamento e configure as seguintes transformações:

- Transformação de Leitura que lê informações do cliente de uma origem de arquivo simples
- Transformação Java como uma transformação ativa que converte dados planos em dados struct e remove dados inconsistentes
- Transformação de Gravação que grava os dados struct em um arquivo Avro

A imagem a seguir mostra o mapeamento com uma transformação de Leitura, uma transformação Java e uma transformação de Gravação.

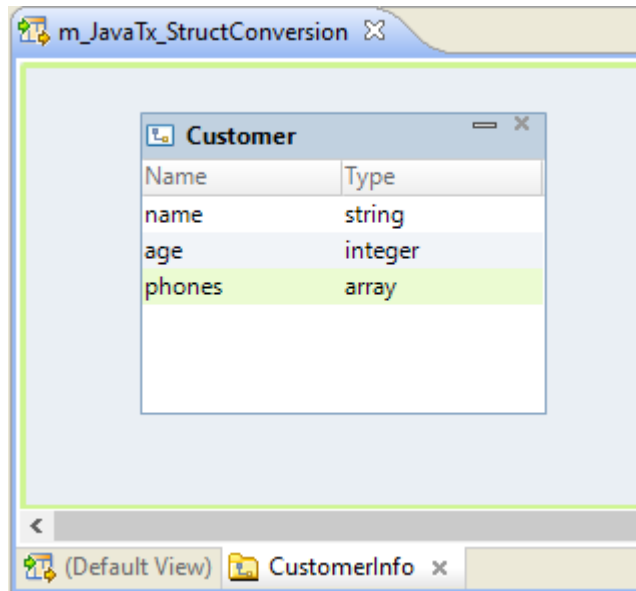


Na guia da biblioteca de definição de tipo do editor de mapeamento, crie uma definição de tipo de dados complex de Cliente. A definição de tipo de dados complex representa o esquema dos dados struct.

Renomeie a biblioteca de definição de tipo para CustomerInfo. Adicione os seguintes elementos à definição do tipo de dados complex:

- nome do tipo de string
- idade do tipo integer
- telefones do tipo array com elementos de string

A imagem a seguir mostra a definição de tipo de dados complex na biblioteca de definição de tipos:

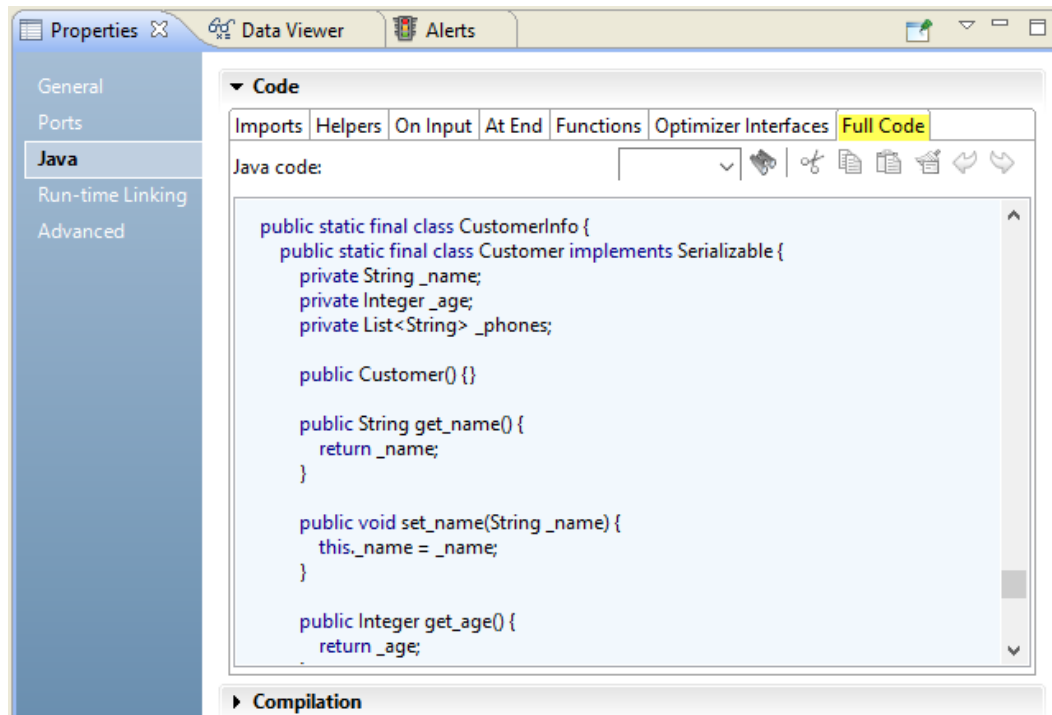


Na transformação Java, inclua uma porta de saída struct e especifique a configuração de tipo da porta para referenciar a definição de tipo de dados complex que você criou. A transformação Java gera uma classe Cliente com setters e getters para ler e definir os campos de membros. A classe contém os seguintes campos de membros:

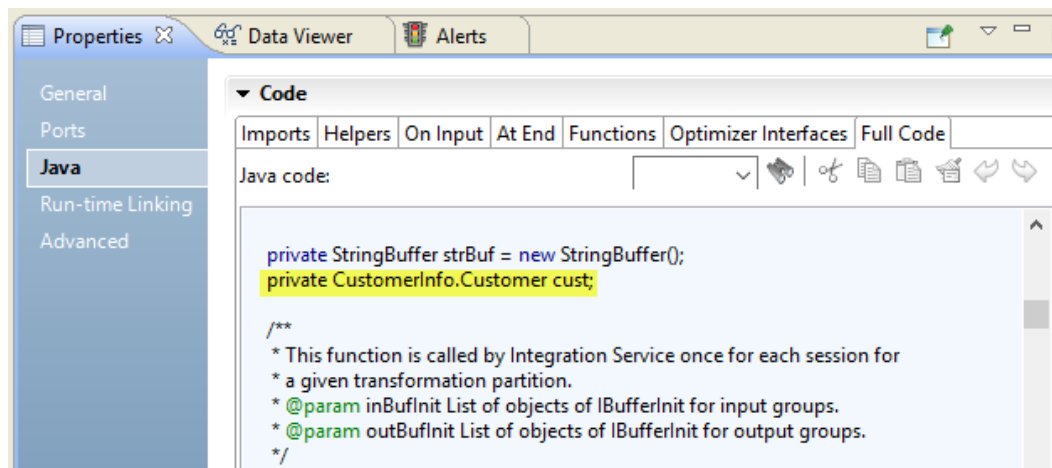
- \_name
- \_age
- \_phones



A imagem a seguir mostra a classe criada para a porta struct na guia **Código Completo** da exibição **Java**:

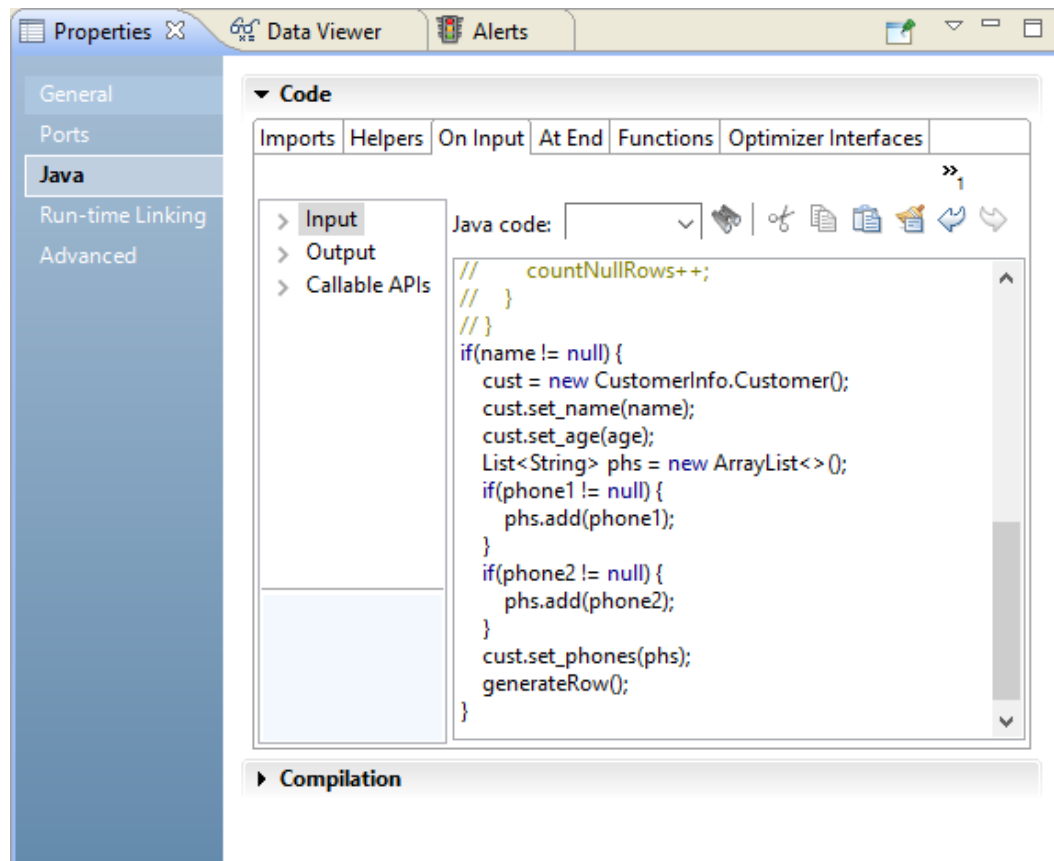


O tipo de dados Java para a porta de struct usa o nome da biblioteca de definição de tipos e a definição de tipo de dados complex. A imagem a seguir mostra o nome do tipo de dados Java `CustomerInfo.Customer` para o campo `cust` no código gerado:



Na exibição **Java** da transformação Java, importe quaisquer pacotes Java de terceiros, incorporados ou personalizados que a transformação requiera. Escreva e compile o código Java para converter os dados planos em dados struct e para remover a linha do cliente, se o nome do cliente for nulo.

A imagem a seguir mostra o código na guia **Na Entrada**:



Valide o mapeamento e execute o mapeamento no mecanismo Spark para gravar os dados transformados na saída do arquivo Avro.

## Transformação Java em um ambiente não nativo

O processamento da transformação Java em um ambiente não nativo depende do mecanismo que executa a transformação.

Considere o suporte para os seguintes mecanismos de tempo de execução não nativos:

- Mecanismo Blaze. Suportado com restrições.
- Mecanismo Spark. Com suporte com restrições em mapeamentos de fluxo e em lote.
- Mecanismo Databricks Spark. Sem suporte.

## Transformação Java no mecanismo Blaze

Para usar arquivos .jar externos em uma transformação Java, execute as seguintes etapas:

1. Copie os arquivos .jar externos para o diretório de instalação do Informatica na máquina do Serviço de Integração de Dados no seguinte local: <diretório de instalação do Informatica>/services/shared/jars. Em seguida, recicle o Serviço de Integração de Dados.

2. Na máquina que hospeda a Developer tool, onde você desenvolve e executa o mapeamento que contém a transformação Java:
  - a. Copie os arquivos .jar externos para um diretório na máquina local.
  - b. Edite a transformação Java para incluir uma instrução de importação apontando para os arquivos .jar locais.
  - c. Atualize o classpath na transformação Java.
  - d. Compile a transformação.

Algumas regras de processamento para o mecanismo Blaze são diferentes das regras de processamento do Serviço de Integração de Dados.

### Particionamento

- As seguintes restrições se aplicam à propriedade Escopo da Transformação:
  - O valor de Transação para o escopo da transformação não é válido.
  - Se você ativar uma porta de entrada para a chave de partição, o escopo da transformação deverá ser definido como Todas as Entradas.
  - Sem Estado deve ser ativado se o escopo da transformação for linha.

## Transformação Java no mecanismo Spark

Algumas regras de processamento para o mecanismo Spark são diferentes das regras de processamento do Serviço de Integração de Dados.

### Restrições gerais

A transformação Java é suportada com as seguintes restrições no mecanismo Spark:

- O código Java na transformação não pode gravar a saída na saída padrão quando você envia a lógica de transformação para o Hadoop. O código Java pode gravar a saída no erro padrão que aparece nos arquivos de log.
- Para valores de data/hora, o mecanismo Spark suporta a precisão de até microssegundos. Se um valor de data/hora contiver nanossegundos, os dígitos à direita serão truncados.

### Particionamento

A transformação Java possui as seguintes restrições quando usada com particionamento:

- A propriedade Particionável deve estar ativada na transformação Java. A transformação não pode ser executada em uma partição.
- As seguintes restrições se aplicam à propriedade Escopo da Transformação:
  - O valor de Transação para o escopo da transformação não é válido.
  - Se você ativar uma porta de entrada para a chave de partição, o escopo da transformação deverá ser definido como Todas as Entradas.
  - Sem Estado deve ser ativado se o escopo da transformação for linha.

### Validação de Mapeamento

A validação do mapeamento falha nas seguintes situações:

- Você faz referência a uma transformação de Pesquisa não conectada de uma expressão dentro de uma transformação Java.
- Você seleciona uma porta de um tipo de dados complex como a partição ou a chave de classificação.

- Você ativa o processamento de nanossegundos em data/hora e a transformação Java contém uma porta de tipo de dados complex com um elemento de um tipo de data/hora. Por exemplo, uma porta de tipo `array<data/time>` não é válida se você ativar o processamento de nanossegundos em data/hora.

O mapeamento falha na seguinte situação:

- A transformação Java e o mapeamento usam diferentes modos de precisão quando a transformação Java contém uma porta decimal ou uma porta complexa com um elemento de um tipo de dados decimal.

Mesmo que a alta precisão esteja ativada no mapeamento, o mapeamento processa os dados no modo de baixa precisão em algumas situações, como quando o mapeamento contém uma porta complexa com um elemento de um tipo de dados decimal ou o mapeamento é um mapeamento de fluxo. Se a alta precisão estiver ativada na transformação Java e no mapeamento, mas o mapeamento processar dados no modo de baixa precisão, o mapeamento falhará.

### Usando arquivos .jar externos

Para usar arquivos .jar externos em uma transformação Java, execute as seguintes etapas:

1. Copie os arquivos .jar externos para o diretório de instalação do Informatica na máquina do Serviço de Integração de Dados no seguinte local:  
`<Diretório de instalação do Informatica>/services/shared/jars`
2. Recicle o Serviço de Integração de Dados.
3. Na máquina que hospeda a Developer tool, onde você desenvolve e executa o mapeamento que contém a transformação Java:
  - a. Copie os arquivos .jar externos para um diretório na máquina local.
  - b. Edite a transformação Java para incluir uma instrução de importação apontando para os arquivos .jar locais.
  - c. Atualize o classpath na transformação Java.
  - d. Compile a transformação.

### Transformação de Java em um mapeamento de fluxo

Os mapeamentos de fluxo têm regras de processamento adicionais que não se aplicam a mapeamentos em lote.

O Serviço de Integração de Dados ignora as seguintes propriedades:

- Particionável. O Serviço de Integração de Dados ignora a propriedade e pode processar a transformação com vários segmentos.
- Sem estado. O Serviço de Integração de Dados ignora a propriedade e mantém a ordem das linhas dos dados de entrada.

## CAPÍTULO 20

# Referência da API da transformação Java

Este capítulo inclui os seguintes tópicos:

- [Visão geral dos métodos API da transformação Java, 337](#)
- [defineJExpression, 338](#)
- [failSession, 339](#)
- [generateRow, 339](#)
- [getInRowType, 340](#)
- [getMetadata, 341](#)
- [incrementErrorCount, 341](#)
- [invokeJExpression, 342](#)
- [isNull, 343](#)
- [logError, 344](#)
- [logInfo, 344](#)
- [resetNotification, 345](#)
- [setNull, 345](#)
- [storeMetadata, 346](#)

## Visão geral dos métodos API da transformação Java

Nas guias de entrada de código do modo de exibição **Javano** Editor, você pode adicionar métodos API ao código Java para definir o comportamento da transformação.

Para adicionar um método API ao código, expanda a lista **APIs que podem ser chamadas** no navegador da guia de entrada de código e clique duas vezes no nome do método que você deseja adicionar ao código.

Como alternativa, você pode arrastar o método do navegador até o trecho de código Java ou inserir manualmente o método API no trecho de código Java.

Você pode adicionar os seguintes métodos API ao código Java em uma transformação Java:

### **defineJExpression**

Define uma expressão Java.

**failSession**

Lança uma exceção com uma mensagem de erro e apresenta falha no da sessão.

**generateRow**

Gera uma linha de saída para transformações Java ativas.

**getInRowType**

Retorna o tipo de entrada da linha atual na transformação.

**incrementErrorCount**

Incrementa a contagem de erros do da sessão.

**invokeJExpression**

Invoca uma expressão Java que você definiu usando o método `defineJExpression`.

**isNull**

Verifica se há um valor nulo em uma coluna de entrada.

**logError**

Grava uma mensagem de erro no log da .

**logInfo**

Grava uma mensagem com informações no log da .

**resetNotification**

Se a máquina do Data Integration Service estiver sendo executada no modo de reinicialização, redefine as variáveis que você usa no código Java depois de executar um mapeamento.

**setNull**

Define como nulo o valor de uma coluna de saída em uma transformação Java ativa ou passiva.

## defineJExpression

Define uma expressão, incluindo a string da expressão e os parâmetros de entrada. Argumentos para o método `defineJExpression` incluem uma matriz de objetos `JExprParamMetadata` que contém os parâmetros de entrada e um valor de string que define a sintaxe de expressão.

Utilize a seguinte sintaxe:

```
defineJExpression(  
    String expression,  
    Object[] paramMetadataArray  
);
```

A tabela a seguir descreve os parâmetros:

Parâmetro	Tipo	Tipo de dados	Descrição
expressão	Entrada	String	String que representa a expressão.
paramMetadataArray	Entrada	Objeto	Matriz de objetos <code>JExprParamMetadata</code> que contém os parâmetros de entrada da expressão.

Você pode adicionar o método `defineJExpression` ao código Java em qualquer guia de entrada de código, exceto as guias **Importações** e **Funções**.

Para usar o método `defineJExpression`, é necessário representar uma matriz de objetos `JExprParamMetadata` que representam os parâmetros de entrada da expressão. Defina os valores de metadados para os parâmetros e transfira a matriz como um parâmetro para o método de `defineJExpression`.

Por exemplo, o seguinte código Java cria uma expressão para consultar o valor de duas strings:

```
JExprParamMetadata params[] = new JExprParamMetadata[2];
params[0] = new JExprParamMetadata(EDataType.STRING, 20, 0);
params[1] = new JExprParamMetadata(EDataType.STRING, 20, 0);
defineJExpression(":lkp.mylookup(x1,x2)",params);
```

**Nota:** Você deve numerar os parâmetros transferidos para a expressão consecutivamente e iniciar os parâmetros com a letra x. Por exemplo, para transferir três parâmetros para uma expressão, nomeie os parâmetros x1, x2 e x3.

## failSession

Lança uma exceção com uma mensagem de erro e apresenta falha noda sessão.

Utilize a seguinte sintaxe:

```
failSession(String errorMessage);
```

A tabela a seguir descreve o parâmetro:

Parâmetro	Tipo de Parâmetro	Tipo de dados	Descrição
errorMessage	Entrada	String	String de mensagem de erro.

Use o método `failSession` para finalizar o mapeamento da . Não use o método `failSession` em um bloco `try/catch` em uma guia de entrada de código.

Você pode adicionar o método `failSession` ao código Java em qualquer guia de entrada de código, exceto as guias **Importações** e **Funções**.

O código Java a seguir mostra como testar a porta de entrada `input1` para um valor nulo e como encerrar o da sessão se for nulo:

```
if(isNull("input1")) {
    failSession("Cannot process a null value for port input1.");
}
```

## generateRow

Gera uma linha de saída para transformações Java ativas.

Utilize a seguinte sintaxe:

```
generateRow();
```

Ao chamar o método `generateRow`, a transformação Java gera uma linha de saída usando o valor atual das variáveis da porta de saída. Se desejar gerar linhas múltiplas que correspondam a uma linha de entrada, você poderá chamar o método `generateRow` mais de uma vez para cada linha de entrada. Se você não usar o método `generateRow` em uma transformação Java ativa, a transformação não gerará linhas de saída.

Você pode adicionar o método `generateRow` ao código Java em qualquer guia de entrada de código, exceto as guias **Importações e Funções**.

Você pode chamar o método `generateRow` somente em transformações ativas. Se você chamar o método `generateRow` em uma transformação passiva, a Data Integration Service gerará um erro.

Use o seguinte código Java para gerar uma linha de saída, modificar os valores das portas de saída e gerar outra linha de saída:

```
// Generate multiple rows.
if(!isNull("input1") && !isNull("input2"))
{
    output1 = input1 + input2;
    output2 = input1 - input2;
}
generateRow();
// Generate another row with modified values.
output1 = output1 * 2;
output2 = output2 * 2;
generateRow();
```

## getInRowType

Retorna o tipo de entrada da linha atual na transformação. O método retorna um valor de inserção, atualização, exclusão ou rejeição.

Utilize a seguinte sintaxe:

```
rowType getInRowType();
```

A tabela a seguir descreve o parâmetro:

Parâmetro	Tipo de Parâmetro	Tipo de dados	Descrição
rowType	Saída	String	Retorna o tipo de estratégia de atualização, que é um dos seguintes valores: <ul style="list-style-type: none"><li>- DELETE</li><li>- INSERT</li><li>- REJECT</li><li>- UPDATE</li></ul>

Você pode adicionar o método `getInRowType` ao código Java na guia **Na entrada** de entrada de código.

Você só pode utilizar o método `getInRowType` em transformações ativas configuradas para definir a estratégia de atualização. Se você chamar esse método em uma transformação ativa que não está configurada para definir a estratégia de atualização, a Data Integration Service gerará um erro.



## getMetadata

Recupera os metadados de transformação de Java no tempo de execução. O método `getMetadata` recupera os metadados que você salva com o método `storeMetadata`, tais como uma condição de filtro que o otimizador passa para a transformação de Java na função `pushFilter`.

Utilize a seguinte sintaxe:

```
getMetadata (String key);
```

A tabela a seguir descreve os parâmetros:

Parâmetro	Tipo de Parâmetro	Tipo de dados	Descrição
chave	Entrada	String	Identifica os metadados. O método <code>getMetadata</code> usa a chave para determinar quais metadados recuperar.

Você pode adicionar o método `getMetadata` ao código Java nas seguintes guias de entrada de código:

- Auxiliar
- Na Entrada
- No Final
- Interfaces do Otimizador
- Funções

Você pode configurar o método `getMetadata` para recuperar as condições de filtro para otimização de push-into. O método `getMetadata` pode recuperar cada condição de filtro que você armazenar no otimizador.

```
// Retrieve a filter condition  
String mydata = getMetadata ("FilterKey");
```

## incrementErrorCount

Incrementa a contagem de erros . Se a contagem de erros atingir o limite de erros, o da sessão apresentará falha.

Utilize a seguinte sintaxe:

```
incrementErrorCount(int nErrors);
```

A seguinte tabela descreve o parâmetro:

Parâmetro	Tipo de Parâmetro	Tipo de dados	Descrição
nErrors	Entrada	Número inteiro	Número para incrementar a contagem de erros .

Você pode adicionar o método `incrementErrorCount` para o código Java em qualquer guia de entrada de código, exceto as guias **Importações** e **Funções**.

O seguinte código Java mostra como incrementar a contagem de erros caso uma porta de entrada de uma transformação tenha um valor nulo:

```
// Check if input employee id and name is null.
if (isNull ("EMP_ID_INP") || isNull ("EMP_NAME_INP"))
{
    incrementErrorCount(1);
    // if input employee id and/or name is null, don't generate a output row for this
    input row
    generateRow = false;
}
```

## invokeJExpression

Invoca uma expressão e retorna o valor para a expressão.

Utilize a seguinte sintaxe:

```
(datatype)invokeJExpression(
    String expression,
    Object[] paramMetadataArray);
```

Os parâmetros de entrada para o método invokeJExpression são um valor de string que representa a expressão e uma matriz de objetos que contém os parâmetros de entrada da expressão.

A tabela a seguir descreve os parâmetros:

Parâmetro	Tipo de Parâmetro	Tipo de dados	Descrição
expressão	Entrada	String	String que representa a expressão.
paramMetadataArray	Entrada	Objeto[]	Matriz de objetos que contém os parâmetros de entrada da expressão.

Você pode adicionar o método de invokeJExpression ao código Java em qualquer guia de entrada de código, exceto as guias **Importações** e **Funções**.

Use as seguintes regras e diretrizes ao usar o método invokeJExpression:

- Tipo de dados de retorno. O tipo de dados de retorno do método invokeJExpression é um objeto. É necessário moldar o valor de retorno da função com o tipo de dados apropriado.  
É possível retornar valores com os tipos de dados Inteiro, Double, String e byte[].

- Tipo de linha. O tipo de linha para valores de retorno do método invokeJExpression é INSERT.  
Para usar um tipo de linha diferente para o valor de retorno, use a interface avançada.

- Valores nulos. Se você transferir um valor nulo como um parâmetro ou se o valor de retorno para o método invokeJExpression for NULL, o valor será tratado como um indicador nulo.

Por exemplo, se o valor de retorno de uma expressão for NULL e o tipo de dados retornado for String, uma string será retornada com um valor de nulo.

- Tipo de dados Data. Você deve converter os parâmetros de entrada com um tipo de dados Data para o tipo de dados String.

Para usar a string em uma expressão como tipo de dados Data, utilize a função `to_date()` para converter a string nesse tipo de dados.

Além disso, é necessário moldar o tipo de retorno de qualquer expressão que retorne um tipo de dados Data como uma String.

O exemplo a seguir concatena strings " John " e " Smith " e retorna a string " Paulo Silva ":

```
(String)invokeJExpression("concat(x1,x2)", new Object [] { "John ", "Smith" });
```

**Nota:** Você deve numerar os parâmetros que você transfere para a expressão consecutivamente e iniciar o parâmetro com a letra x. Por exemplo, para transferir três parâmetros para uma expressão, nomeie os parâmetros x1, x2 e x3.

## isNull

Verifica o valor de uma coluna de entrada para um valor nulo.

Utilize a seguinte sintaxe:

```
Boolean isNull(String satrColName);
```

A tabela a seguir descreve o parâmetro:

Parâmetros	Tipo de Parâmetro	Tipo de dados	Descrição
strColName	Entrada	String	Nome de uma coluna de entrada.

Você pode adicionar o método `isNull` ao código Java na guia de entrada de código **Na Entrada** .

O código Java a seguir mostra como verificar se o valor da coluna de entrada SALARY é NULL antes de adicioná-lo à variável de instância `totalSalaries`:

```
// if value of SALARY is not null
if (!isNull("SALARY")) {
    // add to totalSalaries
    TOTAL_SALARIES += SALARY;
}
```

Como alternativa, use o seguinte código Java para obter os mesmos resultados:

```
// if value of SALARY is not null
String strColName = "SALARY";
if (!isNull(strColName)) {
    // add to totalSalaries
    TOTAL_SALARIES += SALARY;
}
```

# logError

Grava uma mensagem de erro no log da .

Utilize a seguinte sintaxe:

```
logError(String msg);
```

A tabela a seguir descreve o parâmetro:

Parâmetro	Tipo de Parâmetro	Tipo de dados	Descrição
msg	Entrada	String	String de mensagem de erro.

Você pode adicionar o método `logError` ao código Java em qualquer guia de entrada de código, exceto as guias **Importações** e **Funções**.

O código Java a seguir mostra como registrar um erro quando a porta de entrada for nula:

```
// check BASE_SALARY
if (isNull("BASE_SALARY")) {
    logError("Cannot process a null salary field.");
}
```

Quando o código é executado, a seguinte mensagem é exibida no log da :

```
[JTX_1013] [ERROR] Cannot process a null salary field.
```

# logInfo

Grava uma mensagem com informações no log da .

Utilize a seguinte sintaxe:

```
logInfo(String msg);
```

A tabela a seguir descreve o parâmetro:

Parâmetro	Tipo de Parâmetro	Tipo de dados	Descrição
msg	Entrada	String	String de mensagem de informações.

Você pode adicionar o método `logInfo` ao código Java em qualquer guia de entrada de código, exceto as guias **Importações** e **Funções**.

O seguinte código Java mostra como gravar uma mensagem no log da depois que a transformação Java processar um limite de 1000 linhas para a mensagem:

```
if (numRowsProcessed == messageThreshold) {
    logInfo("Processed " + messageThreshold + " rows.");
}
```

# resetNotification

Se a máquina do Data Integration Service estiver sendo executada no modo de reinicialização, redefina as variáveis que você usa no código Java depois de executar um mapeamento.

No modo de reinicialização, o Serviço de Integração de Dados não está desligado, mas é redefinido após uma solicitação, de forma que possa processar a próxima solicitação.

Para uma transformação Java, use o método `resetNotification` para redefinir as variáveis no código Java depois de executar um mapeamento.

Utilize a seguinte sintaxe:

```
public int resetNotification(IGroup group) {  
    return EStatus.value;  
}
```

A tabela a seguir descreve os parâmetros:

Parâmetro	Tipo de Parâmetro	Tipo de dados	Descrição
inteiro	Saída	Estatus. <i>Valor</i>	Valor de retorno, onde <i>valor</i> é um dos seguintes valores: <ul style="list-style-type: none"><li>- SUCESSO Sucesso.</li><li>- FAKHA Falha.</li><li>- NOIMPL. Não implementado.</li></ul>
grupo	Entrada	IGroup	O grupo de entrada.

Você pode adicionar o método `resetNotification` ao código Java na guia de entrada de código na guia **Ajuda**.

O método `resetNotification` não aparece na lista de APIs de chamada (callable).

Por exemplo, supondo que o código Java declara uma variável estática chamada `out5_satatic` e a inicializa como 1, o seguinte código Java redefine a variável `out5_static` como 1, depois da próxima execução de mapeamento:

```
public int resetNotification(IGroup group) {  
    out5_static=1;  
    return EStatus.SUCCESS;  
}
```

Esse método não é necessário. No entanto, se o Data Integration Service é executado no modo de reinicialização e o mapeamento contém uma transformação Java que não implementa o método `resetNotification`, a mensagem de aviso JSDK\_42075 é exibida no log.

# setNull

Define o valor de uma coluna de saída como nulo em uma transformação Java passiva ou ativa.

Utilize a seguinte sintaxe:

```
setNull(String strColName);
```

A tabela a seguir descreve o parâmetro:

Parâmetro	Tipo de Parâmetro	Tipo de dados	Descrição
strColName	Entrada	String	Nome de uma coluna de saída.

O método `setNull` define como nulo o valor de uma coluna de saída em uma transformação Java passiva ou ativa. Depois de definir uma coluna de saída como nula, você não poderá modificar o valor até gerar uma linha de saída.

Você pode adicionar o método `setNull` ao código Java em qualquer guia de entrada de código, exceto as guias **Importações** e **Funções**.

O seguinte código Java mostra como verificar o valor de uma coluna de entrada e define o valor correspondente de uma coluna de saída como nulo:

```
// check value of Q3RESULTS input column
if(isNull("Q3RESULTS")) {
    // set the value of output column to null
    setNull("RESULTS");
}
```

Como alternativa, você pode usar o seguinte código Java para obter os mesmos resultados:

```
// check value of Q3RESULTS input column
String strColName = "Q3RESULTS";
if(isNull(strColName)) {
    // set the value of output column to null
    setNull(strColName);
}
```

## storeMetadata

Armazena os metadados de transformação de Java que você pode recuperar no tempo de execução com o método `getMetadata`.

Utilize a seguinte sintaxe:

```
storeMetadata (String key String data);
```

A tabela a seguir descreve os parâmetros:

Parâmetro	Tipo de Parâmetro	Tipo de dados	Descrição
chave	Entrada	String	Identifica os metadados. O método <code>storeMetadata</code> requer uma chave para identificar os metadados. Defina a chave como qualquer string.
dados	Entrada	String	Os dados que você deseja armazenar como metadados de transformação de Java.

Você pode adicionar o método `storeMetadata` ao código Java para as seguintes guias de entrada de código:

- Auxiliar
- Na Entrada

- No Final
- Interfaces do Otimizador
- Funções

Você pode configurar o método `storeMetadata` em uma transformação ativa para aceitar as condições de filtro para otimização push-into. O método `storeMetadata` armazena uma condição de filtro que o otimizador passa do mapeamento para a transformação de Java.

```
// Store a filter condition  
storeMetadata ("FilterKey", condition);
```

## CAPÍTULO 21

# Expressões Java

Este capítulo inclui os seguintes tópicos:

- [Visão Geral das Expressões Java, 348](#)
- [Usando a caixa de diálogo Definir Função para definir uma expressão, 349](#)
- [Trabalhando com a Interface Simples, 351](#)
- [Trabalhando com a Interface Avançada, 353](#)
- [Referência de API da classe JExpression, 357](#)

## Visão Geral das Expressões Java

Você pode invocar expressões do em uma transformação Java com a linguagem de programação Java.

Use as expressões para estender a funcionalidade de uma transformação Java. Por exemplo, você pode invocar uma expressão em uma transformação Java para procurar os valores das portas de entrada ou saída ou procurar os valores das variáveis da transformação Java.

Para invocar expressões em uma transformação Java, você deve gerar o código Java ou usar os métodos APIs da transformação Java para invocar a expressão. Você invoca a expressão e usa o resultado da mesma na guia de entrada de código apropriada. Você pode gerar o código Java que invoca uma expressão ou usar métodos de API para escrevê-lo.

A tabela a seguir descreve os métodos que você pode usar para criar e invocar expressões em uma transformação Java:

Método	Descrição
Caixa de diálogo <b>Definir Função</b>	Permite criar uma função que invoca uma expressão e gerar o código de uma expressão.
Interface simples	Permite chamar um único método API para invocar uma expressão e obter o resultado da expressão.
Interface avançada	Permite definir a expressão, invocar a expressão e usar o resultado da expressão. Se você estiver familiarizado com a programação orientada a objeto e quiser ter mais controle ao invocar a expressão, use a interface avançada.



## Tipos de Função de Expressão

Você pode criar expressões para uma transformação Java usando a caixa de diálogo **Definir Função** ou usando a interface simples ou avançada.

Você pode inserir expressões que usam variáveis de porta de entrada ou saída, ou variáveis no código Java como parâmetros de entrada.

Se você usar a caixa de diálogo **Definir Função**, poderá validar a expressão antes de usá-la em uma transformação Java.

Você pode invocar os seguintes tipos de funções de expressão em uma transformação Java:

Tipo de função de expressão	Descrição
Funções de linguagem de transformação	Funções do tipo SQL criadas para manipular expressões comuns.
Funções definidas pelo usuário	Funções que você cria na ferramenta Developer do com base em funções de linguagem de transformação.
Funções personalizadas.	Funções criadas por você com a API de Função Personalizada.

Você também pode usar transformações desconectadas e variáveis internas em expressões. Por exemplo, você pode usar uma transformação de pesquisa desconectada em uma expressão.

## Usando a caixa de diálogo Definir Função para definir uma expressão

Ao definir uma expressão Java, você configura a função, cria a expressão e gera o código que chama a expressão.

Você pode definir a função e criar a expressão na caixa de diálogo **Definir Função**.

Para criar uma função de expressão e usar a expressão em uma transformação Java, execute as tarefas de alto nível abaixo:

1. Configure a função que invoca a expressão, incluindo função, nome, descrição e parâmetros. Quando cria a expressão, você utiliza os parâmetros da função.
2. Crie a sintaxe de expressão e valide a expressão.
3. Gere o código Java que invoca a expressão.

O Developer coloca o código na guia de entrada de código **Funções**.

Depois de gerar o código Java, chame a função gerada na guia de entrada de código apropriada para chamar uma expressão ou obter um objeto JExpression, dependendo da interface que você usar, a simples ou a avançada.

**Nota:** Para validar uma expressão ao criá-la, use a caixa de diálogo **Definir Função**.

## Etapa 1. Configurar a Função

Você configura parâmetros de entrada, descrição e nome para a função Java que invoca a expressão.

Use as seguintes regras e diretrizes ao configurar a função:

- Use um nome de função exclusivo que não entre em conflito com funções Java na transformação, nem com palavras-chave Java reservadas.
- Você deve configurar tipo de dados, precisão, escala e nome do parâmetro. Os parâmetros de entrada são os valores passados quando você chama a função no código Java para a transformação.
- Para transferir um tipo de dados Data para uma expressão, use um tipo de dados String para o parâmetro de entrada.

Se uma expressão retornar um tipo de dados Data, você poderá usar o valor de retorno como tipo de dados String na interface simples e tipo de dados String ou Longo na interface avançada.

## Etapa 2. Criar e Validar a Expressão

Ao criar a expressão, utilize os parâmetros que você configurou para a função.

Você também pode usar funções de linguagem da transformação, funções personalizadas ou outras funções definidas pelo usuário na expressão. Você pode criar e validar a expressão nas caixas de diálogo **Definir Função**.

## Etapa 3. Gerar o Código Java para a Expressão

Depois de configurar a função e os parâmetros dela e definir e validar a expressão, você poderá gerar o código Java que invoca a expressão.

O Developer coloca o código Java gerado na guia de entrada de código **Funções**. Use o código Java gerado para chamar as funções que invocam a expressão nas guias de entrada de código. Você pode gerar o código Java simples ou avançado.

Depois de gerar o código Java que invoca uma expressão, você não poderá editar nem revalidar a expressão. Para modificar uma expressão depois de gerar o código, você precisa recriar a expressão.

## Criando uma expressão e gerando código Java usando a caixa de diálogo Definir Função.

Você pode criar uma função que invoca uma expressão na caixa de diálogo **Definir Função**.

Conclua as seguintes etapas para criar uma função que invoca uma expressão:

1. No Developer, abra uma transformação Java ou crie uma nova.
2. Na guia **Código Java**, clique em **Nova Função**.  
A caixa de diálogo **Definir Função** é exibida.
3. Digite um nome de função.
4. Se quiser, insira uma descrição para a expressão.  
Digite até 2.000 caracteres.
5. Crie os argumentos de para a função.  
Ao criar os de parâmetros, configure o nome, o tipo de dado, a precisão e a escala do de parâmetro.
6. Na guia **Expressão**, crie uma expressão com os argumentos que você criou.
7. Para validar a expressão, clique em **Validar**.

8. Como alternativa, insira a expressão na caixa de diálogo **Expressão**. Em seguida, clique em **Validar** para validar a expressão.
9. Para gerar o código Java usando a interface avançada, selecione a opção **Gerar Código Avançado**. Em seguida, clique em **Gerar**.  
O Developer gera a função para invocar a expressão na guia **Funções** da entrada de código.

## Modelos da expressão Java

Você pode gerar o código Java para uma expressão usando o código Java simples ou avançado para uma expressão.

O código Java da expressão é gerado com base no modelo da expressão.

O exemplo a seguir mostra o modelo de uma expressão Java gerado para código Java simples:

```
Object function_name (Java datatype x1[,
                        Java datatype x2 ...] )
                        throws SDK Exception
{
    return (Object)invokeJExpression( String expression,
                                     new Object [] { x1[, x2, ... ]} );
}
```

O exemplo a seguir mostra o modelo de uma expressão Java gerado com a interface avançada:

```
JExpression function_name () throws SDKException
{
    JExprParamMetadata params[] = new JExprParamMetadata[number of parameters];
    params[0] = new JExprParamMetadata (
        EDataType.STRING, // data type
        20, // precision
        0 // scale
    );
    ...
    params[number of parameters - 1] = new JExprParamMetadata (
        EDataType.STRING, // data type
        20, // precision
        0 // scale
    );
    ...
    return defineJExpression(String expression,params);
}
```

## Trabalhando com a Interface Simples

Utilize o método Java API `invokeJExpression` para invocar uma expressão na interface simples.

### invokeJExpression

Invoca uma expressão e retorna o valor para a expressão.

Utilize a seguinte sintaxe:

```
(datatype)invokeJExpression(
    String expression,
    Object[] paramMetadataArray);
```

Os parâmetros de entrada para o método `invokeJExpression` são um valor de cadeia de caracteres que representa a expressão e uma matriz de objetos que contém os parâmetros de entrada da expressão.

A tabela a seguir descreve os parâmetros:

Parâmetro	Tipo de Parâmetro	Tipo de dados	Descrição
expressão	Entrada	String	String que representa a expressão.
paramMetadataArray	Entrada	Objeto[]	Matriz de objetos que contém os parâmetros de entrada da expressão.

Você pode adicionar o método de `invokeJExpression` ao código Java em qualquer guia de entrada de código, exceto as guias **Importações** e **Funções**.

Use as seguintes regras e diretrizes ao usar o método `invokeJExpression`:

- Tipo de dados de retorno. O tipo de dados de retorno do método `invokeJExpression` é um objeto. É necessário moldar o valor de retorno da função com o tipo de dados apropriado.  
É possível retornar valores com os tipos de dados Inteiro, Double, String e byte[].
- Tipo de linha. O tipo de linha para valores de retorno do método `invokeJExpression` é INSERT.  
Para usar um tipo de linha diferente para o valor de retorno, use a interface avançada.
- Valores nulos. Se você transferir um valor nulo como um parâmetro ou se o valor de retorno para o método `invokeJExpression` for NULL, o valor será tratado como um indicador nulo.  
Por exemplo, se o valor de retorno de uma expressão for NULL e o tipo de dados retornado for String, uma string será retornada com um valor de nulo.
- Tipo de dados Data. Você deve converter os parâmetros de entrada com um tipo de dados Data para o tipo de dados String.  
Para usar a string em uma expressão como tipo de dados Data, utilize a função `to_date()` para converter a string nesse tipo de dados.  
Além disso, é necessário moldar o tipo de retorno de qualquer expressão que retorne um tipo de dados Data como uma String.

**Nota:** Você deve numerar os parâmetros que você transfere para a expressão consecutivamente e iniciar o parâmetro com a letra x. Por exemplo, para transferir três parâmetros para uma expressão, nomeie os parâmetros x1, x2 e x3.

## Exemplo de Interface Simples

Você pode definir e chamar expressões que usam o método API `invokeJExpression` nas guias de entrada de código **Auxiliares** e **Na Entrada**.

O exemplo a seguir mostra como concluir uma pesquisa nas portas de entrada NAME e ADDRESS em uma transformação Java e atribuir o valor de retorno à porta de saída COMPANY\_NAME.

Digite o seguinte código na guia de entrada de código **Na Entrada**:

```
COMPANY_NAME = (String)invokeJExpression(":lkp.my_lookup(X1,X2)", new Object []  
{str1 ,str2} );  
generateRow();
```

# Trabalhando com a Interface Avançada

Na interface avançada, você pode usar métodos API orientados a objeto para definir, chamar e obter o resultado de uma expressão.

A tabela a seguir descreve as classes e métodos API que estão disponíveis na interface avançada:

Classe ou método API	Descrição
Classe EDataType	Enumera os tipos de dados de uma expressão.
Classe JExprParamMetadata.	Contém os metadados de cada parâmetro em uma expressão. Os metadados de parâmetro incluem o tipo de dados, a precisão e a escala.
Método API defineJExpression	Define a expressão Inclui a string e os parâmetros da expressão do .
Método API invokeJExpression	Chama uma expressão.
Classe JExpression	Contém os métodos para criar, chamar, obter os metadados e o resultado da expressão e verificar o tipo de dados de retorno.

## Invocando uma expressão com a Interface Avançada

Você pode definir, chamar e obter o resultado de uma expressão usando a interface avançada.

1. Na guia **Auxiliares** ou **Na Entrada** de entrada de código, crie uma instância de Classe JExprParamMetadata para cada de parâmetro da expressão e defina o valor dos metadados. Como alternativa, você pode instanciar o objeto JExprParamMetadata no método defineJExpression.
2. Use o método defineJExpression para obter o objeto JExpression da expressão.
3. Na guia de entrada de código apropriada, invoque a expressão com o método invokeJExpression.
4. Verifique o resultado do valor de retorno com o método isResultNull.
5. Você pode obter o tipo de dados ou os metadados do valor de retorno com os métodos getResultDataType e getResultMetadata.
6. Obtenha o resultado da expressão usando o método API adequado. Você pode usar os métodos getInt, getDouble, getStringBuffer e getBytes.

## Regras e Diretrizes para Trabalhar com a Interface Avançada

Ao trabalhar com a interface avançada, você deve estar ciente das regras e diretrizes.

Use as seguintes regras e diretrizes:

- Se você transferir um valor nulo como um parâmetro ou o resultado de uma expressão for nulo, o valor será tratado como um indicador nulo. Por exemplo, se o resultado de uma expressão for nulo e o tipo de dados de retorno for String, será retornada uma string com um valor nulo. Você pode verificar o resultado de uma expressão usando o método isResultNull.

- Você precisa converter os parâmetros de entrada com um tipo de dados Data para um tipo String antes de usá-los em uma expressão. Para usar a string em uma expressão como tipo de dados Data, utilize a função `to_date()` para converter a string nesse tipo de dados.

Você pode obter o resultado de uma expressão que retorna um tipo de dados Data como String ou Longo.

Para obter o resultado de uma expressão que retorna um tipo de dados Data como String, use o método `getStringBuffer`. Para obter o resultado de uma expressão que retorne um tipo de dados Data como Longo, use o método `getLong`.

## Classe EDataType

Enumera os tipos de dados Java usados nas expressões. Obtém o tipo de dados de retorno de uma expressão ou atribui o tipo de dados para um parâmetro em um objeto `JExprParamMetadata`. Não é necessário instanciar a classe `EDataType`.

A tabela a seguir mostra os valores enumerados para os tipos de dados Java nas expressões:

Tipo de dados	Valor enumerado
INT	1
DOUBLE	2
STRING	3
BYTE_ARRAY	4
DATE_AS_LONG	5

O exemplo de código Java a seguir mostra como usar a classe `EDataType` para atribuir um tipo de dados String a um objeto `JExprParamMetadata`:

```
JExprParamMetadata params[] = new JExprParamMetadata[2];
params[0] = new JExprParamMetadata (
    EDataType.STRING, // data type
    20, // precision
    0 // scale
);
...
```

## Classe JExprParamMetadata

Instancia um objeto que representa os parâmetros de uma expressão e define os metadados para os parâmetros.

Use uma matriz de objetos `JExprParamMetadata` como entrada para o método `defineJExpression` a fim de definir os metadados para os parâmetros de entrada. Você pode criar uma instância do objeto `JExprParamMetadata` na guia de entrada do código **Funções** ou em `defineJExpression`.

Utilize a seguinte sintaxe:

```
JExprParamMetadata paramMetadataArray[] = new JExprParamMetadata[numberOfParameters];
paramMetadataArray[0] = new JExprParamMetadata(datatype, precision, scale);
...
paramMetadataArray[numberOfParameters - 1] = new JExprParamMetadata(datatype, precision,
scale);;
```

A tabela a seguir descreve os argumentos:

Argumento	Tipo de Argumento	Tipo de Dados do Argumento	Descrição
tipo de dados	Entrada	EDataType	Tipo de dados do parâmetro.
precisão	Entrada	Número inteiro	Precisão do parâmetro.
escala	Entrada	Número inteiro	Escala do parâmetro.

Por exemplo, use o segundo código Java para instanciar uma matriz de dois objetos JExprParamMetadata com tipos de dados String, precisão 20 e escala 0:

```
JExprParamMetadata params[] = new JExprParamMetadata[2];
params[0] = new JExprParamMetadata(EDataType.STRING, 20, 0);
params[1] = new JExprParamMetadata(EDataType.STRING, 20, 0);
return defineJExpression(":LKP.LKP_addresslookup(X1,X2)",params);
```

## defineJExpression

Define uma expressão, incluindo a cadeia de caracteres da expressão e os parâmetros de entrada. Argumentos para o método defineJExpression incluem uma matriz de objetos JExprParamMetadata que contém os parâmetros de entrada e um valor de cadeia de caracteres que define a sintaxe de expressão.

Utilize a seguinte sintaxe:

```
defineJExpression(
    String expression,
    Object[] paramMetadataArray
);
```

A tabela a seguir descreve os parâmetros:

Parâmetro	Tipo	Tipo de dados	Descrição
expressão	Entrada	String	String que representa a expressão.
paramMetadataArray	Entrada	Objeto	Matriz de objetos JExprParamMetadata que contém os parâmetros de entrada da expressão.

Você pode adicionar o método defineJExpression ao código Java em qualquer guia de entrada de código, exceto as guias **Importações** e **Funções**.

Para usar o método defineJExpression, é necessário representar uma matriz de objetos JExprParamMetadata que representam os parâmetros de entrada da expressão. Defina os valores de metadados para os parâmetros e transfira a matriz como um parâmetro para o método de defineJExpression.

Por exemplo, o seguinte código Java cria uma expressão para consultar o valor de duas strings:

```
JExprParamMetadata params[] = new JExprParamMetadata[2];
params[0] = new JExprParamMetadata(EDataType.STRING, 20, 0);
params[1] = new JExprParamMetadata(EDataType.STRING, 20, 0);
defineJExpression(":lkp.mylookup(x1,x2)",params);
```

**Nota:** Você deve numerar os parâmetros transferidos para a expressão consecutivamente e iniciar os parâmetros com a letra x. Por exemplo, para transferir três parâmetros para uma expressão, nomeie os parâmetros x1, x2 e x3.

## Classe JExpression

Contém métodos para criar e invocar uma expressão, retornar o valor de uma expressão e verificar o tipo de dados de retorno.

A tabela a seguir lista os métodos na classe JExpression:

Nome do Método	Descrição
invoke	Chama uma expressão.
getResultDataType	Retorna o tipo de dados do resultado da expressão.
getResultMetadata	Retorna metadados do resultado da expressão.
isResultNull	Verifica o valor de resultado de um resultado de expressão.
getInt	Retorna o valor de um resultado de expressão como um tipo de dados Inteiro.
getDouble	Retorna o valor de um resultado de expressão como um tipo de dados Double.
getStringBuffer	Retorna o valor de um resultado de expressão como um tipo de dados de String.
getBytes	Retorna o valor de um resultado de expressão como um tipo de dados byte[].

## Exemplo de Interface Avançada

Você pode usar a interface avançada para criar e invocar uma expressão de pesquisa em uma transformação Java.

O exemplo de código Java a seguir mostra como criar uma função que chama uma expressão e como invocar a expressão para obter o valor de retorno. Este exemplo transfere os valores de duas portas de entrada com um tipo de dados String, NAME e COMPANY à função myLookup. A função myLookup usa uma expressão de pesquisa para pesquisar o valor da porta de saída ADDRESS.

**Nota:** Este exemplo parte do pressuposto de que você tem uma transformação de pesquisa não conectada no mapeamento chamada LKP\_addresslookup.

Use o seguinte código Java na guia **Auxiliares**:

```
JExpression addressLookup() throws SDKException
{
    JExprParamMetadata params[] = new JExprParamMetadata[2];
    params[0] = new JExprParamMetadata (
        EDataType.STRING,          // data type
        50,                        // precision
        0                           // scale
    );
    params[1] = new JExprParamMetadata (
        EDataType.STRING,          // data type
        50,                        // precision
        0                           // scale
    );
    return defineJExpression(":LKP.LKP_addresslookup(X1,X2)",params);
}
JExpression lookup = null;
boolean isJExprObjCreated = false;
```



Use o seguinte código Java na guia **Na Entrada** para invocar a expressão e retornar o valor da porta ADDRESS:

```
...
if(!iisJExprObjCreated)
{
    lookup = addressLookup();
    isJExprObjCreated = true;
}
lookup = addressLookup();
lookup.invoke(new Object [] {NAME,COMPANY}, ERowType.INSERT);
EDatatype addressDataType = lookup.getResultDataType();
if(addressDataType == EDatatype.STRING)
{
    ADDRESS = (lookup.getStringBuffer()).toString();
} else {
    logError("Expression result datatype is incorrect.");
}
...
```

## Referência de API da classe JExpression

A classe JExpression contém os métodos API que permitem a você criar e invocar uma expressão, retornar o valor de uma expressão e verificar o tipo de dados de retorno.

A classe JExpression contém os seguintes métodos de API:

- `getBytes`
- `getDouble`
- `getInt`
- `getLong`
- `getResultDataType`
- `getResultMetadata`
- `getStringBuffer`
- `invoke`
- `isResultNull`

### getBytes

Retorna o valor de um resultado de expressão como um tipo de dados `byte[]`. Obtém o resultado de uma expressão que criptografa os dados com a função `AES_ENCRYPT`.

Utilize a seguinte sintaxe:

```
objectName.getBytes();
```

Use o seguinte exemplo de código Java para obter o resultado de uma expressão que criptografa os dados binários usando a função `AES_ENCRYPT`, onde `JExprEncryptData` é um objeto `JExpression`:

```
byte[] newBytes = JExprEncryptData.getBytes();
```

## getDouble

Retorna o valor de um resultado de expressão como um tipo de dados Double.

Utilize a seguinte sintaxe:

```
objectName.getDouble();
```

Use o seguinte exemplo de código Java para obter o resultado de uma expressão que retorna um valor de salário como um double, onde JExprSalary é um objeto JExpression:

```
double salary = JExprSalary.getDouble();
```

## getInt

Retorna o valor de um resultado de expressão como um tipo de dados Inteiro.

Utilize a seguinte sintaxe:

```
objectName.getInt();
```

Por exemplo, use o seguinte código Java para obter o resultado de uma expressão que retorna um número de ID de funcionário como um inteiro, onde findEmpID é um objeto JExpression:

```
int empID = findEmpID.getInt();
```

## getLong

Retorna o valor de um resultado de expressão como um tipo de dados Longo. Obtém o resultado de uma expressão que usa um tipo de dados Data.

Utilize a seguinte sintaxe:

```
objectName.getLong();
```

Por exemplo, use o seguinte código Java para obter o resultado de uma expressão que retorna um valor Data como um tipo de dados Longo, onde JExprCurrentDate é um objeto JExpression:

```
long currDate = JExprCurrentDate.getLong();
```

## getResultDataType

Retorna o tipo de dados de um resultado da expressão. Retorna um valor EDataType.

Utilize a seguinte sintaxe:

```
objectName.getResultDataType();
```

Use o seguinte exemplo de código Java para invocar uma expressão e atribuir o tipo de dados do resultado ao tipo de dados da variável:

```
myObject.invoke(new Object[] { NAME,COMPANY }, ERowType INSERT);  
EDataType dataType = myObject.getResultDataType();
```

## getResultMetadata

Retorna os metadados em um resultado de expressão. Você pode usar getResultMetadata para obter a precisão, a escala e o tipo de dados de um resultado de expressão. Você pode atribuir os metadados do valor de retorno de uma expressão a um objeto JExprParamMetadata. Use os métodos de objeto getScale, getPrecision e getDataType para recuperar os metadados do resultado.

Utilize a seguinte sintaxe:

```
objectName.getResultMetadata();
```

Use o seguinte código Java para atribuir a escala, a precisão e o tipo de dados do valor de retorno de `myObject` às variáveis:

```
JExprParamMetadata myMetadata = myObject.getResultMetadata();
int scale = myMetadata.getScale();
int prec = myMetadata.getPrecision();
int datatype = myMetadata.getDataType();
```

**Nota:** O método de objeto `getDataType` retorna o valor inteiro do tipo de dados, conforme enumerado em `EDataType`.

## getStringBuffer

Retorna o valor de um resultado de expressão como um tipo de dados `String`.

Utilize a seguinte sintaxe:

```
objectName.getStringBuffer();
```

Use o código Java a seguir para obter o resultado de uma expressão que retorna duas strings concatenadas, onde `JExprConcat` é um objeto `JExpression`:

```
String result = JExprConcat.getStringBuffer();
```

## invoke

Chama uma expressão. Os argumentos para invocação incluem um objeto que define os parâmetros de entrada e o tipo de linha. Você deve instanciar um objeto `JExpression` antes de usar o método de invocar. Para o tipo de linha, use `ERowType.INSERT`, `ERowType.DELETE` e `eRowType.UPDATE`.

Utilize a seguinte sintaxe:

```
objectName.invoke(
    new Object[] { param1[, ... paramN ]},
    rowType
);
```

A tabela a seguir descreve os argumentos:

Argumento	Tipo de dados	Entrada/Saída	Descrição
<code>objectName</code>	<code>JExpression</code>	Entrada	Nome do objeto <code>JExpression</code> .
parâmetros	-	Entrada	Matriz de objetos que contém os valores de entrada para a expressão.

Por exemplo, você cria uma função na guia de entrada de código **Funções** chamada `address_lookup()` que retorna um objeto `JExpression` que representa a expressão. Use o seguinte código para invocar a expressão que usa as portas de entrada `NAME` e `COMPANY`:

```
JExpression myObject = address_lookup();
myObject.invoke(new Object[] { "NAME", "COMPANY" }, ERowType.INSERT);
```

## isResultNull

Verifica o valor de um resultado de expressão.

Utilize a seguinte sintaxe:

```
objectName.isResultNull();
```

Use o seguinte código Java para invocar uma expressão e atribuir seu valor de retorno da expressão ao endereço da variável se o valor de retorno não for nulo:

```
JExpression myObject = address_lookup();
myObject.invoke(new Object[] { NAME,COMPANY }, ERowType INSERT);
if(!myObject.isResultNull()) {
    String address = myObject.getStringBuffer();
}
```

## CAPÍTULO 22

# Transformação de Associador

Este capítulo inclui os seguintes tópicos:

- [Transformação de Associador - Visão Geral, 361](#)
- [Propriedades Avançadas da Transformação de Associador, 362](#)
- [Caches de Associador, 363](#)
- [Portas de Transformação de Associador, 364](#)
- [Transformações de Associador em Mapeamentos Dinâmicos, 365](#)
- [Seletores de Portas em uma Transformação de Associador, 365](#)
- [Definindo uma Condição de Associação, 368](#)
- [Tipos de Associação, 372](#)
- [Entrada Classificada para uma Transformação de Associador, 375](#)
- [Associando Dados da Mesma Origem, 379](#)
- [Bloqueando os Pipelines de Origem, 381](#)
- [Transformação de Associador - Dicas de Desempenho, 382](#)
- [Regras e Diretrizes para uma Transformação de Associador, 383](#)
- [Transformação de Associador em um ambiente não nativo , 384](#)

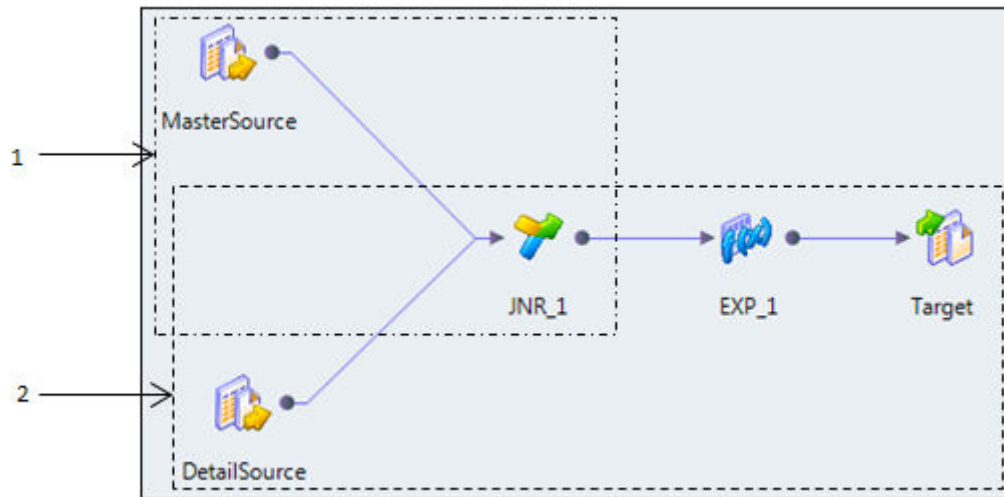
## Transformação de Associador - Visão Geral

A transformação de Associador une os dados de origem de duas origens heterogêneas relacionadas de diferentes localizações ou de diferentes sistemas de arquivo. Também é possível associar dados da mesma origem. A transformação de Associador é uma transformação ativa de vários grupos.

A transformação Joiner associa origens a pelo menos uma coluna correspondente. A transformação Joiner utiliza uma condição que corresponde a um ou mais pares de colunas entre as duas origens.

Os dois pipelines de entrada contam com um pipeline mestre e um pipeline de detalhes, ou uma ramificação mestre e de detalhes. O pipeline mestre termina na transformação Joiner, ao passo que o pipeline de detalhes segue para o destino.

A seguinte figura mostra os pipelines mestre e de detalhes em um mapeamento com uma transformação de Associador:



1. Pipeline Mestre
2. Pipeline de Detalhes

Para associar mais de duas origens em um mapeamento, associe a saída da transformação Joiner com outro pipeline de origem. Adicione transformações de Joiner ao mapeamento até ter associado todos os pipelines de origem.

## Propriedades Avançadas da Transformação de Associador

Configure propriedades que ajudem a determinar como o Serviço de Integração de Dados processa dados para a Transformação de Associador.

Configure as seguintes propriedades na guia **Avançado**:

### Tamanho de Cache de Dados de Associador

Quantidade de memória que o Serviço de Integração de Dados aloca para o cache de dados para a transformação no início da execução do mapeamento. Selecione Automático para que o Serviço de Integração de Dados calcule automaticamente os requisitos de memória no tempo de execução. Insira um valor específico em bytes quando você ajustar o tamanho do cache. O padrão é Automático.

### Tamanho de Cache de Índice de Associador

Quantidade de memória que o Serviço de Integração de Dados aloca para o cache de índice para a transformação no início da execução do mapeamento. Selecione Automático para que o Serviço de Integração de Dados calcule automaticamente os requisitos de memória no tempo de execução. Insira um valor específico em bytes quando você ajustar o tamanho do cache. O padrão é Automático.

### Diretório de Cache

O diretório onde o Serviço de Integração de Dados cria os arquivos de cache de índice e de dados. Verifique se o diretório existe e se contém espaço em disco suficiente para os arquivos de cache.

Insira vários diretórios separados por ponto-e-vírgula para melhorar o desempenho durante o particionamento do cache. O particionamento do cache cria um cache separado para cada partição que processa a transformação.

O padrão é o parâmetro do sistema CacheDir. Você pode configurar outro parâmetro do sistema ou parâmetro definido pelo usuário para essa propriedade.

#### **Entrada Classificada**

Indica que os dados de entrada são pré-classificados por grupos. Escolha Entrada Classificada para associar os dados classificados. O uso da entrada classificada pode aumentar o desempenho.

#### **Ordem de Classificação Mestra**

Especifica a ordem de classificação dos dados da origem mestra. Escolha Crescente se os dados da origem mestra estiverem na ordem crescente. Se você escolher Crescente, ative também a entrada classificada. O padrão é Automático.

#### **Nível de Rastreamento**

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.

#### **TÓPICOS RELACIONADOS:**

- [“Tamanho do Cache” na página 72](#)

## Caches de Associador

Quando você executa um mapeamento que usa uma transformação de Associador, o Serviço de Integração de Dados cria um cache de índice e um cache de dados na memória para processar a transformação. Se o Serviço de Integração de Dados exigir mais espaço que o disponível no cache de memória, ele armazenará os dados excedentes nos arquivos de cache.

Quando você executa um mapeamento que usa uma transformação de Associador, o Serviço de Integração de Dados lê as linhas das origens mestra e de detalhes simultaneamente e cria caches de índice e de dados com base nas linhas mestras. O Serviço de Integração de Dados executa a associação com base nos dados de origem detalhados e nos dados principais armazenados em cache.

O tipo de transformação de Associador determina o número de linhas que o Serviço de Integração de Dados armazena no cache.

A seguinte tabela descreve as informações que o Serviço de Integração de Dados armazena nos caches para tipos diferentes de transformação de Associador:

Tipo de Transformação de Associador	Cache de Índice	Cache de Dados
Entrada Não Classificada	Armazena todas as linhas mestras na condição de associação com chaves de índice exclusivas.	Armazena todas as linhas mestras.
Entrada Classificada com Diferentes Origens	Armazena 100 linhas mestras na condição de associação com chaves de índice exclusivas.	Armazena linhas mestras que correspondam às linhas armazenadas no cache de índice. Se os dados principais contiverem várias linhas com a mesma chave, o Serviço de Integração de Dados armazenará mais de 100 linhas no cache de dados.
Entrada Classificada com a Mesma Origem	Armazena todas as linhas mestras ou detalhadas na condição de associação com chaves exclusivas. Armazena linhas detalhadas se o Serviço de Integração de Dados processa o pipeline de detalhes mais rápido do que o pipeline mestre. Caso contrário, armazena linhas mestras. O número de linhas que ele armazena depende das taxas de processamento dos pipelines mestre e de detalhes. Se um pipeline processar as linhas de forma mais rápida do que outro, o Serviço de Integração de Dados armazenará em cache todas as linhas que já foram processadas. O serviço mantém as linhas em cache até outro pipeline concluir o processamento de suas linhas.	Armazena dados para as linhas armazenadas no cache de índice. Se o cache de índice armazenar chaves para o pipeline mestre, o cache de dados armazenará os dados para o pipeline mestre. Se o cache de índice armazenar chaves para o pipeline de detalhes, o cache de dados armazenará os dados para o pipeline de detalhes.

## Portas de Transformação de Associador

Uma transformação de Associador tem diferentes tipos de porta que determinam como o Serviço de Integração de Dados executa a associação.

Uma transformação de Associador tem os seguintes tipos de porta:

### Mestra

Portas que se vinculam à origem mestra no mapeamento.

### Detalhe

Portas que se vinculam à origem de detalhes no mapeamento.

### Porta Dinâmica

Recebe ou retorna portas em um mapeamento dinâmico. Uma porta dinâmica pode receber uma ou mais colunas de uma transformação em upstream e criar uma porta gerada para cada coluna. Uma porta de saída dinâmica pode retornar uma ou mais portas geradas. Você pode definir regras de entrada para determinar quais colunas são recebidas por uma porta dinâmica.

Você pode converter porta mestra em porta de detalhes e vice-versa. Ao mudar o tipo de uma porta, você muda o tipo de todas as portas. Portanto, quando você converte uma porta mestra em porta de detalhes, converte todas as portas mestras em portas de detalhes e todas as portas de detalhes em portas mestras.



# Transformações de Associador em Mapeamentos Dinâmicos

É possível usar uma transformação de Associador em um mapeamento dinâmico. Você pode fazer referência a portas dinâmicas e portas geradas em condições de associação.

Um mapeamento dinâmico é um mapeamento no qual as origens, os destinos e a lógica de transformação podem mudar em tempo de execução. É possível definir parâmetros e regras para alterar a estrutura dos dados. Quando uma transformação de Associador é utilizada em um mapeamento dinâmico, a estrutura da origem pode mudar. As portas de entrada e as portas na condição de associação também mudam.

As seguintes tarefas podem ser realizadas para configurar uma transformação de Associador em um mapeamento dinâmico:

## **Definir portas dinâmicas.**

Defina portas dinâmicas e portas geradas para acomodar as diferentes colunas de entrada de uma origem dinâmica. Você pode incluir portas dinâmicas ou portas geradas na condição de associação.

## **Definir seletores de portas.**

Defina um seletor de porta que contenha as portas a serem usadas na condição de associação. Configure regras de seleção que determinem as portas a serem incluídas no seletor de portas. Você pode parametrizar o seletor de portas para incluir portas específicas em tempo de execução.

## **Parametrizar a condição de associação.**

Você pode parametrizar a condição de associação inteira. Configure um parâmetro de expressão para cada condição de associação necessária.

Para obter mais informações sobre mapeamentos dinâmicos, consulte o *Guia de Mapeamentos do Informatica Developer*.

# Seletores de Portas em uma Transformação de Associador

Quando a transformação de Associador gerou portas, você precisa configurar uma condição de associação que seja válida quando a transformação tiver portas geradas diferentes em tempo de execução.

Por exemplo, um mapeamento dinâmico inclui uma transformação de Associador com a seguinte condição de Associação:

```
CustomerID = CustomerNo
```

CustomerID é uma porta gerada na transformação de Associador. Como o mapeamento tem uma origem dinâmica, ele pode ser executado com vários formatos de arquivos de origem diferentes. A coluna que contém o número do cliente tem um nome diferente em cada arquivo de origem: CustomerID, CustomerNum ou CustNO.

É possível criar um seletor de portas na transformação de Associador para acomodar os diferentes nomes de coluna de cliente da origem dinâmica. Configure um seletor de porta com uma regra de seleção que inclua qualquer nome de porta com o prefixo "Cust".

Em seguida, configure a condição de Associação para incluir o nome do seletor de portas em vez do nome da coluna CustomerID:

```
Customer_PortSelector = CustomerNo
```

A condição de associação é válida com qualquer nome de porta que comece com "Cust".

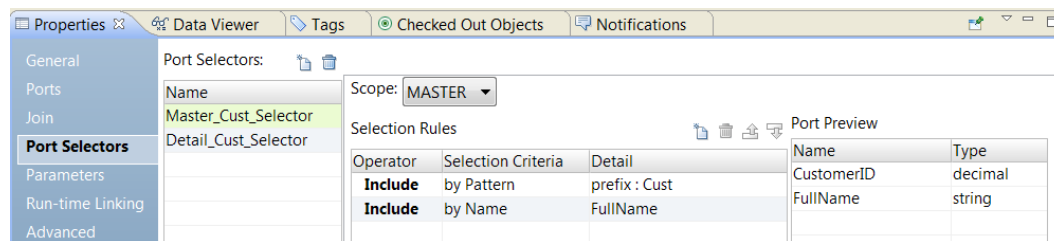
Um seletor de portas pode conter uma ou mais portas. A condição de associação poderá incluir várias portas se o grupo mestre e o grupo detalhado nessa condição contiverem o mesmo número de portas.

## Regras de Seleção

Ao configurar um seletor de portas, você define regras de seleção para determinar quais portas geradas devem ser incluídas. As regras de seleção são semelhantes às regras de entrada que você pode configurar para portas dinâmicas.

Um seletor de portas pode incluir portas ou portas geradas. Configure um seletor de portas na guia **Seletor de Portas**.

A seguinte imagem mostra a guia **Seletor de Portas**:



Configure as seguintes propriedades para um seletor de portas:

### Nome

Identifica o seletor de portas. Você pode criar vários seletores de portas em uma transformação e fazer referência a eles em expressões.

### Escopo

Identifica um grupo de portas ao qual o seletor de portas se aplica. Escolha um escopo Mestre ou Detalhado.

### Regras de Seleção

Determina as portas a serem incluídas no seletor de portas. Quando você cria as regras de seleção, o painel **Visualização de Portas** mostra as portas qualificadas a partir das portas de entrada atuais. Essas portas podem mudar. Configure as regras de seleção para acomodar portas de diferentes origens.

Você pode criar regras de seleção com base nos seguintes critérios:

#### Operador

Inclui ou exclui as portas que são retornadas pelas portas de seleção. O padrão é incluir. Você deve incluir portas antes de poder excluí-las.

#### CrITÉRIOS de Seleção

O tipo de regra de seleção que você deseja criar. Você pode criar uma regra com base no tipo de porta ou no nome da coluna. Para incluir portas com base no nome da coluna, procure nomes específicos ou procure um padrão de caracteres no nome.

#### Detalhe

Os valores a serem aplicados aos critérios de seleção. Se os critérios de seleção forem por nome de coluna, configure a cadeia ou o nome a ser pesquisado. Se os critérios de seleção forem por tipo de porta, selecione os tipos de porta a serem incluídos.

A seguinte tabela descreve os critérios de seleção e como especificar os detalhes dos critérios:

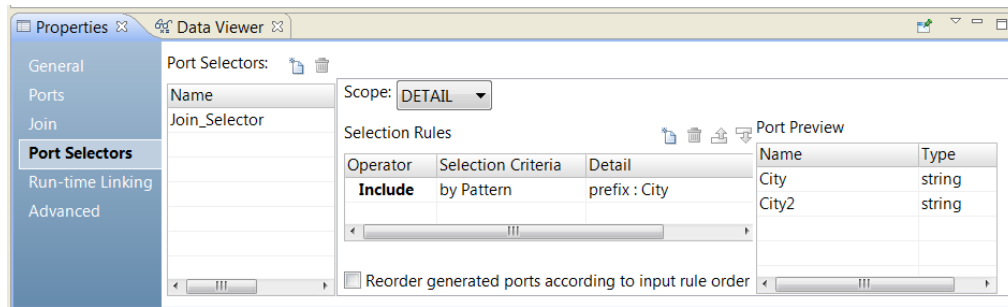
<b>Crítérios de Seleção</b>	<b>Descrição</b>	<b>Detalhe</b>
Tudo	Inclui todas as portas.	Nenhum detalhe é necessário.
Nome	Filtra portas com base no nome da porta.	Selecione os nomes de portas de uma lista de valores ou use um parâmetro do tipo Porta ou Lista de Portas.
Tipo	Filtra portas com base no tipo de dados de cada uma.	Selecione tipos de dados de uma lista.
Padrão	Filtra portas por uma cadeia de caracteres no nome ou por uma expressão regular.	Escolha o prefixo, o sufixo ou a expressão regular como o tipo de padrão para o nome da porta. Em seguida, insira um valor para o padrão ou use um parâmetro do tipo Cadeia.

## Criando um Seletor de Portas

Crie um seletor de portas para determinar quais portas devem ser usadas em uma expressão dinâmica, uma condição de pesquisa ou uma condição de associador.

1. Clique na guia **Seletores de Portas**.
2. Na área **Seletores de Portas**, clique em **Novo**.  
A Developer tool cria um seletor de portas com uma regra de seleção padrão que inclui todas as portas.
3. Na área **Seletores de Portas**, mude o nome do seletor de portas para um nome exclusivo.
4. Se estiver trabalhando na transformação de Associador ou na transformação de Pesquisa, escolha o escopo.  
As portas disponíveis mudam de acordo com o grupo de portas que você escolhe.
5. Na área **Regras de Seleção**, selecione um **Operador**.
  - Incluir. Crie uma regra que inclua portas para o seletor de portas. Você deve incluir portas antes de poder excluí-las.
  - Excluir. Crie uma regra que exclua portas específicas do seletor de portas.
6. Escolha os **Crítérios de Seleção**.
  - Por Nome. Selecione portas específicas por nome. Você pode selecionar os nomes das portas em uma lista de portas no escopo.
  - Por Tipo. Selecione portas por tipo. Você pode selecionar um ou mais tipos de dados.
  - Por Padrão. Selecione portas por um padrão de caracteres no nome da porta. É possível pesquisar com caracteres específicos ou criar uma expressão regular.

A seguinte imagem mostra a guia Seletor de Portas:



7. Clique na coluna **Detalhes**.

A caixa de diálogo **Detalhes da Regra de Entrada** é exibida.

8. Selecione os valores pelos quais filtrar as portas.

- Por Nome. Opte por criar uma lista de portas por valor ou de acordo com um parâmetro. Clique em **Escolher** para selecionar as portas na lista.
- Por Tipo. Selecione um ou mais tipos de dados de uma lista. A área **Visualização de Portas** mostra portas dos tipos que você seleciona.
- Por Padrão. Opte por pesquisar o prefixo ou sufixo do nome da porta em busca de um padrão específico de caracteres. Ou, opte por criar uma expressão regular com a qual pesquisar. Configure um parâmetro ou configure o padrão com base no qual pesquisar.

A área **Visualização de Portas** mostra as portas no seletor de portas à medida que você configura as regras.

9. Para reordenar as portas no seletor de portas, selecione **Reordenar portas geradas de acordo com a ordem das regras de entrada**.

## Definindo uma Condição de Associação

A condição de associação contém portas de ambas as origens de entrada que o Serviço de Integração de Dados usa para associar duas linhas.

Dependendo do tipo de associação selecionada, o Serviço de Integração de Dados adiciona a linha ao conjunto de resultados ou a descarta. A transformação de Associador produz conjuntos de resultados com base no tipo de associação, na condição e nas origens de dados de entrada.

Antes que você defina uma condição de associação, verifique se as origens mestre e detalhada foram configuradas para um desempenho ideal. Durante uma execução de mapeamento, o Serviço de Integração de Dados compara cada linha da origem mestra com a origem de detalhes. Para aumentar o desempenho de uma Transformação Joiner não ordenada, use a origem com menos linhas que a origem mestra. Para aumentar o desempenho de uma Transformação Joiner ordenada, use a origem com menos valores-chave duplicados como o mestre.

Use uma ou mais portas das origens de entrada de uma transformação Joiner na condição de associação. Portas adicionais aumentam o tempo necessário para a associação de duas origens. A ordem das portas na condição pode afetar o desempenho da transformação Joiner. Se você usar várias portas na condição de associação, o Serviço de Integração de Dados comparará as portas na ordem que você especificar.

Se você associar os tipos de dados Char e Varchar, o Serviço de Integração de Dados contará qualquer espaço que preencher os valores Char como parte da cadeia:

```
Char(40) = "abcd"  
Varchar(40) = "abcd"
```

O valor Char é preenchido por "abcd" com 36 espaços em branco, e o Serviço de Integração de Dados não faz a associação dos dois campos, porque o campo Char contém espaços à direita.

**Nota:** A transformação Joiner não corresponde valores nulos. Por exemplo, se EMP\_ID1 e EMP\_ID2 contiverem uma linha com um valor nulo, o Serviço de Integração de Dados não os considerará como uma correspondência e não associará as duas linhas. Para juntar linhas com valores nulos, substitua a entrada nula por valores padrão e, em seguida, junte os valores padrão.

Você pode definir um tipo de condição simples ou avançada. Também pode definir um parâmetro de expressão. Um parâmetro de expressão é um parâmetro que contém uma expressão de associação. Você pode alterar o valor do parâmetro em tempo de execução com um parâmetro de mapeamento.

## Tipo de Condição Simples

Defina um tipo de condição simples para uma transformação de Associador classificada ou não classificada.

Uma condição simples inclui uma ou mais condições que comparam as origens mestra e de detalhes especificadas. Uma condição simples deve usar o seguinte formato:

```
<master_port> operator <detail_port>
```

Para uma transformação de Associador classificada, a condição deve usar o operador de igualdade.

Para uma transformação de Associador não classificada, a condição pode usar qualquer um dos seguintes operadores: =, !=, >, >=, <, <=.

Por exemplo, se duas origens com tabelas chamadas EMPLOYEE\_AGE e EMPLOYEE\_POSITION contiverem números de ID de funcionário, a seguinte condição corresponderá linhas com funcionários listadas em ambas as origens:

```
EMP_ID1 = EMP_ID2
```

A Developer tool valida os tipos de dados em uma condição simples. Ambas as portas da condição devem ter o mesmo tipo de dados. Se você precisar usar duas portas na condição com tipos de dados que não correspondem, converta os tipos de dados para que correspondam.

Você pode configurar uma lista de condições de associação em uma condição simples. Quando várias condições de associação são configuradas, todas elas devem ser verdadeiras para fazer a associação.

Por exemplo, é possível configurar as seguintes instruções em uma condição simples:

```
StoreID = StoreNO  
Dept = Department  
Salary > Commission
```

Se você visualizar as mesmas instruções como uma Condição Avançada, a condição de associação aparecerá como a seguinte expressão:

```
StoreID = StoreNO AND Dept = Department AND (Salary > Commission)
```

## Tipo de Condição Avançada

Defina um tipo de condição avançada para uma transformação de Associador não classificada.

Uma condição avançada pode incluir qualquer expressão avaliada como um valor booleano ou numérico.

Uma condição avançada podem incluir qualquer um dos seguintes operadores: =, !=, >, >=, <, <=.

É possível inserir uma constante para a condição de associação. O equivalente numérico de FALSE é zero (0). Qualquer valor diferente de zero equivale a TRUE. Por exemplo, a transformação contém uma porta com o nome NUMBER\_OF\_UNITS com um tipo de dados numérico. Você configura uma condição de filtro para retornar FALSE se o valor de NUMBER\_OF\_UNITS for igual a zero. Caso contrário, a condição retorna TRUE.

**Nota:** Não é possível usar uma única porta dinâmica ou um seletor de portas como valor booleano para uma condição de associação.

Para inserir uma expressão na condição de associação, escolha o tipo de condição Avançado na guia **Associação**. Use o Editor de Expressão para incluir portas, parâmetros, expressões, seletores de portas e operadores na condição. É possível usar portas geradas. Você poderá inserir apenas uma porta no Editor de Expressão se o tipo de porta for numérico. No entanto, não é possível inserir um seletor de portas como expressão.

Por exemplo, você deseja associar origens correspondendo o nome completo de um funcionário. A origem mestra inclui uma porta FirstName e uma LastName. A origem de detalhes inclui uma porta FullName. Defina a seguinte condição para concatenar as portas mestras e corresponder o nome completo em ambas as origens:

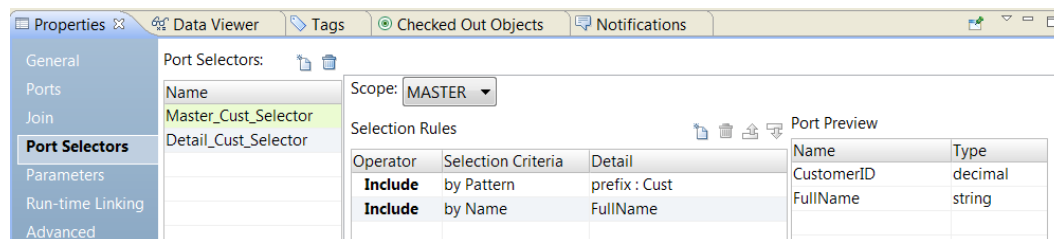
```
CONCAT(FirstName, LastName) = FullName
```

## Seletores de Portas em Condições de Associação

É possível incluir seletores de portas em uma condição de associação. A condição de junção deve fazer referência a um seletor de portas do grupo mestre e um a seletor de portas do grupo detalhado.

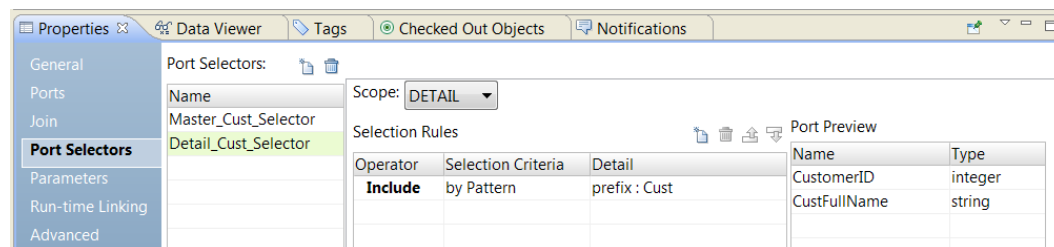
Por exemplo, a transformação de Associador tem portas dinâmicas. Talvez seja necessário comparar várias portas geradas na condição de associação.

A seguinte imagem mostra os campos no seletor de portas para o grupo mestre:



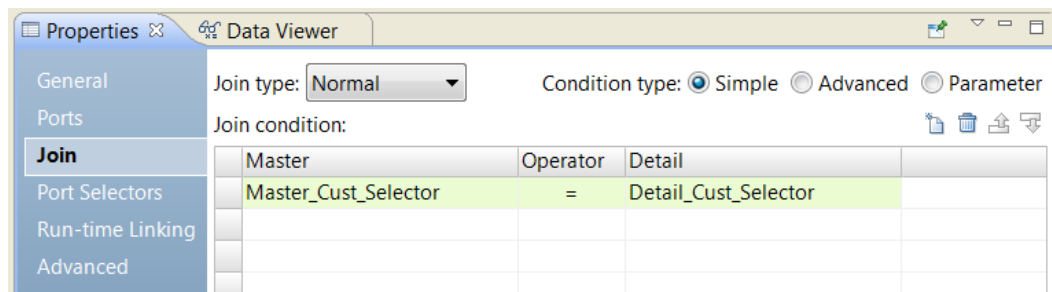
Master\_Cust\_Selector contém as portas CustomerID e FullName.

A seguinte imagem mostra os campos no seletor de portas para o grupo detalhado:



Detail\_Cust\_Selector contém as portas CustomerNo e CustFullName. Essas portas têm o prefixo Cust.

Crie a seguinte condição de associação simples:



A condição de associação compara cada porta em Master\_Cust\_Selector com Detail\_Cust\_Selector. A condição de associação é: CustomerID = CustomerNo AND FullName = CustFullName.

Cada seletor de porta deve conter o mesmo número de portas. As portas devem ser do mesmo tipo.

**Nota:** Se você alterar o escopo de um seletor de portas, e uma condição de associação de tipo simples não for mais válida, a Developer tool poderá mudar o tipo de condição para avançado. É possível mudar o tipo de condição de associação de volta para um tipo simples na guia **Associação**.

## Portas Dinâmicas em Condições de Associação

É possível fazer referência a uma porta dinâmica em um seletor de portas.

Uma porta dinâmica pode conter uma ou mais portas geradas. Se a condição de associação contiver portas dinâmicas, o número de portas mestras deverá ser igual ao número de portas detalhadas.

Por exemplo, uma porta dinâmica A tem 2 portas geradas:

```
CustomerID
OrderID
```

A porta dinâmica B também tem 2 portas geradas:

```
CustomerNo
OrderNo
```

A seguinte condição de associação é válida:

DynamicPortA = DynamicPortB

A condição de associação se expande para a seguinte expressão:

```
CustomerID = CustomerNo AND OrderID = OrderNo
```

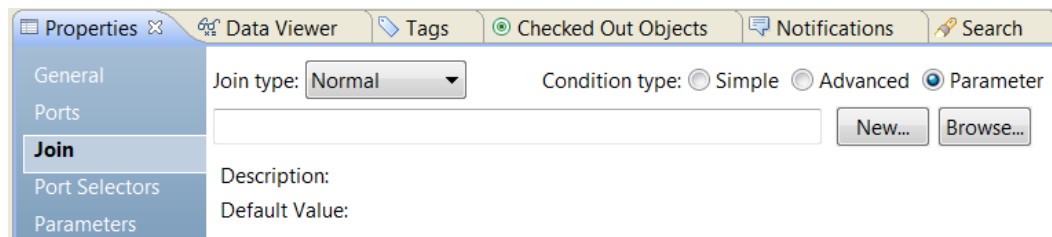
É possível fazer referência a um seletor de portas e a uma porta dinâmica em uma condição de associação quando o seletor de portas contém o mesmo número de portas que a porta dinâmica.

## Parâmetro de Expressão

Você pode definir um parâmetro de expressão que contém uma condição de associação. É possível escolher o parâmetro como a condição de associação na transformação de Associador.

Para usar um parâmetro para a condição de associação, selecione o tipo de condição de parâmetro na guia Associação.

A seguinte imagem mostra onde selecionar o tipo de condição de parâmetro:



Você pode procurar um parâmetro existente ou pode criar um parâmetro. Para criar um parâmetro, clique em **Novo** e defina o parâmetro. Crie a expressão no editor de expressão.

**Nota:** Um parâmetro de expressão não pode conter outros parâmetros. Se você incorporar um parâmetro em um parâmetro de expressão, o Serviço de Integração de Dados emitirá um erro de validação de tempo de execução.

## Tipos de Associação

Em uma transformação de Associador, a associação pode se originar de diferentes tipos de origens.

A transformação de Associador oferece suporte para os seguintes tipos de associação:

- Normal
- Externa mestra
- Externa de detalhes
- Externa completa

**Nota:** Uma associação externa normal ou mestre trabalha mais rápido do que uma associação externa completa ou externa de detalhes.

Se um conjunto de resultados apresentar campos que não contenham dados em nenhuma das origens, a transformação de Associador preencherá os campos vazios com valores nulos. Se você souber que um campo retornará NULL e você não deseja inserir NULLs no destino, poderá definir um valor padrão para a porta correspondente.

## Associação Normal

Com uma associação normal, o Data Integration Service descarta todas as linhas de dados da origem mestra e de detalhes que não tenham correspondência, com base na condição.

Por exemplo, você tem duas fontes de dados para peças automotivas chamadas PARTS\_SIZE e PARTS\_COLOR.

A fonte de dados PARTS\_SIZE é a origem mestra e contém os seguintes dados:

PART_ID1	DESCRIPTION	SIZE
1	Seat Cover	Large
2	Ash Tray	Small
3	Floor Mat	Medium



A fonte de dados PARTS\_COLOR é a origem de detalhes e contém os seguintes dados:

PART_ID2	DESCRIPTION	COLOR
1	Seat Cover	Blue
3	Floor Mat	Black
4	Fuzzy Dice	Yellow

Para associar as duas tabelas correspondendo PART\_IDs nas duas origens, defina a condição da seguinte forma:

```
PART_ID1 = PART_ID2
```

Ao associar essas tabelas com uma associação normal, o conjunto de resultados apresenta os seguintes dados:

PART_ID	DESCRIPTION	SIZE	COLOR
1	Seat Cover	Large	Blue
3	Floor Mat	Medium	Black

O exemplo a seguir mostra a instrução SQL equivalente:

```
SELECT * FROM PARTS_SIZE, PARTS_COLOR WHERE PARTS_SIZE.PART_ID1 = PARTS_COLOR.PART_ID2
```

## Associação Externa Mestra

Uma associação externa mestra mantém todas as linhas de dados da origem de detalhes e as linhas correspondentes da origem mestra. Ela descarta as linhas sem correspondência da origem mestra.

Quando você associa as tabelas de amostra com uma associação externa mestra e a mesma condição, o conjunto de resultados inclui os seguintes dados:

PART_ID	DESCRIPTION	SIZE	COLOR
1	Seat Cover	Large	Blue
3	Floor Mat	Medium	Black
4	Fuzzy Dice	NULL	Yellow

Como nenhum tamanho foi especificado para Fuzzy Dice, o Data Integration Service preenche o campo com um NULL.

O exemplo a seguir mostra a instrução SQL equivalente:

```
SELECT * FROM PARTS_SIZE RIGHT OUTER JOIN PARTS_COLOR ON (PARTS_COLOR.PART_ID2 = PARTS_SIZE.PART_ID1)
```

## Associação Externa Detalhada

Uma associação externa de detalhes mantém todas as linhas de dados da origem mestra e as linhas correspondentes da origem de detalhes. Ela descarta as linhas sem correspondência da origem de detalhes.

Quando você associa as tabelas de amostra com uma associação externa de detalhes e a mesma condição, o conjunto de resultados inclui os seguintes dados:

PART_ID	DESCRIPTION	SIZE	COLOR
1	Seat Cover	Large	Blue
2	Ash Tray	Small	NULL
3	Floor Mat	Medium	Black

Como nenhuma cor foi especificada para Ash Tray, o Data Integration Service preenche o campo com NULL.

O exemplo a seguir mostra a instrução SQL equivalente:

```
SELECT * FROM PARTS_SIZE LEFT OUTER JOIN PARTS_COLOR ON (PARTS_SIZE.PART_ID1 =  
PARTS_COLOR.PART_ID2)
```

## Associação Externa Completa

Uma associação externa completa mantém todas as linhas de dados das origens mestra e de detalhes.

Quando você associa as tabelas de amostra com uma associação externa completa e a mesma condição, o conjunto de resultados inclui os seguintes dados:

PARTED	DESCRIPTION	SIZE	Color
1	Seat Cover	Large	Blue
2	Ash Tray	Small	NULL
3	Floor Mat	Medium	Black
4	Fuzzy Dice	NULL	Yellow

Como nenhuma cor foi especificada para Ash Tray e nenhum tamanho foi especificado para Fuzzy Dice, o Data Integration Service preenche os campos com NULL.

O exemplo a seguir mostra a instrução SQL equivalente:

```
SELECT * FROM PARTS_SIZE FULL OUTER JOIN PARTS_COLOR ON (PARTS_SIZE.PART_ID1 =  
PARTS_COLOR.PART_ID2)
```

# Entrada Classificada para uma Transformação de Associador

Você pode aumentar o desempenho da transformação de Associador com a opção de entrada classificada. Você usa a entrada classificada quando os dados estão classificados.

Quando você configura a transformação de Associador para utilizar dados classificados, o Data Integration Service aumenta o desempenho, minimizando a entrada e a saída de disco. Você vê o máximo de aumento no desempenho quando trabalha com grandes conjuntos de dados.

Para configurar um mapeamento para usar dados classificados, você estabelece e mantém uma ordem de classificação no mapeamento de modo que o Data Integration Service possa usar os dados classificados quando processar a transformação de Associador. Conclua as seguintes etapas para configurar o mapeamento:

1. Configure a ordem de classificação dos dados que deseja associar.
2. Adicione transformações que mantêm a ordem dos dados classificados.
3. Configure a transformação de Associador para usar dados classificados e configure a condição de associação para usar as portas de origem de classificação. A origem de classificação representa a origem dos dados classificados.

## Configurando a Ordem de Classificação

Configure a ordem de classificação para garantir que o Data Integration Service passe os dados classificados para a transformação de Associador.

Para configurar a ordem de classificação, use um dos seguintes métodos:

- Use arquivos simples classificados. Quando os arquivos simples contiverem dados classificados, verifique se a ordem das colunas de classificação corresponde em cada arquivo de origem.
- Use dados relacionais classificados. Use portas classificadas no objeto de dados relacionais para classificar as colunas a partir do banco de dados de origem. Configure a ordem das portas classificadas igualmente em cada objeto de dados relacionais.
- Use uma transformação de Ordenador para classificar dados relacionais ou de arquivo simples. Coloque uma transformação de Ordenador nos pipelines mestre e de detalhes. Configure cada transformação de Ordenador para usar a mesma ordem das portas de chave de classificação e a mesma direção da ordem de classificação.

Se você passar dados não classificados ou classificados incorretamente para uma transformação de Associador configurada para usar dados classificados, a execução do mapeamento falhará. O Data Integration Service registra o erro no arquivo de log.

## Adicionando Transformações ao Mapeamento

Adicione transformações ao mapeamento que mantêm a ordem dos dados classificados em uma transformação de Associador.

Você pode colocar a transformação de Associador diretamente após a origem da classificação, para manter os dados classificados.

Ao adicionar transformações entre a origem da classificação e a transformação de Associador, use as seguintes diretrizes para manter os dados classificados:

- Não coloque nenhuma das seguintes transformações entre a origem da classificação e a transformação de Associador:
  - Classificação
  - União
  - Agregador Não Classificado
  - Mapplet que contém uma das transformações precedentes
- Você poderá colocar uma transformação de Agregador classificada entre a origem da classificação e a transformação de Associador se usar as seguintes diretrizes:
  - Configure a transformação de Agregador para entrada classificada.
  - Use as mesmas portas para o grupo por colunas, na transformação de Agregador, que as portas na origem da classificação.
  - O grupo por portas deve estar na mesma ordem que as portas na origem da classificação.
- Ao associar o conjunto de resultados de uma transformação de Associador com outro pipeline, verifique se a saída de dados da primeira transformação de Associador é classificada.

## Regras e Diretrizes para Condições de Associação

Determinadas regras e diretrizes aplicam-se quando você cria condições de associação para uma transformação de Associador classificada.

Use as seguintes diretrizes ao criar condições de associação:

- Você deve definir um tipo de condição simples que use o operador de igualdade.
- Se você utilizar uma transformação de Agregador classificada entre a origem da classificação e a transformação de Associador, trate a transformação de Agregador classificada como origem da classificação ao definir a condição de associação.
- As portas usadas na condição de associação devem corresponder às portas na origem da classificação.
- Quando você configura várias condições de associação, as portas da primeira condição de associação devem corresponder às primeiras portas na origem da classificação.
- Quando você configura várias condições, a ordem das condições deve corresponder à ordem das portas na origem da classificação, e não se deve ignorar nenhuma porta.
- O número de portas classificadas na origem da classificação pode ser maior ou igual ao número de portas na condição de associação.
- Se você associar portas com tipos de dados Decimais, a precisão de cada porta deverá pertencer ao mesmo intervalo de precisão.

É possível usar um dos seguintes intervalos válidos de precisão:

- Decimais 0-18
- Decimais 19-28
- Decimais 29-38
- Decimais 39 e acima

Por exemplo, se você definir a condição `DecimalA = DecimalB` onde `DecimalA` tem precisão 15 e `DecimalB` tem precisão 25, a condição não será válida.

## Exemplo de uma Condição de Associação e Ordem de Classificação

Este exemplo mostra uma transformação de Associador que associa os pipelines mestre e de detalhes com portas classificadas.

Configure transformações de Ordenador nos pipelines mestre e de detalhes com as seguintes portas classificadas:

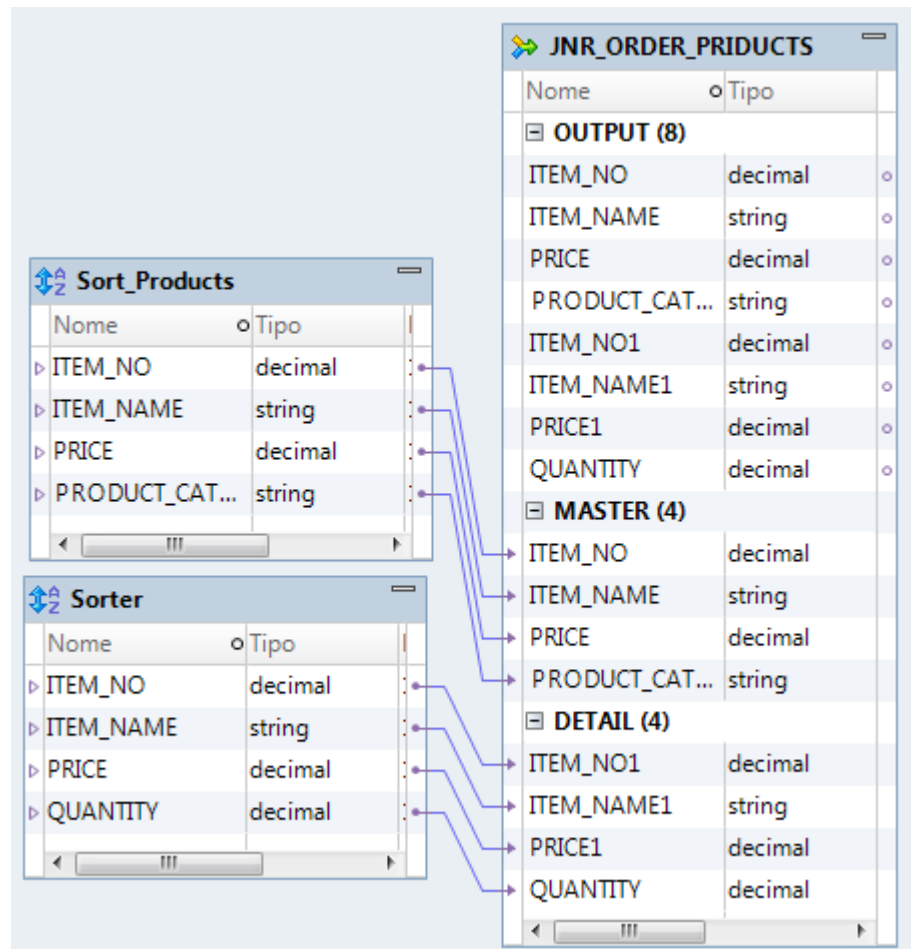
- ITEM\_NO
- ITEM\_NAME
- PRICE

Ao configurar a condição de associação, use as seguintes orientações para manter a ordem de classificação:

- É necessário usar ITEM\_NO na primeira condição de associação.
- Caso se adicione uma segunda condição de associação, será necessário usar ITEM\_NAME.
- Caso você deseje usar PRICE em uma condição de associação, também será necessário usar ITEM\_NAME na segunda condição de associação.

Caso você ignore ITEM\_NAME e faça a associação em ITEM\_NO e PRICE, a ordem de classificação será perdida, e o Serviço de Integração de Dados falhará na execução do mapeamento.

A figura a seguir mostra um mapeamento configurado para classificação e associação nas portas ITEM\_NO, ITEM\_NAME e PRICE:



Ao usar a transformação Joiner para associar os pipelines mestre e de detalhes, é possível configurar qualquer uma das seguintes condições de associação:

ITEM\_NO = ITEM\_NO

OU

ITEM\_NO = ITEM\_NO1

ITEM\_NAME = ITEM\_NAME1

OU

ITEM\_NO = ITEM\_NO1

ITEM\_NAME = ITEM\_NAME1

PRICE = PRICE1

# Associando Dados da Mesma Origem

Você poderá associar dados da mesma origem se quiser efetuar um cálculo em parte dos dados e associar os dados transformados com os dados originais.

Ao associar dados da mesma origem, você pode manter os dados originais e transformar partes desses dados dentro de um mapeamento. É possível associar dados da mesma origem das seguintes maneiras:

- Associar duas ramificações do mesmo pipeline.
- Associar duas instâncias da mesma origem.

## Associando Duas Ramificações do Mesmo Pipeline

Ao associar dados da mesma origem, você pode criar duas ramificações do pipeline.

Ao ramificar um pipeline, é necessário adicionar uma transformação entre a entrada de mapeamento e a transformação de Associador em pelo menos uma ramificação do pipeline. É necessário associar dados classificados e configurar a transformação de Associador para entrada classificada.

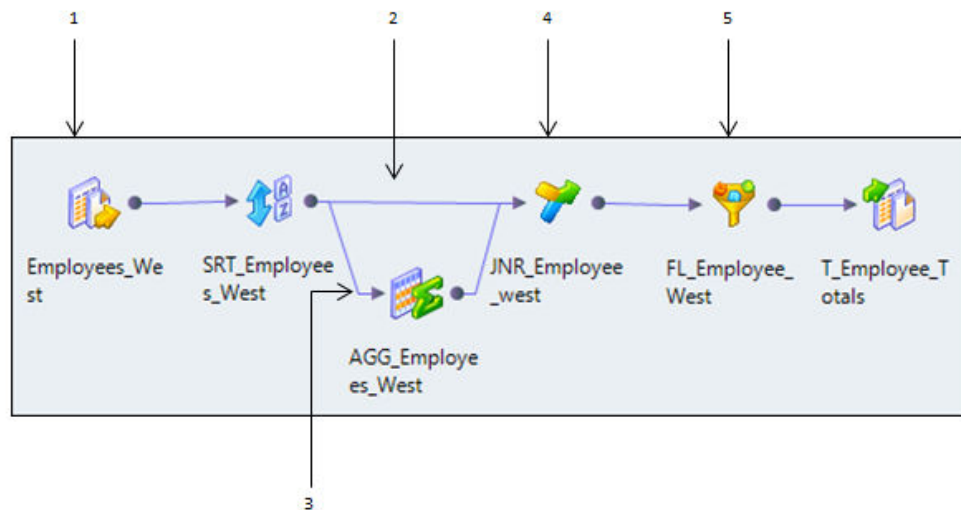
Por exemplo, há uma origem com as seguintes portas:

- Funcionário
- Departamento
- Total de Vendas

No destino, você deseja visualizar os funcionários que geraram vendas maiores que as vendas médias de seus departamentos. Para fazer isso, crie um mapeamento com as seguintes transformações:

- Transformação de Ordenador. Classifica os dados.
- Transformação de Agregador classificada. Calcula a média dos dados de vendas e do grupo por departamento. Ao efetuar essa agregação, você perde os dados de cada funcionário. Para manter os dados dos funcionários, é necessário passar uma ramificação do pipeline para a transformação de Agregador e passar uma ramificação com os mesmos dados para a transformação de Associador, a fim de manter os dados originais. Ao associar as duas ramificações do pipeline, você associa os dados agregados aos dados originais.
- Transformação de Associador classificada. Associa os dados agregados classificados com os dados originais.

- Transformação de Filtro. Compara os dados médios das vendas com os dados de vendas para cada funcionário e filtra os funcionários com vendas inferiores ao que está acima da média.



1. Employees\_West Source
2. Ramificação de pipeline 1
3. Ramificação de Pipeline 2
4. Transformação de Associador classificada
5. Filtrar os funcionários com menos do que as vendas acima da média

A associação de duas ramificações poderá diminuir o desempenho se a transformação de Associador receber dados de uma ramificação muito mais tarde do que de outra ramificação. A transformação de Associador armazena em cache todos os dados da primeira ramificação e gravará o cache em disco se ele ficar cheio. Assim, a transformação de Associador deve ler os dados do disco ao receber os dados da segunda ramificação.

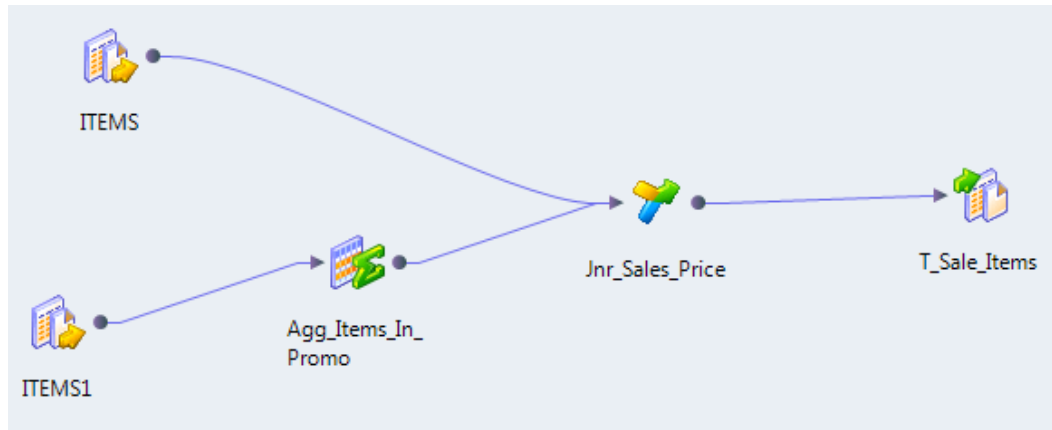
## Associando Duas Instâncias da Mesma Origem

Você pode associar dados da mesma origem, criando uma segunda instância da origem.

Depois de criar a instância da segunda origem, é possível associar os pipelines das duas instâncias de origem. Se desejar associar dados não associados, é necessário criar duas instâncias da mesma origem e associar os pipelines.



A figura a seguir mostra duas instâncias da mesma origem associadas com uma transformação Joiner:



Quando você associa duas instâncias da mesma origem, o Serviço de Integração de Dados lê os dados de origem de cada instância de origem. O desempenho pode ser inferior à associação de duas ramificações de um pipeline.

## Diretrizes para Associar Dados da Mesma Origem

Determinadas diretrizes aplicam-se quando você decide se deve associar ramificações de um pipeline ou associar duas instâncias de uma origem.

Use as seguintes diretrizes ao decidir se vai associar ramificações de um pipeline ou associar duas instâncias de uma origem:

- Associe duas ramificações de um pipeline quando tiver uma grande origem ou se puder ler os dados de origem apenas uma vez.
- Associe duas ramificações de um pipeline ao usar dados classificados. Se os dados de origem não estiverem classificados e você usar uma transformação de Ordenador para classificar os dados, ramifique o pipeline após classificar os dados.
- Associe duas instâncias de uma origem quando precisar adicionar uma transformação de bloqueio ao pipeline entre a origem e a transformação de Associador.
- Associe duas instâncias de uma origem se houver a possibilidade de um pipeline processar mais lentamente do que o outro.
- Associe duas instâncias de uma origem se precisar associar dados não classificados.

## Bloqueando os Pipelines de Origem

Quando você executa um mapeamento com uma transformação de Associador, o Data Integration Service bloqueia e desbloqueia os dados de origem com base na configuração do mapeamento e no fato de você ter ou não configurado a transformação de Associador para entrada classificada.

## Transformação de Associador Não Classificada

Quando o Data Integration Service processa uma transformação de Associador não classificada, ele lê todas as linhas mestras antes de ler as linhas de detalhes. O Data Integration Service bloqueia a origem de detalhes enquanto armazena linhas da origem mestra em cache.

Depois que o Data Integration Service lê e armazena em cache todas as linhas mestras, ele desbloqueia a origem de detalhes e lê as linhas de detalhes. Alguns mapeamentos com transformações de Associador não classificadas violam a validação do fluxo de dados.

## Transformação de Associador Classificada

Quando o Data Integration Service processa uma transformação de Associador classificada, ele bloqueia dados com base na configuração do mapeamento. A lógica de bloqueio é possível se a entrada mestra e a de detalhes para a transformação de Associador provêm de origens diferentes.

O Data Integration Service usará a lógica de bloqueio para processar a transformação de Associador se puder fazê-lo sem bloquear todas as origens de um grupo de ordem de carregamento de destino simultaneamente. Caso contrário, ele não usará a lógica de bloqueio. Em vez disso, armazenará mais linhas no cache.

Quando o Data Integration Service puder usar a lógica de bloqueio para processar a transformação de Associador, ele armazenará menos linhas no cache, aumentando o desempenho.

### Armazenando Linhas Mestras em Cache

Quando o Data Integration Service processa uma transformação de Associador, ele lê as linhas de ambas as origens simultaneamente e cria o índice e o cache de dados com base nas linhas mestras.

Em seguida, o Data Integration Service executa a associação com base nos dados da origem de detalhes e nos dados do cache. O número de linhas que o Data Integration Service armazena no cache depende dos dados de origem e de você ter ou não configurado a transformação de Associador para entrada classificada.

Para aumentar o desempenho de uma transformação de Associador não classificada, use a origem com menos linhas como origem mestra. Para aumentar o desempenho de uma transformação de Associador classificada, use a origem com menos valores de chave duplicados como mestra.

## Transformação de Associador - Dicas de Desempenho

Use as dicas para aumentar o desempenho da transformação de Associador.

As transformações de Associador podem reduzir o desempenho porque precisam de espaço adicional no tempo de execução para manter resultados intermediários. Você pode exibir as informações de contador de desempenho de Unificador para determinar se é necessário otimizar as transformações de Unificador.

Use as seguintes dicas para aumentar o desempenho com a transformação de Associador:

**Designe a origem mestra como a origem com menos valores de chave duplicados.**

Quando o Serviço de Integração de Dados processa uma transformação de Associador classificada, ele armazena em cache linhas de cem chaves exclusivas por vez. Se a origem mestra contiver muitas linhas com o mesmo valor de chave, o Serviço de Integração de Dados precisará armazenar mais linhas em cache, que poderá diminuir o desempenho.

**Designe a origem mestra como a origem com menos linhas.**

A transformação de Associador compara cada linha da origem de detalhes com a origem mestra. Quanto menos linhas houver na origem mestra, menor será o número de iterações da comparação de associações. Isso acelerará o processo de associação.

**Quando possível, faça associações em um banco de dados.**

Realizar uma associação em um banco de dados é mais rápido do que fazê-lo durante a execução do mapeamento. O tipo de associação de banco de dados usado pode afetar o desempenho. As associações normais são mais rápidas do que as associações externas e geram menos linhas. Em alguns casos, não é possível fazer a associação no banco de dados, como a associação de tabelas de dois bancos de dados diferentes ou de sistemas de arquivos simples.

**Quando possível, associe dados classificados.**

Configure a transformação de Associador para usar entrada classificada. O Serviço de Integração de Dados aumenta o desempenho, minimizando a entrada e a saída de disco. O maior aumento no desempenho ocorre quando você trabalha com grandes conjuntos de dados. No caso de uma transformação de Associador não classificada, designe a origem com menos linhas como a origem mestra.

**Otimizar a condição de associação.**

O Serviço de Integração de Dados tenta diminuir o tamanho do conjunto de dados de um operando de associação lendo as linhas a partir do grupo menor, localizando as linhas correspondentes no grupo maior e depois realizando a operação de associação. A diminuição do tamanho do conjunto de dados melhora o desempenho do mapeamento porque o Serviço de Integração de Dados deixa de ler linhas desnecessárias da origem de grupo maior. O Serviço de Integração de Dados move a condição de associação para a origem de grupo maior e lê apenas as linhas que correspondem ao grupo menor.

**Usar o método de otimização com semi-associação.**

Use o método de otimização com semi-associação para melhorar o desempenho do mapeamento quando um grupo de entrada possui muito mais linhas do que o outro e quando o grupo maior tem muitas linhas sem correspondência no grupo menor com base na condição de associação.

## Regras e Diretrizes para uma Transformação de Associador

Determinadas regras e diretrizes aplicam-se quando você usa uma transformação de Associador.

A transformação de Associador aceita entrada da maioria das transformações. Entretanto, não será possível usar uma transformação de Associador quando qualquer um dos pipelines de entrada contiver uma transformação de Estratégia de Atualização.

# Transformação de Associador em um ambiente não nativo

O processamento da transformação de Associador em um ambiente não nativo depende do mecanismo que executa a transformação.

Considere o suporte para os seguintes mecanismos de tempo de execução não nativos:

- Mecanismo Blaze. Suportado com restrições.
- Mecanismo Spark. Com suporte com restrições em mapeamentos de fluxo e em lote.
- Mecanismo Databricks Spark. Suportado com restrições.

## Transformação de Associador no mecanismo Blaze

A validação do mapeamento falha nas seguintes situações:

- A transformação contém uma associação de desigualdade e a associação do lado do mapa está desativada.
- A expressão de transformação de Associador faz referência a uma transformação de Pesquisa não conectada.

A associação do lado do mapa é desativada quando a transformação de Associador é configurada para associação externa de detalhes ou associação externa completa.

## Transformação de Associador no mecanismo Spark

A validação do mapeamento falha na seguinte situação:

- A condição de associação é do tipo de dados binary ou contém expressões binárias.

## Transformação de Associador em um mapeamento de fluxo

Os mapeamentos de fluxo têm regras de processamento adicionais que não se aplicam a mapeamentos em lote.

### Validação de Mapeamento

A validação do mapeamento falha nas seguintes situações:

- Uma transformação de Associador está downstream de uma transformação de Agregador.
- Uma transformação de Associador está downstream de uma transformação de Classificação.
- Um pipeline de fluxo contém mais de uma transformação de Associador.
- Uma transformação de Associador une dados de pipelines de fluxo e não fluxo.

### Diretrizes gerais

Para especificar uma condição de associação, selecione a função `TIME_RANGE` no tipo de condição Avançada na guia Associar e insira uma expressão de condição de associação. A função `TIME_RANGE` determina o intervalo de tempo para os eventos de fluxo a serem associados.

## Transformação de Associador no mecanismo Databricks Spark

A validação do mapeamento falha na seguinte situação:

- A condição de associação é do tipo de dados binary ou contém expressões binárias.

## CAPÍTULO 23

# Transformação de Gerador de Chaves

Este capítulo inclui os seguintes tópicos:

- [Transformação de Gerador de Chaves - Visão Geral, 386](#)
- [Estratégia de Soundex, 387](#)
- [Estratégia de String, 388](#)
- [Estratégia de NYSIIS, 388](#)
- [Portas de Saída do Gerador de Chaves, 389](#)
- [Configurando uma Estratégia de Agrupamento, 389](#)
- [Propriedades de Criação de Chave, 390](#)
- [Propriedades Avançadas da Transformação do Gerador de Chaves, 390](#)
- [Transformação de Gerador de Chaves em um ambiente não nativo, 391](#)

## Transformação de Gerador de Chaves - Visão Geral

A transformação de Gerador de Chaves é uma transformação ativa que organiza os registros em grupos com base nos valores de dados de uma coluna que você selecionar. Use esta transformação para classificar registros antes de passá-los para a transformação de Correspondência.

A transformação de Gerador de Chaves usa uma estratégia de agrupamento para criar chaves de grupo para a coluna selecionada. As estratégias são String, Soundex e NYSIIS. Registros com valores comuns no campo selecionado têm um valor comum de chave de grupo. A transformação de Correspondência processa registros com valores comuns de chave de grupo juntos. Isso permite uma análise mais rápida de duplicata na transformação de Correspondência.

O número de operações de comparação que a transformação de Correspondência deve executar cresce exponencialmente com o número de registros no conjunto de dados. Esse crescimento exponencial pode consumir quantidades significativas de recursos de computação. Ao criar chaves de grupo, a transformação de Gerador de Chaves permite que a transformação de Correspondência compare registros em grupos menores, reduzindo o tempo de processamento.

Ao executar a correspondência de campo, selecione uma coluna para geração de chave de grupo que tenha probabilidade de fornecer grupos úteis para suas necessidades de correspondência. Por exemplo, uma coluna Sobrenome tem probabilidade de fornecer dados de chave de grupo mais significativos do que uma

coluna Nome. Porém, não use a coluna Sobrenome se você pretende selecionar essa coluna para análise de duplicata na transformação de Correspondência.

A transformação de Gerador de Chaves também pode criar uma ID exclusiva para cada registro. Cada registro que entra na transformação de Correspondência deve conter uma ID exclusiva. Use a transformação de Gerador de Chaves para criar IDs para seus dados, se elas não existirem.

## Estratégia de Soundex

A estratégia de Soundex analisa palavras e cria chaves de grupo a partir de códigos que representam a pronúncia das palavras.

Os códigos Soundex começam com a primeira letra da palavra, seguida por uma série de números que representam consoantes sucessivas. Use a estratégia de Soundex para atribuir o mesmo código a palavras que têm sons semelhantes. Configure a profundidade do Soundex para definir o número de caracteres alfanuméricos que a estratégia retorna.

Essa estratégia concentra-se no som das palavras e não na ortografia, podendo agrupar ortografias alternativas e pequenas variações ortográficas. Por exemplo, os códigos Soundex para `Smyth` e `Smith` são os mesmos.

A estratégia de Soundex também pode agrupar palavras pronunciadas incorretamente. Por exemplo, os códigos Soundex para os nomes `Edmonton` e `Edmonson` são os mesmos.

## Propriedades da Estratégia de Soundex

Configure as propriedades da estratégia de Soundex para determinar as configurações de Soundex usadas pela transformação de Gerador de Chaves para criar uma chave de grupo.

A tabela a seguir descreve as propriedades da estratégia de Soundex:

Propriedade	Descrição
<b>Profundidade de Soundex</b>	Determina o número de caracteres alfanuméricos retornado pela estratégia de Soundex. A profundidade padrão é 3. Essa profundidade cria um código Soundex composto pela primeira letra da string e dois números que representam os dois sons de consoantes distintos seguintes.

### TÓPICOS RELACIONADOS:

- [“Propriedades da Estratégia de String” na página 388](#)
- [“Propriedades de Criação de Chave” na página 390](#)
- [“Configurando uma Estratégia de Agrupamento” na página 389](#)

# Estratégia de String

A estratégia de String cria chaves de grupo a partir de substrings em dados de entrada.

Você pode especificar o comprimento e a localização de uma substring dentro da coluna de entrada. Por exemplo, você pode configurar essa estratégia para criar uma chave a partir dos quatro primeiros caracteres na string de entrada.

## Propriedades da Estratégia de String

Configure as propriedades da estratégia de String para determinar as substrings usadas pela transformação de Gerador de Chaves para criar uma chave de grupo.

A tabela a seguir descreve as propriedades da estratégia de String:

Propriedade	Descrição
Iniciar pela esquerda	Configura a transformação para ler o campo de entrada da esquerda para a direita.
Iniciar pela direita	Configura a transformação para ler o campo de entrada da direita para a esquerda.
Posição inicial	Especifica o número de caracteres a ignorar. Por exemplo, se você inserir 3 para a <b>Posição inicial</b> , a substring começará no quarto caractere do campo de entrada a partir do lado que você especificar.
Comprimento	Especifica o comprimento da string a ser usada como uma chave de grupo. Insira 0 para usar o campo de entrada inteiro.

### TÓPICOS RELACIONADOS:

- [“Propriedades da Estratégia de Soundex” na página 387](#)
- [“Propriedades de Criação de Chave” na página 390](#)
- [“Configurando uma Estratégia de Agrupamento” na página 389](#)

# Estratégia de NYSIIS

A estratégia de NYSIIS analisa palavras e cria chaves de grupo a partir de letras que representam a pronúncia das palavras.

Enquanto a estratégia de Soundex considera apenas a primeira vogal de uma string, a estratégia de NYSIIS analisa vogais em toda a string. A estratégia de NYSIIS converte todas as letras em um de seis caracteres e converte a maioria das vogais na letra A.



# Portas de Saída do Gerador de Chaves

As portas de saída da transformação de Gerador de Chaves criam IDs e grupos de chaves usados pela transformação de Correspondência para processar registros.

A tabela a seguir descreve as portas de saída para a transformação de Gerador de Chaves:

Propriedade	Descrição
<b>SequenceID</b>	Cria um ID que identifica cada registro no conjunto de dados de origem.
<b>GroupKey</b>	Cria as chaves de grupo usadas pela transformação de Correspondência para processar registros.

Ao criar uma transformação de Gerador de Chave reutilizável, use a exibição **Visão Geral** para exibir as portas. Quando você adicionar uma transformação não reutilizável para um mapeamento, use a guia **Portas** da exibição **Propriedades** para exibir as portas.

## Configurando uma Estratégia de Agrupamento

Para configurar uma estratégia de agrupamento, edite as propriedades na exibição **Estratégias**.

Antes de configurar uma estratégia de Gerador de Chaves, adicione portas de entrada à transformação de Gerador de Chaves.

1. Selecione a exibição **Estratégias**.
2. Clique no botão **Novo**.
3. Selecione uma estratégia de agrupamento.
4. Clique em **OK**.
5. Na coluna **Entradas**, selecione uma porta de entrada.
6. Configure as propriedades da estratégia clicando na seta de seleção no campo de propriedades.
7. Configure as propriedades de criação de chave.

### TÓPICOS RELACIONADOS:

- [“Propriedades da Estratégia de Soundex” na página 387](#)
- [“Propriedades da Estratégia de String” na página 388](#)
- [“Propriedades de Criação de Chave” na página 390](#)

# Propriedades de Criação de Chave

Configure as propriedades de criação de chave apropriadas para os dados analisados.

A tabela a seguir descreve as propriedades de criação de chave:

Propriedade	Descrição
<b>Classificar resultados</b>	Classifica a saída da transformação de Gerador de Chaves usando o campo GroupKey. Para operações de correspondência de campo, você deverá selecionar esta opção ou verificar se está fornecendo dados classificados à transformação de Correspondência. Não selecione esta opção para operações de correspondência de identidade.
<b>Gerar chave de sequência automaticamente</b>	Gera um campo de chave de sequência usando a ordem dos dados de entrada.
<b>Usar campo como chave de sequência</b>	Gera um campo de sequência para uma coluna que você especificar.
<b>Campo de chave de sequência</b>	Especifica o nome do campo de chave de sequência.

## TÓPICOS RELACIONADOS:

- [“Propriedades da Estratégia de Soundex” na página 387](#)
- [“Propriedades da Estratégia de String” na página 388](#)
- [“Configurando uma Estratégia de Agrupamento” na página 389](#)

# Propriedades Avançadas da Transformação do Gerador de Chaves

A transformação de Gerador de Chaves contém propriedades avançadas que determinam o comportamento da memória cache e o nível de rastreamento.

Você pode configurar as seguintes propriedades avançadas:

## Diretório de Arquivo de Cache

Especifica o diretório no qual o Serviço de Integração de Dados grava os dados temporários da transformação atual. O Serviço de Integração de Dados grava os arquivos temporários no diretório quando o volume de dados de entrada é maior do que a memória disponível do sistema. O Serviço de Integração de Dados exclui os arquivos temporários depois que o mapeamento é executado.

Você pode inserir um caminho de diretório na propriedade ou usar um parâmetro para identificar o diretório. Especifique um caminho local na máquina do Serviço de Integração de Dados. O Serviço de Integração de Dados deve ser capaz de gravar no diretório. O valor padrão é o parâmetro do sistema CacheDir.

## Tamanho do Arquivo de Cache

Determina a quantidade de memória do sistema que o Serviço de Integração de Dados usa para classificar os dados de entrada na transformação. Você pode usar um parâmetro para especificar o tamanho do arquivo de cache.

Antes de classificar os dados, o Serviço de Integração de Dados aloca a quantidade de memória especificada por você. Se a operação de classificação gerar uma maior quantidade de dados, o Serviço de Integração de Dados gravará o excesso de dados no diretório de cache. Se a operação de classificação exigir mais memória do que a memória do sistema e o armazenamento de arquivos podem fornecer, o mapeamento falhará.

Se você não especificar um tamanho de arquivo de cache, a transformação aplicará o valor máximo de memória nas opções de execução do Serviço de Integração de Dados.

**Nota:** Se você inserir um valor igual ou maior que 65536, a transformação lerá o valor em bytes. Se você inserir um valor menor, a transformação lerá o valor em megabytes.

#### Nível de Rastreamento

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.

## Transformação de Gerador de Chaves em um ambiente não nativo

O processamento da transformação de Gerador de Chaves em um ambiente não nativo depende do mecanismo que executa a transformação.

Considere o suporte para os seguintes mecanismos de tempo de execução não nativos:

- Mecanismo Blaze. Suportado sem restrições.
- Mecanismo Spark. Com suporte sem restrições em mapeamentos em lote. Sem suporte em mapeamentos de fluxo.
- Mecanismo Databricks Spark. Sem suporte.

## CAPÍTULO 24

# Transformação de Rotulador

Este capítulo inclui os seguintes tópicos:

- [Transformação de Rotulador - Visão Geral, 392](#)
- [Quando Usar uma Transformação de Rotulador, 393](#)
- [Uso de Dados de Referência na Transformação de Rotulador, 394](#)
- [Estratégias de Transformação de Rotulador, 396](#)
- [Portas da Transformação de Rotulador, 397](#)
- [Propriedades de Rotulagem de Caractere, 397](#)
- [Propriedades de Rotulagem de Token, 400](#)
- [Configurando uma Estratégia de Rotulagem de Caractere, 403](#)
- [Configurando uma Estratégia de Rotulagem de Token, 403](#)
- [Propriedades Avançadas da Transformação de Rotulador, 404](#)
- [Transformação de Rotulador em um ambiente não nativo, 404](#)

## Transformação de Rotulador - Visão Geral

A transformação de Rotulador é uma transformação passiva que analisa os campos da porta de entrada e grava rótulos de texto que descrevem os dados em cada campo.

Use uma transformação de Rotulador quando quiser entender os tipos de informação que uma porta contém. Você usa uma transformação de Rotulador quando você não sabe os tipos de informação em uma porta ou quando quer identificar registros que não contêm os tipos esperados de informação em uma porta.

Um rótulo é uma string de um ou mais caracteres que descreve uma string de entrada. Você configura a transformação de Rotulador para atribuir rótulos a strings de entrada com base nos dados que cada string contém.

Ao configurar a transformação, você especifica os tipos de caractere ou string que devem ser pesquisados, bem como o rótulo que a transformação grava como saída quando encontra o caractere ou string associado(a). Você insere os tipos de caractere e string a serem pesquisados quando configura uma operação de rotulagem. Também é possível usar objetos de dados de referência para especificar os caracteres, strings e rótulos.

Você configura a transformação para executar a rotulagem de caractere ou token:

### Rotulagem de Caractere

Grava um rótulo que descreve a estrutura de caracteres da string de entrada, incluindo pontuação e espaços. A transformação grava um único rótulo para cada linha em uma coluna. Por exemplo, a transformação de Rotulador pode rotular o código postal 10028 como "nnnnn", onde "n" representa um caractere numérico.

### Rotulagem de Token

Grava um rótulo que descreve o tipo de informação na string de entrada. A transformação grava um rótulo para cada token identificado nos dados de entrada. Por exemplo, você pode configurar a transformação de Rotulador para rotular a string "John J. Smith" com os tokens "Palavra Init Palavra".

Um token é um valor delimitado em uma string de entrada.

Quando o Rotulador encontra um caractere ou string que corresponde a um rótulo especificado, ele grava o nome do rótulo em uma nova porta de saída.

A transformação de Rotulador usa dados de referência para identificar caracteres e tokens. Você seleciona o objeto de dados de referência ao configurar uma operação em uma estratégia de Rotulador.

## Quando Usar uma Transformação de Rotulador

A transformação de Rotulador grava um rótulo descritivo para cada valor em uma porta.

Os exemplos a seguir descrevem alguns dos tipos de análise que você pode executar com uma transformação de Rotulador.

### Localizar registros com dados de contato

Configure a transformação com uma tabela de referência que contém uma lista de nomes. Crie um a estratégia de rotulagem de token para rotular qualquer string que corresponda a um valor na tabela de referência. Quando você examinar os dados de saída, qualquer registro que contiver o rótulo terá probabilidade de identificar uma pessoa.

### Localizar registros de empresas

Configure a transformação com um conjunto de tokens que contém uma lista de sufixos de empresas, como Inc., Corp. e Ltd. Crie um a estratégia de rotulagem de token para rotular qualquer string que corresponda a um valor na tabela de referência. Quando você examinar os dados de saída, qualquer registro que contiver o rótulo terá probabilidade de identificar uma empresa.

**Nota:** Use um conjunto de tokens de sufixos de empresas se quiser identificar qualquer nome de empresa. Você poderá usar uma tabela de referência de nomes de empresas se tiver certeza de que a tabela contém todas as empresas que você deseja identificar. Por exemplo, você pode usar uma tabela de referência que lista as corporações que estão na Bolsa de Valores de Nova York.

### Localizar dados de número de telefone

Configure a transformação com um conjunto de caracteres que define a estrutura de caracteres de um número de telefone. Por exemplo, você pode usar um conjunto de caracteres que reconhece diferentes padrões de símbolos de pontuação e dígitos como números de telefone dos Estados Unidos. É possível examinar os dados para localizar registros que não contenham os dígitos corretos para um número de telefone.

Os rótulos de caracteres podem usar os seguintes caracteres para analisar os dados da coluna:

```
c=punctuation character n=digit s=space
```

A seguinte tabela mostra estruturas de número de telefone de amostra:

Estrutura de Caracteres	Número de telefone
cnncsnnncnnncnnnn	(212) 555-1212
nnnnnnnnnn	2125551212
cnncnnncnnnn	+212-555-1212

## Uso de Dados de Referência na Transformação de Rotulador

O Informatica Developer é instalado com diferentes tipos de objetos de dados de referência que podem ser usados com a transformação de Rotulador. Você também pode criar objetos de dados de referência.

Quando você adiciona um objeto de dados de referência à estratégia da transformação de Rotulador, a transformação pesquisa os dados de entrada da estratégia em busca de valores no objeto de dados de referência. A transformação substituirá qualquer valor encontrado por um valor válido do objeto de dados de referência ou por um valor que você especificar.

A seguinte tabela descreve os tipos de dados de referência que você pode usar:

Tipo de Dados de Referência	Descrição
Conjuntos de caracteres	Identifica diferentes tipos de caracteres, como letras, números e símbolos de pontuação. Use nas operações de rotulagem de caractere.
Modelos probabilísticos	Adiciona recursos de correspondência parcial às operações de rótulo de token. A transformação pode usar um modelo probabilístico para inferir o tipo de informação em uma string. Para ativar os recursos de correspondência parcial, você compila o modelo probabilístico na ferramenta Developer. Use em operações de rotulagem de token.
Tabelas de referência	Localiza strings que correspondem às entradas em uma tabela de banco de dados. Use em operações de rotulagem de caractere e de token.
Expressões regulares	Identifica strings que correspondem às condições definidas por você. Você pode usar uma expressão regular para localizar uma string dentro de uma string maior. Use em operações de rotulagem de token.
Conjuntos de tokens	Identifica as strings com base nos tipos de informação que elas contêm. Use em operações de rotulagem de token. O Informatica instala com os conjuntos de tokens diferentes tipos de definições de token, como definições de palavra, número de telefone, código postal e código do produto.

## Conjuntos de Caracteres

Um conjunto de caracteres contém expressões que identificam caracteres específicos e intervalos de caracteres. Você pode usar os conjuntos de caracteres em transformações de Rotulador e de Analisador que usam o modo de análise de tokens.

Os intervalos de caracteres especificam um intervalo sequencial de códigos de caractere. Por exemplo, o intervalo de caracteres "[A-C]" corresponde aos caracteres maiúsculos "A", "B" e "C". Esse intervalo de caracteres não corresponde aos caracteres minúsculos "a", "b" ou "c".

Use conjuntos de caracteres para identificar um caractere específico ou um intervalo de caracteres como parte das operações de rotulagem ou análise de tokens. Por exemplo, você pode rotular todos os numerais em uma coluna que contém números de telefone. Após rotular os números, você pode identificar padrões com uma transformação de Analisador e gravar os padrões problemáticos em saídas separadas.

## Modelos Probabilísticos

Um modelo probabilístico identifica tokens pelos tipos de informações que eles contêm e pelas posições que ocupam em uma cadeia de entrada.

Um modelo probabilístico contém valores de dados de referência e de rótulos. Os valores de dados de referência representam os dados em uma porta de entrada que você conectar à transformação. Os valores de rótulos descrevem os tipos de informações contidas nos valores de dados de referência. Atribua um rótulo para cada valor de dados de referência no modelo.

Para vincular os valores de dados de referência aos rótulos em um modelo probabilístico, compile o modelo. O processo de compilação gera uma série de associações lógicas entre os valores de dados e os rótulos. Quando você executa um mapeamento que lê o modelo, o Serviço de Integração de Dados aplica a lógica de modelo aos dados de entrada da transformação. O Serviço de Integração de Dados retorna o rótulo descreve os valores de dados de entrada com mais precisão.

Você pode criar um modelo probabilístico na Developer tool. O repositório do Modelo armazena o objeto do modelo probabilístico. A Developer tool grava os valores de dados, os rótulos e os dados de compilação em um arquivo na estrutura de diretórios da Informatica.

## Tabelas de Referência

Uma tabela de referência é uma tabela de banco de dados que contém pelo menos duas colunas. Uma coluna contém o padrão ou versão necessária de um valor de dados, e outras colunas contêm versões alternativas do valor. Quando você adiciona uma tabela de referência a uma transformação, a transformação pesquisa os dados da porta de entrada em busca de valores que também apareçam na tabela. Você pode criar tabelas com quaisquer dados que sejam úteis para o projeto de dados no qual você está trabalhando.

## Expressões Regulares

No contexto de operações de rotulagem, uma expressão regular é aquela que pode ser usada para identificar uma string específica nos dados de entrada. Você pode usar expressões regulares em transformações de Rotulador que usam o modo de rotulagem de token.

Transformações de Rotulador usam expressões regulares para corresponder a um padrão de entrada e criar um único rótulo. Expressões regulares que têm várias saídas não geram vários rótulos.

## Conjuntos de Tokens

Um conjunto de tokens contém expressões que identificam tokens específicos. Você pode usar conjuntos de tokens em transformações de Rotulador que usam o modo de rotulagem de token.

Use conjuntos de tokens para identificar tokens específicos como parte de operações de rotulagem de token. Por exemplo, você pode usar um conjunto de tokens para rotular todos os endereços de e-mail que usam um formato "NomeConta@NomeDomínio". Após rotular os tokens, você pode usar a transformação de Analisador para gravar endereços de e-mail nas portas de saída que especificar.

A ferramenta Developer inclui conjuntos de tokens definidos pelo sistema que podem ser usados para identificar uma ampla variedade de padrões. Exemplos de conjuntos de tokens definidos pelo sistema:

- Palavras
- Números
- Números de telefone
- Endereços de e-mail
- Códigos postais
- Números de identificação nacionais, como Números de Seguridade Social
- Números de cartão de crédito

## Estratégias de Transformação de Rotulador

Use estratégias de rotulagem para atribuir rótulos a dados de entrada. Para configurar uma estratégia de rotulagem, edite as configurações na exibição **Estratégias** de uma transformação de Rotulador.

Ao criar uma estratégia de rotulagem, você adiciona uma ou mais operações. Cada operação implementa uma tarefa de rotulagem específica.

A transformação de Rotulador fornece um assistente que você usa para criar estratégias. Ao criar uma estratégia de rotulagem, você escolhe entre o modo de rotulagem de caractere ou de token. Em seguida, você adiciona operações específicas para esse modo de rotulagem.

**Importante:** Você pode alterar a ordem das operações e estratégias. A ordem das operações dentro de uma estratégia pode alterar a saída de uma estratégia, pois cada operação lê os resultados da operação anterior.

## Operações de Rotulagem de Caractere

Use operações de rotulagem de caractere para criar rótulos que descrevem os padrões de caracteres em seus dados.

Você pode adicionar os seguintes tipos de operações a uma estratégia de rotulagem de caractere:

### **Caracteres de Rótulo que Usam Conjuntos de Caracteres**

Caracteres de rótulo que usam conjuntos de caracteres predefinidos, como dígitos ou caracteres alfabéticos. Você pode selecionar conjuntos de caracteres Unicode e não Unicode.

### **Caracteres de Rótulo que Usam Tabela de Referência**

Caracteres de rótulo com rótulos personalizados de uma tabela de referência.



## Operações de Rotulagem de Token

Use operações de rotulagem de token para criar rótulos que descrevem strings em seus dados.

A transformação de Rotulador pode identificar e rotular vários tokens em uma string de entrada. Por exemplo, você pode configurar a transformação de Rotulador para usar os conjuntos de tokens Número de Telefone dos EUA e Endereços de E-mail. Quando a transformação de Rotulador processa a string de entrada "555-555-1212 alguém@algumlugar.com", a string de saída é "USPHONE EMAIL".

Você pode adicionar os seguintes tipos de operações de rotulagem de token a uma estratégia de rotulagem:

### **Rótulo com Tabela de Referência**

Rotule strings que correspondem a entradas da tabela de referência.

### **Tokens de Rótulo com Conjunto de Tokens**

Rotule padrões de strings que correspondem aos dados do conjunto de token ou aos dados do modelo probabilístico.

## Portas da Transformação de Rotulador

Selecione as portas de entrada e saída necessárias para as operações de rotulagem que você configura na transformação.

As transformações de Rotulador usam a seguinte porta:

### **Portas de Entrada**

Lê a string inserida de objetos upstream.

### **Portas de Saída Rotuladas**

Grava os rótulos definidos pelas operações de transformação.

### **Portas de Saída com Token**

Passa as strings de entrada que correspondem a cada rótulo na saída. Selecione esta porta se for adicionar uma transformação de Analisador downstream da transformação de Rotulador em um mapplet ou mapeamento, e se for configurar a transformação de Analisador para executar no modo de análise baseada em padrões. A transformação de Analisador associa a saída de rotulagem de token aos dados nas portas de saída com token.

### **Portas de Saída de Pontuação**

Selecione para gravar os valores de pontuação gerados por técnicas de correspondência probabilística em uma operação de rotulagem de token.

Quando você executa uma operação de rotulagem de token que usa um modelo probabilístico, a operação gera uma pontuação numérica para cada string rotulada. A pontuação representa o grau de similaridade entre a string de entrada e os padrões definidos no modelo probabilístico.

## Propriedades de Rotulagem de Caractere

Configure propriedades para operações de rotulagem de caractere na exibição **Estratégias** da transformação de Rotulador.

## Propriedades Gerais

As propriedades gerais aplicam-se a todas as operações de rotulagem de caractere que você define na estratégia. Use as propriedades gerais para nomear a estratégia e especificar as portas de entrada e saída.

A seguinte tabela descreve as propriedades gerais:

Propriedade	Descrição
Nome	Fornecer um nome para a estratégia.
Entradas	Identifica as portas de entrada que as operações de estratégia podem ler.
Saídas	Identifica as portas de saída nas quais as operações de estratégia podem gravar.
Descrição	Fornecer uma descrição de texto da estratégia. Esta é uma propriedade opcional.

## Propriedades da Tabela de Referência

Ao definir uma estratégia de rotulagem de caractere, você pode adicionar operações para rotular com conjuntos de caracteres e com tabelas de referência. Use as propriedades da tabela de referência para especificar como a transformação usa as tabelas de referência.

A seguinte tabela descreve as propriedades da tabela de referência:

Propriedade	Descrição
Nome	Fornecer um nome para a operação.
Tabela de Referência	Especifica as tabelas de referência usadas pela transformação para rotular caracteres.
Rótulo	Especifica o texto de substituição para caracteres de entrada que correspondem a entradas da tabela de referência.
Substituir outros rótulos na estratégia	Determina se esta operação de rotulagem substitui outras operações de rotulagem.

## Propriedades do Conjunto de Caracteres

Ao definir uma estratégia de rotulagem de caractere, você pode adicionar operações para rotular com conjuntos de caracteres e com tabelas de referência. Use as propriedades do conjunto de caracteres para especificar como a transformação usa os conjuntos de caracteres.

A seguinte tabela descreve as propriedades do conjunto de caracteres:

Propriedade	Descrição
Nome	Fornecer um nome para a operação.
Selecionar Conjuntos de Caracteres	Especifica os conjuntos de caracteres usados pela transformação para rotular cadeias. Você pode ignorar o texto de substituição das cadeias de entrada que correspondem ao conjunto de caracteres. Clique na seta de seleção na coluna <b>Rótulo</b> para inserir o texto de substituição personalizado.
Filtrar texto	Usa caracteres ou caracteres curinga que você insere para filtrar a lista de conjuntos de caracteres.
Adicionar Conjunto de Caracteres	Selecione para definir um conjunto de caracteres personalizado.
Editar	Edite o conteúdo de um conjunto de caracteres personalizado.
Importar	Permite que você crie cópias não reutilizáveis de conjuntos de caracteres que são armazenadas em conjuntos de conteúdo. As alterações no conjunto de caracteres original não atualizam as cópias armazenadas na transformação de Rotulador.
Remover	Exclui um conjunto de caracteres personalizado.
Especificar Ordem de Execução	Define a ordem na qual a operação aplica os conjuntos de tokens aos dados. Use as setas Para Cima e Para Baixo para alterar a ordem.

## Propriedades de Filtro

Você pode especificar valores para ignorar durante uma operação de rotulagem. Use as propriedades de **Ignorar Texto** para especificar valores aos quais as operações de rotulagem não se aplicarão.

A seguinte tabela descreve as propriedades de filtro:

Propriedade	Descrição
Termo de Pesquisa	Especifica as strings que a transformação filtra antes de executar a rotulagem. Use este recurso para especificar as exceções à sua estratégia de rotulagem definida.
Distinção Entre Maiúsculas e Minúsculas	Determina se as strings filtradas devem corresponder à distinção entre maiúsculas e minúsculas do termo de pesquisa.
Maiúscula	Converte as strings filtradas em maiúsculas.
Início	Especifica a posição do caractere para começar a pesquisar a string filtrada.
Término	Especifica a posição do caractere para terminar de pesquisar a string filtrada.

# Propriedades de Rotulagem de Token

Configure propriedades para operações de rotulagem de token na exibição **Estratégias** da transformação de Rotulador.

## Propriedades Gerais

As propriedades gerais aplicam-se a todas as operações de rotulagem de token que você define na estratégia. Use as propriedades gerais para nomear a estratégia, especificar as portas de entrada e saída e especificar se a estratégia permite técnicas de correspondência probabilística.

A seguinte tabela descreve as propriedades gerais:

Propriedade	Descrição
Nome	Fornece um nome para a estratégia.
Entradas	Identifica a porta de saída que as operações de estratégia podem ler.
Saídas	Identifica a porta de saída na qual as operações de estratégia podem gravar.
Descrição	Descreve a estratégia. A propriedade é opcional.
Usar técnicas de correspondência probabilística	Especifica que a estratégia pode usar um modelo probabilístico para identificar os tipos de token.
Inversão Ativada	Indica que a estratégia lê os dados de entrada da direita para a esquerda. Esta propriedade é desativada para correspondência probabilística.
Delimitadores	Especifica os caracteres que a transformação usa para avaliar substrings em dados de entrada. O padrão é espaço. A propriedade está desativada na rotulagem probabilística.
Campo de Saída com Token	Indica que a estratégia grava vários rótulos em uma porta de saída. Selecione este campo para criar dados de entrada para análise baseada em padrões na transformação de Analisador.
Campo de Saída de Pontuação	Identifica o campo que contém os valores de pontuação gerados na correspondência probabilística. Defina o campo de saída de pontuação ao selecionar a opção para usar técnicas de correspondência probabilística.
Delimitador de Saída	Especifica um caractere para separar os valores de dados na porta de saída. O padrão é dois-pontos.

## Propriedades do Conjunto de Tokens

As propriedades do conjunto de token se aplicam quando você configura uma operação de rotulagem para usar conjuntos de tokens.

A seguinte tabela descreve as propriedades gerais:

Propriedade	Descrição
Selecionar Conjuntos de Tokens	Especifica os conjuntos de tokens usados pela transformação para rotular cadeias.
Filtrar texto	Filtra a lista de conjuntos de tokens ou de expressões regulares. Use caracteres de texto e caracteres curinga como um filtro.
Adicionar Conjunto de Tokens	Use para definir os conjuntos de tokens personalizados.
Adicionar Expressão Regular	Use para definir as expressões regulares que correspondem a um padrão de entrada.
Editar	Edita o conteúdo de um conjunto de tokens personalizado ou uma expressão regular.
Importar	Importa uma cópia não reutilizável de um conjunto de tokens ou de uma expressão regular de uma pasta no repositório do Modelo. Se você atualizar o objeto de origem do conjunto de tokens ou da expressão regular, o Serviço de Integração de Dados não atualizará a cópia não reutilizável.
Remover	Remove um conjunto de tokens personalizado ou uma expressão regular.
Especificar Ordem de Execução	Define a ordem na qual a operação aplica os conjuntos de tokens ou as expressões regulares aos dados. Use as setas Para Cima e Para Baixo para alterar a ordem.

## Propriedades de Rótulo Personalizadas

Ao configurar uma operação de rotulagem de token, você pode selecionar a exibição **Rótulo Personalizado** para criar rótulos para termos de pesquisa específicos.

A seguinte tabela descreve as propriedades de rótulo personalizadas:

Propriedade	Descrição
Termo de Pesquisa	Identifica a string a ser pesquisada.
Distinção Entre Maiúsculas e Minúsculas	Especifica se os dados de entrada devem corresponder à distinção entre maiúsculas e minúsculas do termo de pesquisa.
Rótulo Personalizado	Identifica o rótulo personalizado a ser aplicado.

## Propriedades da Correspondência Probabilística

Ao selecionar as opções para usar técnicas de correspondência probabilística, você pode adicionar um modelo probabilístico à operação de rotulagem. Não é possível adicionar um modelo probabilístico a uma estratégia que usa um conjunto de tokens ou uma tabela de referência.

A seguinte tabela descreve as propriedades associadas à correspondência probabilística:

Propriedade	Descrição
Nome	Fornecer um nome para a operação.
Filtrar Texto	Usa caracteres ou caracteres curinga que você insere para filtrar a lista de modelos probabilísticos no repositório.
Modelo Probabilístico	Identifica o modelo probabilístico a ser usado na operação.

## Propriedades da Tabela de Referência

As propriedades da tabela de referência aplicam-se quando você configura uma operação de rotulagem para usar uma tabela de referência.

A seguinte tabela descreve as propriedades da tabela de referência:

Propriedade	Descrição
Nome	Fornecer um nome para a operação.
Tabela de Referência	Especifica a tabela de referência usada pela operação para rotular tokens.
Rótulo	Especifica o texto que a operação grava em uma nova porta quando uma string de entrada corresponde a uma entrada da tabela de referência.
Distinção Entre Maiúsculas e Minúsculas	Determina se as strings de entrada devem corresponder à distinção entre maiúsculas e minúsculas das entradas da tabela de referência.
Substituir Correspondências por Valores Válidos	Substitui as strings rotuladas pela entrada da coluna Válido na tabela de referência.
Modo	Determina o método de rotulagem de token. Selecione Inclusivo para rotular strings de entrada que correspondem a entradas da tabela de referência. Selecione Exclusivo para rotular strings de entrada que não correspondem a entradas da tabela de referência.
Definir Prioridade	Determina se as operações de rotulagem da tabela de referência têm precedência sobre as operações de rotulagem do conjunto de tokens em uma estratégia. Se você definir esta propriedade, a transformação executará a rotulagem da tabela de referência antes da rotulagem do conjunto de tokens, e a análise do conjunto de tokens não poderá substituir a análise de rótulo da tabela de referência.

# Configurando uma Estratégia de Rotulagem de Caractere

Para configurar uma estratégia de rotulagem, edite as configurações na exibição **Estratégias** da transformação de Rotulador.

1. Selecione a exibição **Estratégias** e clique em **Novo** para criar uma estratégia.  
O assistente de **Estratégia** é aberto.
2. Insira um nome para a estratégia.
3. Clique nos campos **Entradas** e **Saídas** para definir portas para a estratégia.
4. Opcionalmente, insira uma descrição da estratégia.
5. Selecione o modo de rotulagem de caractere.
6. Clique em **Avançar**.
7. Selecione o tipo de operação de rotulagem de caractere que você deseja configurar. Você pode configurar as seguintes operações:
  - Rotular caracteres usando tabelas de referência.
  - Rotular caracteres usando conjuntos de caracteres.
8. Clique em **Avançar**.
9. Configure as propriedades da operação e clique em **Avançar**.
10. Opcionalmente, configure as propriedades de **Ignorar Texto**.
11. Clique em **Avançar** para adicionar mais operações à estratégia ou clique em **Concluir**.  
Você pode alterar a ordem em que a transformação processa estratégias e operações. Na exibição **Estratégias**, selecione uma estratégia ou operação e clique em **Mover para Cima** ou **Mover para Baixo**.

# Configurando uma Estratégia de Rotulagem de Token

Para configurar uma estratégia de rotulagem, edite as configurações na exibição **Estratégias** da transformação de Rotulador.

1. Selecione a exibição **Estratégias** e clique em **Novo** para criar uma estratégia.  
O assistente de **Estratégia** é aberto.
2. Insira um nome para a estratégia.
3. Clique nos campos **Entradas** e **Saídas** para definir portas para a estratégia.
4. Opcionalmente, insira uma descrição da estratégia.
5. Selecione o modo de rotulagem de token.  
Verifique ou edite as propriedades do modo selecionado.
6. Clique em **Avançar**.
7. Selecione o tipo de operação de rotulagem de token que você deseja configurar. Você pode configurar as seguintes operações:

- Rotular tokens com um conjunto de tokens.
  - Rotular tokens com uma tabela de referência.
  - Rotular tokens usando correspondência probabilística.
8. Clique em **Avançar**.
  9. Configure as propriedades da operação e clique em **Avançar**.  
Se você configurar a estratégia para usar correspondência probabilística, insira um rótulo a ser usado para tokens identificados como dados residuais.
  10. Opcionalmente, configure as propriedades de **Rótulo Personalizado**.
  11. Clique em **Avançar** para adicionar mais operações à estratégia ou clique em **Concluir**.  
Você pode alterar a ordem em que a transformação processa estratégias e operações. Na exibição **Estratégias**, selecione uma estratégia ou operação e clique em **Mover para Cima** ou **Mover para Baixo**.

## Propriedades Avançadas da Transformação de Rotulador

Configure propriedades que ajudam a determinar como o Serviço de Integração de Dados processa dados para a transformação de Rotulador.

Você pode configurar níveis de rastreamento para logs.

Configure a seguinte propriedade na guia **Avançado**:

### Nível de Rastreamento

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.

## Transformação de Rotulador em um ambiente não nativo

O processamento da transformação de Rotulador em um ambiente não nativo depende do mecanismo que executa a transformação.

Considere o suporte para os seguintes mecanismos de tempo de execução não nativos:

- Mecanismo Blaze. Suportado sem restrições.
- Mecanismo Spark. Com suporte sem restrições em mapeamentos em lote. Sem suporte em mapeamentos de fluxo.
- Mecanismo Databricks Spark. Sem suporte.



## CAPÍTULO 25

# Transformação de Pesquisa

Este capítulo inclui os seguintes tópicos:

- [Transformação de Pesquisa - Visão Geral, 405](#)
- [Pesquisas Conectadas e Não Conectadas, 406](#)
- [Desenvolvendo uma Transformação de Pesquisa, 409](#)
- [Consulta de Pesquisa, 409](#)
- [Filtro da Origem de Pesquisa, 412](#)
- [Condição de Pesquisa, 412](#)
- [Cache de Pesquisa, 415](#)
- [Propriedades de Consulta, 416](#)
- [Transformações de Pesquisa em Mapeamentos Dinâmicos, 416](#)
- [Definir Portas Dinâmicas, 417](#)
- [Alterar a Origem da Pesquisa, 417](#)
- [Seletores de Portas, 422](#)
- [Propriedades de Tempo de Execução, 428](#)
- [Propriedades avançadas, 429](#)
- [Criando uma Transformação de Pesquisa Reutilizável, 430](#)
- [Criando uma Transformação de Pesquisa Não Reutilizável, 431](#)
- [Criando uma Transformação de Pesquisa Não Conectada, 432](#)
- [Exemplo de Pesquisa Não Conectada, 433](#)
- [Transformação de Pesquisa em um ambiente não nativo, 435](#)

## Transformação de Pesquisa - Visão Geral

A transformação de Pesquisa é uma transformação passiva ou ativa que procura os dados em um arquivo simples, objeto de dados lógicos, tabela de referência ou tabela relacional. A transformação de Pesquisa pode retornar uma ou várias linhas de uma pesquisa.

Antes de criar uma transformação de Pesquisa, crie a origem da pesquisa. Importe um arquivo simples ou uma tabela de banco de dados relacional como um objeto de dados físicos. Ou crie um objeto de dados lógicos ou uma tabela de referência para usar como origem da pesquisa. Quando você cria uma transformação de Pesquisa, a Developer tool adiciona as colunas do objeto de dados ou da tabela de referência como portas de pesquisa na transformação. Depois de criar a transformação, configure uma ou

mais portas de saída para retornar os resultados da pesquisa. Configure as condições de pesquisa e outras propriedades de pesquisa.

Quando você executa um mapeamento ou visualização de dados, o Serviço de Integração consulta a origem da pesquisa. O Serviço de Integração consulta a origem da pesquisa com base nas portas de pesquisa da transformação, nas propriedades da pesquisa e na condição da pesquisa. A transformação de Pesquisa retorna o resultado da pesquisa para o destino ou outra transformação.

Você pode configurar uma transformação de Pesquisa conectada ou não conectada. Uma transformação conectada se conecta a outra transformação no mapeamento. Uma transformação não conectada recebe entrada de uma expressão: LKP em outra transformação. Se a transformação de Pesquisa executa uma pesquisa em um objeto de dados lógicos, é necessário configurar uma transformação de Pesquisa conectada. Conecte as portas de entrada da transformação de Pesquisa a uma transformação upstream ou a uma origem upstream. Conecte as portas de saída a uma transformação downstream ou a um destino downstream.

É possível usar várias transformações de Pesquisa em um mapeamento.

Você pode executar as seguintes tarefas com uma transformação de Pesquisa:

- Obter um valor relacionado. Obtenha um valor da origem da pesquisa com base em um valor nos dados de entrada. Por exemplo, os dados de entrada contêm uma ID de funcionário. Obtenha o nome do funcionário da origem de pesquisa por ID de funcionário.
- Recuperar várias linhas de uma tabela de pesquisa.
- Executar um cálculo. Recuperar um valor de uma tabela de pesquisa e usar o valor em um cálculo. Por exemplo, recupere uma porcentagem de imposto sobre vendas, calcule um imposto e retorne o imposto a um destino.
- Executar uma pesquisa não conectada com uma expressão: LKP em uma transformação que aceita expressões. Filtrar os resultados com outra expressão na transformação.
- Parametrize a fonte de pesquisa e a condição de pesquisa para usar uma transformação de Pesquisa em um mapeamento dinâmico.

## Pesquisas Conectadas e Não Conectadas

Você pode configurar uma transformação de Pesquisa conectada ou não conectada. Uma transformação de Pesquisa conectada é uma transformação que tem portas de entrada e saída que você conecta a outras transformações em um mapeamento. Uma transformação de Pesquisa não conectada aparece no mapeamento, mas não está conectada a outras transformações.

Uma transformação de Pesquisa não conectada recebe a entrada do resultado de uma expressão: LKP em uma transformação, como as transformações de Expressão ou de Agregador. A expressão :LKP transmite argumentos para a transformação de Pesquisa, que retorna um resultado. A expressão :LKP pode transmitir os resultados da pesquisa para outra expressão na transformação de Expressão ou de Agregador para filtrar os resultados.

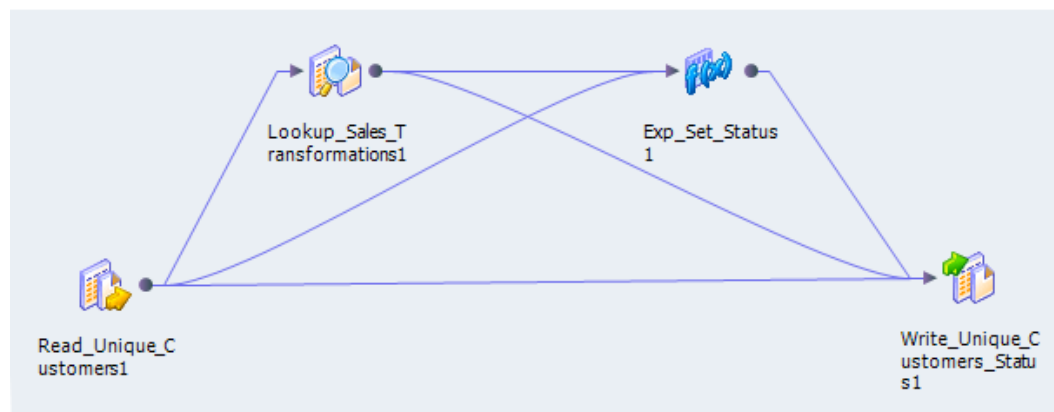
A tabela a seguir lista as diferenças entre pesquisas conectadas e desconectadas:

Pesquisa Conectada	Pesquisa Desconectada
Recebe os valores de entrada diretamente do pipeline.	Recebe os valores de entrada do resultado de uma expressão :LKP em outra transformação.
Usar um cache dinâmico ou estático.	Usar um cache estático.
O cache inclui as colunas de origem da pesquisa na condição de pesquisa e as colunas de origem da pesquisa que são portas de saída.	O cache inclui todas as portas de pesquisa e de saída na condição de pesquisa e a porta de pesquisa/retorno.
Retorna várias colunas da mesma linha ou insere linhas no cache de pesquisa dinâmica.	Retorna uma coluna de cada linha para uma porta de retorno.
Se não houver correspondência para a condição de pesquisa, o Serviço de Integração retornará o valor padrão para todas as portas de saída. Se você configurar o cache dinâmico, o Serviço de Integração vai inserir linhas no cache ou deixá-lo inalterado.	Se não houver correspondência para a condição de pesquisa, o Serviço de Integração retornará NULL.
Se houver correspondência para a condição de pesquisa, o Serviço de Integração retornará o resultado da condição de pesquisa para todas as portas de pesquisa/saída. Se você configurar o cache dinâmico, o Serviço de Integração atualizará a linha no cache ou a deixará inalterada.	Se uma correspondência ocorrer para a condição de pesquisa, o Serviço de Integração retornará o resultado da condição de pesquisa para a porta de retorno.
Transmite vários valores de saída para outra transformação. Vincular as portas de pesquisa/saída a outra transformação.	Retorna um valor de saída para outra transformação. A porta de retorno da transformação de Pesquisa transmite o valor para a porta que contém a expressão :LKP na outra transformação.
Suporta valores padrão definidos pelo usuário.	Não suporta valores padrão definidos pelo usuário.

## Pesquisas Conectadas

Uma transformação de Pesquisa conectada é uma transformação de Pesquisa que está conectada a uma origem ou destino em um mapeamento.

A seguinte figura mostra um mapeamento com uma transformação de Pesquisa conectada:



Quando você executa um mapeamento que contém uma transformação de Pesquisa conectada, o Serviço de Integração realiza as seguintes etapas:

1. O Serviço de Integração transmite valores de outra transformação para as portas de entrada na transformação de Pesquisa.
2. Para cada linha de entrada, o Serviço de Integração consulta o cache ou a origem da pesquisa com base nas portas de pesquisa e na condição de pesquisa na transformação.
3. Se a transformação não estiver armazenada no cache, ou se utilizar um cache estático, o Serviço de Integração retorna valores da consulta de pesquisa.

Se a transformação utilizar um cache dinâmico, o Serviço de Integração insere a linha no cache quando não encontrar a linha no cache. Quando o Serviço de Integração encontrar a linha no cache, ele a atualiza no cache ou a mantém sem alterações. Ele sinaliza a linha como inserção, atualização ou sem alterações.

4. O Serviço de Integração retorna dados da consulta e os transmite para a próxima transformação no mapeamento.

Se a transformação utilizar um cache dinâmico, é possível transferir linhas a uma transformação de filtro ou de roteador para filtrar novas linhas para o destino.

**Nota:** Este capítulo discorre sobre Transformações Pesquisa Conectada, salvo indicação em contrário.

## Pesquisas Não Conectadas

Uma transformação de Pesquisa não conectada é uma transformação de Pesquisa que não está conectada a uma origem ou destino no mapeamento. Chame a pesquisa com uma expressão: LKP em uma transformação que permita expressões.

A sintaxe para a expressão de pesquisa é :LKP lookup\_transformation\_name(argumento, argumento, ...)

A ordem em que você lista cada argumento deve corresponder à ordem das condições de pesquisa na transformação Pesquisa. A transformação de Pesquisa retorna o resultado da consulta por meio da porta de retorno da transformação de pesquisa. A transformação que chama a pesquisa recebe o valor do resultado da pesquisa na porta que contém a expressão: LKP. Se a consulta de pesquisa não conseguir retornar um valor, a porta receberá um valor nulo.

Ao executar uma pesquisa não conectada, você pode executar a mesma pesquisa várias vezes em um mapeamento. É possível testar os resultados da pesquisa em outra expressão e filtrar linhas com base nos resultados.

Quando você executa um mapeamento que contém uma transformação de Pesquisa não conectada, o Serviço de Integração executa as seguintes etapas:

1. Uma transformação de Pesquisa não conectada recebe valores de entrada do resultado de uma expressão :LKP em outra transformação, como uma transformação de Agregador, uma transformação de Expressão ou uma transformação de Estratégia de Atualização.
2. O Serviço de Integração consulta o cache ou a origem da pesquisa com base nas portas de pesquisa e na condição da transformação de Pesquisa.
3. O Serviço de Integração retorna um valor pela porta de retorno da transformação de Pesquisa.
4. O Serviço de Integração transmite esse valor de retorno à porta que contém a expressão :LKP.

# Desenvolvendo uma Transformação de Pesquisa

Ao desenvolver uma transformação de Pesquisa, é necessário considerar fatores como o tipo de origem da pesquisa e a condição de pesquisa.

Considere os seguintes fatores ao desenvolver uma transformação de Pesquisa:

- Se você quer criar a transformação a partir de um arquivo simples, um objeto de dados lógicos, uma tabela de referência ou um objeto de dados relacionais. Antes de criar uma transformação de Pesquisa, crie a origem da pesquisa. Importe um arquivo simples ou uma tabela de banco de dados relacional como um objeto de dados físicos. Ou crie um objeto de dados lógicos ou uma tabela de referência para usar como origem da pesquisa.
- As portas de saída para a transformação.
- As condições de pesquisa na transformação.
- Se você deseja que o Integration Service armazene os dados de pesquisa no cache. O Serviço de Integração pode armazenar dados em cache para arquivos simples, tabelas de referência ou objetos de dados relacionais.

## Consulta de Pesquisa

O Serviço de Integração consulta a pesquisa com base nas portas e nas propriedades que você configura na transformação de Pesquisa. O Serviço de Integração executa uma consulta de pesquisa padrão quando a primeira linha é inserida na transformação de Pesquisa.

Se você usar uma pesquisa em relação a uma tabela relacional, poderá substituir a consulta de pesquisa. É possível usar a substituição para adicionar uma cláusula WHERE ou transformar os dados de pesquisa antes que eles sejam armazenados em cache.

Se você configurar uma substituição SQL e um filtro na consulta de pesquisa, o Serviço de Integração ignorará esse filtro.

## Consulta de Pesquisa Padrão

A consulta de pesquisa padrão contém as seguintes instruções:

### **SELECT**

A instrução SELECT inclui todas as portas de pesquisa no mapeamento. Para exibir a instrução SELECT da consulta de pesquisa, selecione a propriedade Usar Consulta Personalizada.

### **ORDER BY**

A cláusula ORDER BY organiza as colunas na mesma ordem em que elas aparecem na transformação de Pesquisa. O Serviço de Integração gera a cláusula ORDER BY. Não é possível ver isso quando você gera o SQL padrão.

## Substituição SQL para uma Consulta de Pesquisa

É possível substituir a consulta de pesquisa para uma pesquisa relacional. Você pode adicionar uma cláusula WHERE e transformar os dados de pesquisa antes que eles sejam armazenados em cache.

É possível usar palavras reservadas e barras invertidas em nomes de tabela e de coluna.

Você pode inserir uma consulta para substituir completamente a consulta de pesquisa padrão. Ou exibir e editar a consulta de pesquisa padrão. A consulta de pesquisa padrão inclui as portas de pesquisa, de saída e de retorno.

Quando você usa uma substituição SQL, a consulta acrescenta a cláusula ORDER BY 1. A cláusula ordena os dados para fornecer os valores inicial e final de forma confiável para outras cláusulas.

**Nota:** Você pode validar manualmente o SQL executando a seguinte consulta em um utilitário de linha de comando do Hive:

```
CREATE VIEW <table name> (<port list>) AS <SQL>
```

onde:

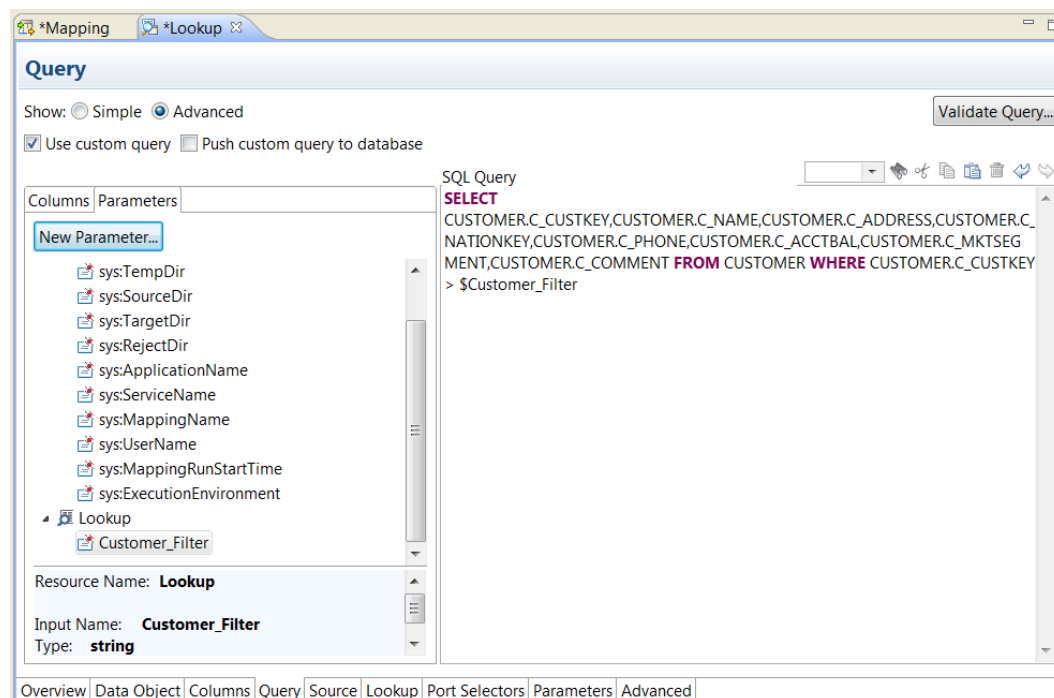
- <table name> é um nome da sua escolha
- <port list> é a lista de portas delimitadas por vírgula na origem
- <SQL> é a consulta para validar

## Parâmetros em uma Consulta de Substituição SQL

Você pode usar parâmetros do sistema ou parâmetros definidos pelo usuário na consulta de pesquisa de uma transformação de Pesquisa. O Editor SQL fornece uma lista dos parâmetros do sistema e dos parâmetros definidos pelo usuário que você pode selecionar.

É possível procurar ou criar parâmetros definidos pelo usuário na guia **Consulta** de uma transformação de Pesquisa. Defina um valor padrão para cada parâmetro. Você pode substituir um valor padrão vinculando um parâmetro de mapplet ou de mapeamento ao parâmetro de transformação depois de adicionar a transformação de Pesquisa a um mapeamento.

A seguinte imagem mostra a guia Consulta da transformação de Pesquisa:



## Palavras reservadas

Se qualquer nome de pesquisa ou de coluna contiver uma palavra reservada para o banco de dados, como MONTH ou YEAR, o mapeamento falhará com erros de banco de dados quando o Serviço de Integração executar o SQL nesse banco de dados.

Você pode criar e manter um arquivo de palavras reservadas, o reswords.txt, no diretório de instalação do Serviço de Integração. Quando o Serviço de Integração inicializa um mapeamento, ele pesquisa o arquivo reswords.txt, coloca as palavras reservadas entre aspas e executa o SQL nos bancos de dados de pesquisa, origem e destino.

Talvez seja necessário ativar alguns bancos de dados, como o Microsoft SQL Server e o Sybase, para usar padrões SQL-92 no que diz respeito a identificadores entre aspas. Use o SQL do ambiente para emitir o comando. Por exemplo, com o Microsoft SQL Server, use o seguinte comando:

```
SET QUOTED_IDENTIFIER ON
```

## Orientações para Substituir a Consulta de Pesquisa

Certas regras e diretrizes são aplicáveis quando você substitui uma consulta de pesquisa.

Considere as seguintes diretrizes ao substituir a consulta SQL de pesquisa:

- É possível substituir a consulta SQL de pesquisa para pesquisas relacionais.
- Adicione um filtro de pesquisa de origem para filtrar as linhas que são adicionadas ao cache de pesquisa. Isso garante que o Serviço de Integração insira ao cache dinâmico e à tabela de destino linhas que coincidam com a cláusula WHERE.
- Se várias transformações de Pesquisa compartilharem um cache de pesquisa, use a mesma substituição SQL de pesquisa para cada uma dessas transformações.
- Se um nome de tabela ou coluna na consulta de pesquisa contiver uma palavra reservada, coloque essa palavra entre aspas.
- Para substituir a consulta de pesquisa para uma pesquisa sem cache, opte por retornar qualquer valor quando o Serviço de Integração encontrar várias correspondências.
- Não é possível adicionar ou excluir nenhuma das colunas da instrução SQL padrão.
- A Developer tool não valida a sintaxe da consulta SQL. Se a substituição SQL em uma consulta de pesquisa desconectada não for válida, o mapeamento falhará.

## Substituindo a Consulta de Pesquisa

Você pode substituir a consulta SQL de pesquisa padrão de forma a criar uma consulta personalizada na origem de pesquisa.

1. Na exibição **Propriedades**, selecione a guia **Consulta**.
2. Selecione **Avançado**.
3. Selecione **Usar Consulta Personalizada**.
4. Edite a consulta de pesquisa na área Consulta SQL.  
É possível clicar duas vezes em um nome de tabela, em um nome de coluna ou em um parâmetro para adicioná-lo à consulta.
5. Clique em **Validar Consulta** para validar a consulta de pesquisa.
6. Selecione **Enviar Consulta Personalizada ao Banco de Dados** para executar a consulta de pesquisa no banco de dados.

## Filtro da Origem de Pesquisa

Você pode configurar um Filtro da Origem de Pesquisa para uma transformação Pesquisa relacional que tenha o cache ativado. Adicione o filtro de origem de pesquisa para limitar o número de pesquisas que o Serviço de Integração realiza em uma tabela de origem da pesquisa.

Quando você configurar um Filtro da Origem de Pesquisa, o Serviço de Integração executa pesquisas com base nos resultados da instrução do filtro. Por exemplo, talvez você precise recuperar o sobrenome de cada funcionário com um ID maior que 510.

Você configura o seguinte filtro de origem de pesquisa na coluna EmployeeID:

```
EmployeeID >= 510
```

EmployeeID é uma porta de entrada na transformação Pesquisa. Quando o Serviço de Integração lê a linha de origem, ele executa uma pesquisa no cache quando o valor de EmployeeID é maior que 510. Quando EmployeeID for menor ou igual a 510, a transformação Pesquisa não recupera o sobrenome.

Quando você adiciona um filtro de origem de pesquisa à consulta de Pesquisa para um mapeamento configurado para a otimização de empilhamento, o Serviço de Integração cria uma exibição para representar a substituição SQL. O Serviço de Integração executa uma consulta de SQL contra esta exibição, para forçar a lógica de transformação para o banco de dados.

## Filtrando Linhas de Origem em uma Pesquisa

Você pode configurar um Filtro da Origem de Pesquisa para uma transformação Pesquisa relacional que tenha o cache ativado. Filtre linhas de origem em uma pesquisa para limitar o número de pesquisas realizadas pelo Serviço de Integração em uma tabela de origem de pesquisa.

1. Na exibição **Propriedades**, selecione a guia **Consulta**.
2. Na opção **Filtro**, clique em **Editar**.
3. No SQL Editor, selecione as portas de entrada ou insira qualquer porta da transformação Pesquisa que você deseja filtrar.
4. Insira uma condição de filtro.  
Não inclua a palavra-chave WHERE na condição do filtro. Inclua os parâmetros e as variáveis de mapeamento de string nos identificadores de string.
5. Clique em **Validar Consulta** para validar a sintaxe da condição de filtro.

## Condição de Pesquisa

O Serviço de Integração de Dados procura dados na origem de pesquisa com base em uma condição de pesquisa. Ao configurar uma condição de pesquisa em uma transformação de Pesquisa, você compara o valor de uma ou mais colunas nos dados de origem com os valores na origem da pesquisa ou cache.

Por exemplo, os dados de origem contêm um employee\_number. A tabela de origem da pesquisa contém employee\_ID, first\_name e last\_name. Você configura a seguinte condição de pesquisa:

```
employee_ID = employee_number
```

Para cada employee\_number, o Serviço de Integração de Dados retorna a coluna employee\_ID, last\_name e first\_name a partir da origem da pesquisa.



O Serviço de Integração de Dados pode retornar mais de uma linha da origem da pesquisa. Você configura a seguinte condição de pesquisa:

```
employee_ID > employee_number
```

O Serviço de Integração de Dados retorna linhas para todos os números employee\_ID maiores que o número do funcionário de origem.

### Valores Nulos em uma Pesquisa de Objeto de Dados

Quando uma entrada para a condição de Pesquisa é NULL, uma transformação de Pesquisa de objeto de dados retorna uma única linha com valores nulos para portas somente de saída e valores de linhas de entrada para portas de passagem.

Por exemplo, a seguinte condição de pesquisa realiza uma pesquisa em uma fonte de dados que contém uma ou mais linhas nas quais o valor para employee\_ID é NULL:

```
employee_ID = employee_number
```

Nesse exemplo, use uma tabela de pesquisa com os seguintes dados:

EMPLOYEE_ID	LAST_NAME
1294765	Hara
1356356	Carver
1407207	NULL
1570348	Draper
NULL	Limonov

Você compara os seguintes valores de entrada da sua fonte de dados com a tabela de pesquisa:

```
EMPLOYEE_NUMBER  
-----  
1294765  
1356356  
1407207  
1648246  
NULL
```

Neste exemplo, a condição de Pesquisa gera os seguintes resultados:

```
1294765,Hara  
1356356,Carver  
1407207,NULL  
NULL,NULL  
NULL,NULL
```

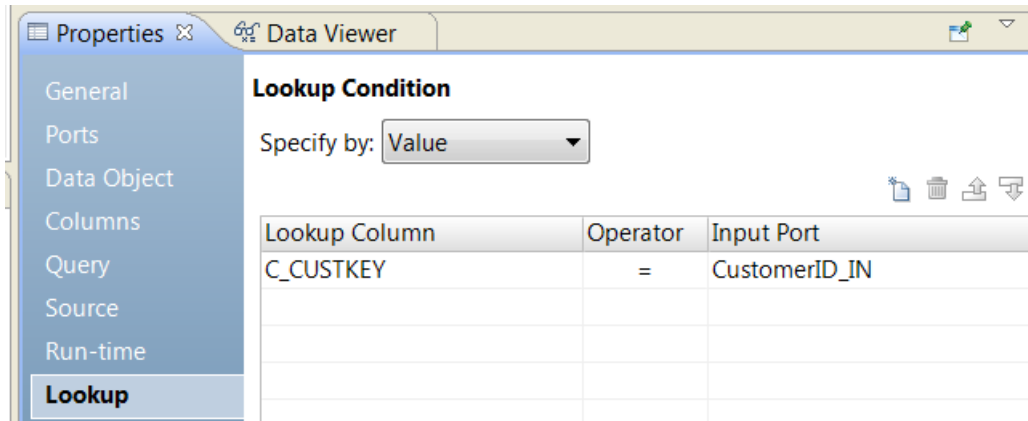
A condição de Pesquisa encontra correspondências entre EMPLOYEE\_ID e EMPLOYEE\_NUMBER para as duas primeiras linhas. Para a terceira fila, a fonte de pesquisa contém uma linha com um valor NULL que não participa da condição de pesquisa. Ela corresponde a condição de pesquisa e retorna um resultado com o valor NULL para a coluna que não é de pesquisa.

Para as quarta e quinta linhas, a condição de Pesquisa não encontra uma correspondência e retorna NULL para ambos os valores. Para a quinta linha, observe que a condição de Pesquisa não encontra uma correspondência novamente, pois NULL não corresponde nada, inclusive NULL.

## Configurar a Condição de Pesquisa

A condição de Pesquisa é uma expressão que determina quais linhas devem ser recuperadas da origem de pesquisa. Configure a condição de pesquisa na guia **Pesquisa** da exibição **Propriedades**.

A imagem a seguir mostra uma condição de pesquisa para realizar uma pesquisa por número de cliente:



É possível configurar as seguintes opções na guia **Pesquisa**:

### Especificar por

Escolha **Valor** para selecionar os nomes da coluna de pesquisa e da porta de entrada. Selecione **Parâmetro** para configurar um parâmetro de expressão de forma a definir a condição de pesquisa.

### Coluna de pesquisa

A coluna na origem de pesquisa a ser correspondida com uma coluna na linha de entrada. É possível incluir várias colunas na condição de pesquisa.

### Operador

O operador que determina a condição a ser procurada entre a coluna de pesquisa e a porta de entrada. Os operadores incluem =, !=, >, <, >=, <=.

### Porta de Entrada

A porta de entrada que contém o valor a ser procurado na origem de pesquisa. Você pode comparar mais de uma porta de entrada com as portas na origem de pesquisa.

## Regras e Diretrizes para Condições de Transformação de Pesquisa

Determinadas regras e diretrizes aplicam-se quando você insere uma condição para uma transformação de Pesquisa.

Considere as seguintes regras e diretrizes ao inserir uma condição para uma transformação de Pesquisa.

- Os tipos de dados das colunas em uma condição de pesquisa devem coincidir.
- Use uma única porta de entrada para cada porta de pesquisa na condição de pesquisa. É possível usar a mesma porta de entrada em mais de uma condição em uma transformação.
- Se você usar um seletor de portas ou uma porta dinâmica em uma condição de pesquisa, esta última levará em consideração todas as portas da expressão.
- É possível usar uma porta de entrada dinâmica ou um seletor de portas como a porta de entrada de uma condição de pesquisa. O número de portas geradas na porta de entrada deve ser igual ao número de portas na coluna de pesquisa.

- Durante o processamento de uma transformação de Pesquisa com várias condições de pesquisa, o Serviço de Integração retorna linhas que correspondem a todas as condições de pesquisa.
- É possível criar um parâmetro de expressão para parametrizar a condição de pesquisa em uma transformação de pesquisa não reutilizável.
- Para melhorar o desempenho da pesquisa, insira as condições na seguinte ordem:
  - Igual a (=)
  - Menor que (<), maior que (>), menor que ou igual a (<=), maior que ou igual a (>=)
  - Diferente de (!=)
- Use um dos seguintes operadores ao criar uma condição de pesquisa: =, >, <, >=, <= ou !=.
- O Serviço de Integração processa correspondências da pesquisa de forma diferente dependendo de você configurar a transformação de Pesquisa com um cache de pesquisa dinâmica, um cache de pesquisa estática ou uma pesquisa sem cache.
- O Serviço de Integração corresponde valores nulos. Por exemplo, o Serviço de Integração determinará que uma porta de pesquisa é igual a uma porta de entrada se ambas tiverem valores nulos.
- Se as colunas em uma condição de pesquisa forem tipos de dados Decimais, a precisão de cada coluna deverá pertencer ao mesmo intervalo de precisão. Os intervalos de precisão válidos incluem:
  - Decimais 0-18
  - Decimais 19-28
  - Decimais 29-38
  - Decimais 39 e acima

Por exemplo, se você definir a condição `DecimalA = DecimalB`, em que `DecimalA` tem uma precisão de 15 e `DecimalB` tem uma precisão de 25, a condição de pesquisa não será válida.

## Cache de Pesquisa

Você pode melhorar o desempenho armazenando em cache uma origem de pesquisa grande ou tabelas de pesquisa pequenas. Quando você armazena em cache a origem de pesquisa, o Serviço de Integração consulta o cache de pesquisa em vez da origem de pesquisa para cada linha de entrada.

Dependendo dos seus requisitos comerciais, você pode criar diferentes tipos de caches de pesquisa. Você pode criar um cache estático ou dinâmico. É possível criar um cache persistente ou não persistente. Você pode compartilhar um cache entre várias transformações de Pesquisa.

Se a transformação de Pesquisa estiver em um mapeamento dinâmico, você poderá ter um cache persistente ou não persistente. Quando você persiste um cache e muda a origem de pesquisa com um parâmetro, o mapeamento falha. O mapeamento também falhará se você alterar o arquivo de controle para uma origem de pesquisa de arquivo simples.

**Nota:** Não é possível usar um cache de pesquisa dinâmica ou um cache de pesquisa persistente quando a transformação de Pesquisa contém uma porta dinâmica ou uma origem de pesquisa parametrizada.

## Propriedades de Consulta

Configure propriedades de consulta para exibir ou alterar a consulta de pesquisa em uma tabela de pesquisa relacional. Você pode aplicar um filtro na pesquisa ou personalizar a consulta de pesquisa.

A tabela a seguir descreve propriedades de consulta para transformações Pesquisa que executam pesquisas relacionais:

Propriedade	Descrição
Simples	Selecione para exibir a consulta de pesquisa e aplicar um filtro na pesquisa.
Avançado	Selecione para exibir a consulta, personalizá-la ou executá-la no banco de dados que contém a tabela de pesquisa relacional.
Filtro	Insira um filtro para reduzir o número de linhas que o Serviço de Integração consulta. Você deve selecionar a opção <b>simples</b> para exibir essa opção.
Usar Consulta Personalizada	Selecione para substituir a consulta de pesquisa. Você deve selecionar a opção <b>Avançado</b> para exibir essa opção.
Enviar Consulta Personalizada ao Banco de Dados	Selecione para executar a consulta no banco de dados que contém a tabela de pesquisa relacional. Você deve selecionar a opção <b>Avançado</b> para exibir essa opção.
Consulta SQL	Exibe a consulta SQL usada para executar a pesquisa. Você pode personalizar a consulta SQL. Você deve selecionar a opção <b>Avançado</b> para exibir essa opção.

## Transformações de Pesquisa em Mapeamentos Dinâmicos

É possível usar uma transformação de Pesquisa em um mapeamento dinâmico. Você pode configurar portas dinâmicas para receber e retornar portas diferentes com base nos dados de origem. A origem e a condição de pesquisa podem ser parametrizadas para que uma pesquisa seja realizada com base em portas diferentes.

Um mapeamento dinâmico é um mapeamento no qual as origens, os destinos e a lógica de transformação podem mudar em tempo de execução. É possível definir parâmetros e regras para alterar a estrutura dos dados. Quando você usa uma transformação de Pesquisa em um mapeamento dinâmico, as portas de entrada dessa transformação podem mudar de acordo com os dados de origem. A estrutura da origem de pesquisa e das portas na condição de pesquisa pode mudar.

**Nota:** Quando a transformação de Pesquisa contém uma porta dinâmica ou uma origem de pesquisa parametrizada, não é possível persistir o cache de pesquisa. Também não é possível configurar um cache dinâmico.

As seguintes tarefas podem ser realizadas para uma transformação de Pesquisa de forma a usá-la em um mapeamento dinâmico:

### Definir portas dinâmicas

Defina portas dinâmicas e portas geradas para acomodar mudanças nas colunas de entrada.

### Parametrizar a origem de pesquisa

Atribua um parâmetro para o objeto de dados que define a origem de pesquisa. É possível parametrizar a origem de pesquisa em uma transformação de Pesquisa não reutilizável.

### Definir seletores de portas

Defina um seletor de portas que especifique as portas a serem usadas na condição de pesquisa. Você pode parametrizar as portas do seletor de portas em uma transformação de Pesquisa não reutilizável.

### Parametrizar a condição de pesquisa

Crie um parâmetro de expressão e defina um valor padrão que contenha uma expressão completa.

Para obter mais informações sobre mapeamentos dinâmicos, consulte o *Guia de Mapeamentos do Informatica Developer*.

## Definir Portas Dinâmicas

É possível definir portas dinâmicas na transformação de Pesquisa.

Você pode fazer referência a uma porta dinâmica na coluna de entrada de uma condição de pesquisa. Se a porta dinâmica contiver várias portas geradas, você poderá usar um seletor de porta para o elemento de coluna de pesquisa da condição de pesquisa. A porta de entrada dinâmica deve conter o mesmo número de portas que o seletor de portas na condição de pesquisa.

Se a porta dinâmica contiver um valor, você poderá usar uma única porta no elemento da coluna de pesquisa da condição de pesquisa.

É possível fazer referência a portas geradas na condição de pesquisa. No entanto, se uma origem no mapeamento dinâmico mudar, a porta gerada talvez não existirá. O mapeamento falhará.

## Alterar a Origem da Pesquisa

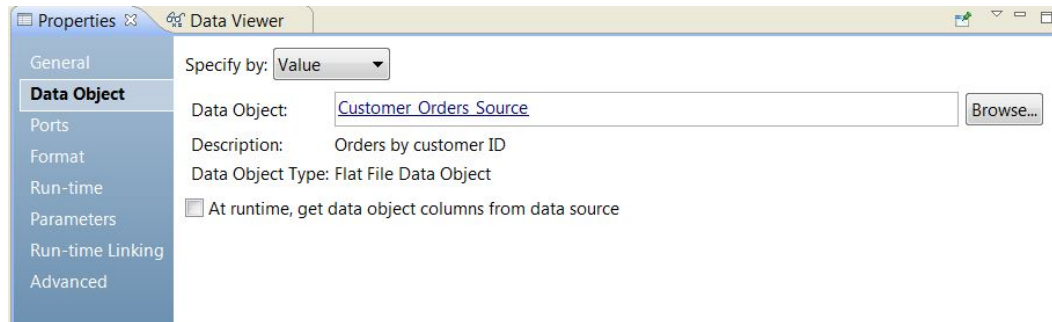
Você pode alterar o objeto de dados que é a origem de pesquisa de uma transformação de Pesquisa reutilizável. É possível configurar um parâmetro que determina o objeto de dados a ser usado como origem de pesquisa em tempo de execução.

Quando você cria uma transformação de um objeto de dados físicos, informações sobre esse objeto de dados são exibidas na guia **Objeto de Dados** das propriedades da transformação. Você pode clicar no nome desse objeto de dados para exibir a definição do objeto de dados físicos no repositório do Modelo.

É possível alterar o objeto de dados da transformação procurando um objeto de dados físicos diferente no repositório do Modelo. Quando você altera o objeto de dados, a transformação usa as propriedades de tempo de execução e as propriedades avançadas do objeto de dados selecionado.

É possível atualizar a estrutura do objeto de dados em tempo de execução com base em alterações na fonte de dados. A fonte de dados é o arquivo físico ou a tabela de banco de dados que o objeto de dados representa. Quando você permite que o Serviço de Integração de Dados obtenha colunas de dados da fonte de dados, o Serviço de Integração de Dados examina a estrutura da fonte de dados. O Serviço de Integração de Dados atualiza as portas de objetos de dados na instância de transformação com base na fonte de dados. O Serviço de Integração de Dados não altera a definição do objeto de dados físicos no repositório do Modelo.

A seguinte imagem mostra a guia **Objeto de Dados**:



A guia **Objeto de Dados** possui os seguintes campos:

#### Especificar por

Escolha **Valor** para inserir um nome de objeto de dados específico. Escolha **Parâmetro** para parametrizar o objeto de dados.

#### Objeto de dados

O nome do objeto de dados no repositório do Modelo. Você pode clicar no link **Objeto de Dados** para abrir a definição de objeto de dados do repositório. Você também pode procurar um objeto de dados diferente no repositório do Modelo.

#### Descrição

A descrição do objeto de dados no repositório. Somente leitura.

#### Tipo de objeto de dados

Descreve o tipo de objeto de dados, como um objeto de dados de arquivo simples, um objeto de tabela relacional ou um objeto de dados personalizado.

#### Em tempo de execução, obter colunas de objetos da fonte de dados

O Serviço de Integração de Dados busca alterações de metadados e definição de dados do arquivo de dados ou da tabela a que o objeto de dados se refere e atualiza a estrutura do objeto de dados para a instância de transformação no tempo de execução.

Para visualizar como o Serviço de Integração de Dados busca as alterações de metadados e definição de dados no tempo de execução, visualize o mapeamento com parâmetros resolvidos.

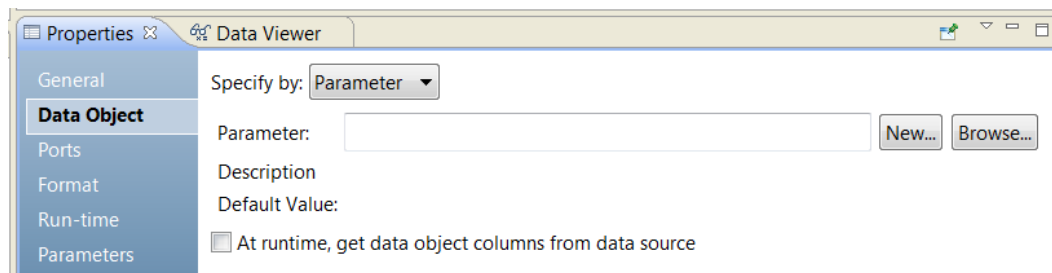
## Parametrizar a Origem de Pesquisa

Você pode configurar um parâmetro para a Origem de pesquisa em uma transformação de Pesquisa não reutilizável.

Para parametrizar um objeto de dados, escolha **Especificar por Parâmetro** na guia **Objeto de Dados**. As propriedades na guia **Objeto de Dados** são modificadas.

Para parametrizar o objeto de dados, crie um parâmetro de tipo de recurso ou procure um parâmetro de recurso que você já tenha criado. O valor padrão do parâmetro é o nome do objeto de dados físicos no repositório do Modelo. Ao criar um valor de parâmetro padrão, você seleciona um nome de objeto de dados físicos em uma lista de objetos de dados no repositório.

A seguinte imagem mostra a guia **Objeto de Dados** quando você especifica o objeto de dados por um parâmetro:



A guia **Objeto de Dados** tem as seguintes opções por parâmetro:

**Parâmetro**

O nome de um parâmetro de recurso que você configurou como o objeto de dados. Somente leitura.

**Descrição**

A descrição do parâmetro. Somente leitura.

**Novo**

Crie um parâmetro de recurso. Procure e selecione um objeto de dados no repositório do Modelo para o valor padrão do parâmetro.

**Procurar**

Procure um parâmetro de recurso e selecione esse parâmetro.

**Valor padrão**

O valor padrão do parâmetro de recurso que você configurou para o objeto de dados. O valor padrão é um nome de objeto de dados físicos e o caminho para o objeto no repositório do Modelo. Somente leitura.

## Conflitos de nomes de porta com a porta de pesquisa

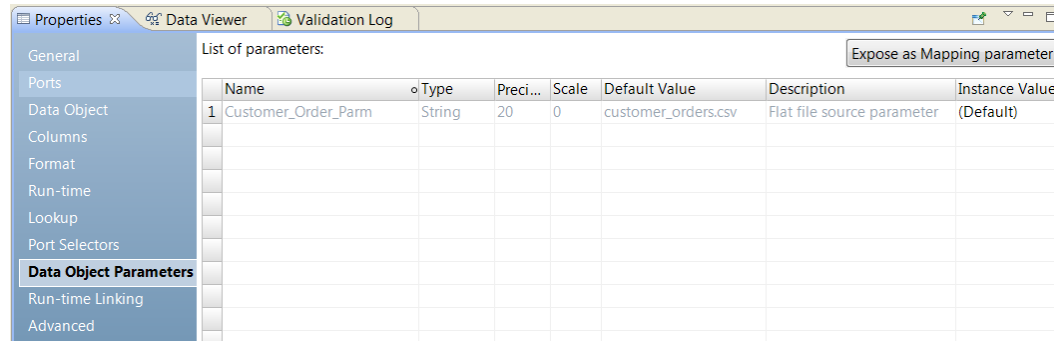
Quando você parametriza uma origem de pesquisa, uma porta de entrada de transformação de Pesquisa pode ter um conflito de nome com uma porta nessa origem.

Quando uma porta de entrada de transformação de Pesquisa tem um conflito de com a porta de pesquisa na origem de pesquisa, a Developer tool não renomeia nenhuma das portas. A Developer tool exibe um erro de validação. Você deve alterar o nome da porta de entrada na transformação de Pesquisa ou remover a porta da transformação.

## Origens de Pesquisa que Contêm Parâmetros

Você pode criar uma origem de pesquisa a partir de um objeto de dados físicos que contém parâmetros. Ao adicionar o objeto de dados físicos a um mapeamento, os parâmetros aparecem na guia **Parâmetros do Objeto de Dados**.

A seguinte imagem mostra a guia **Parâmetros do Objeto de Dados** na transformação de Pesquisa:



A imagem mostra Customer\_Order\_Parm. Customer\_Order\_Parm é um parâmetro para o nome do arquivo de origem em um objeto de dados de arquivo simples. Para substituir o nome do arquivo de origem no mapeamento, associe Customer\_Order\_Parm a um parâmetro no mapeamento. Clique em **Expor como Parâmetro de Mapeamento** para criar um parâmetro duplicado no mapeamento.

## Configurar Parâmetros em um Objeto de Dados Duplicado

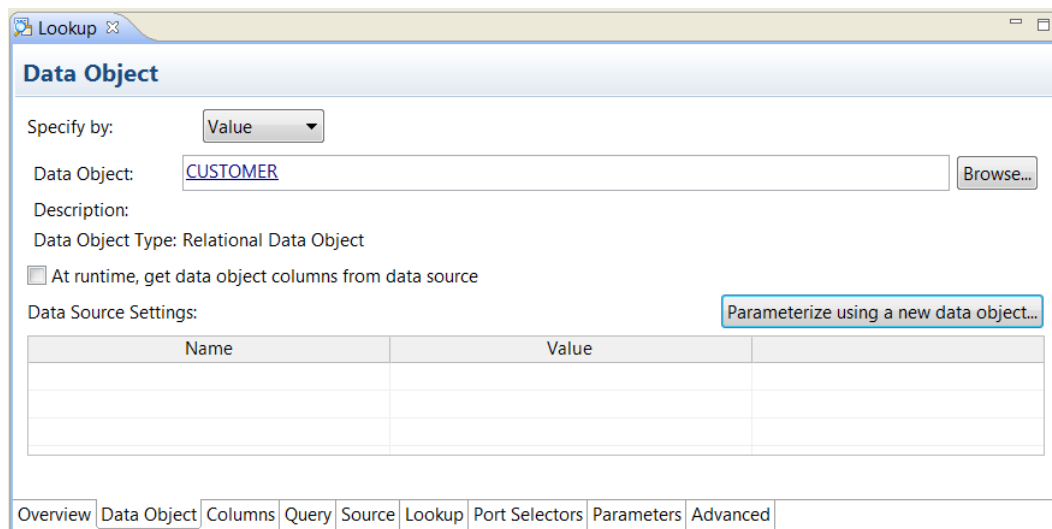
Você pode criar um objeto de dados duplicado no repositório e parametrizar as propriedades desse objeto de dados duplicado. Defina valores padrão para as propriedades, como a conexão, o nome do recurso, o proprietário da tabela ou o nome do arquivo de controle.

Você pode criar um objeto de dados duplicado para objetos de dados relacionais e objetos de dados de arquivo simples. Um objeto de dados duplicado pode ser criado em uma transformação de Pesquisa reutilizável e em uma transformação de Pesquisa não reutilizável.

Crie um objeto de dados duplicado na guia **Objeto de Dados** da transformação de Pesquisa. Você poderá criar um objeto duplicado se especificar o objeto de dados como um valor. Ao criar o objeto de dados duplicado, você substitui o nome do objeto de dados na transformação de Pesquisa pelo nome desse objeto de dados duplicado. A Developer tool cria parâmetros para as propriedades do objeto de dados. A Developer tool solicita a especificação de valores padrão para os parâmetros. A sintaxe do nome do objeto de dados duplicado é: <Nome do objeto original>\_Param.

A seguinte imagem mostra o botão **Parametrizar Usando Novo Objeto de Dados** na guia **Objeto de Dados** para objetos de dados relacionais e de arquivo simples:

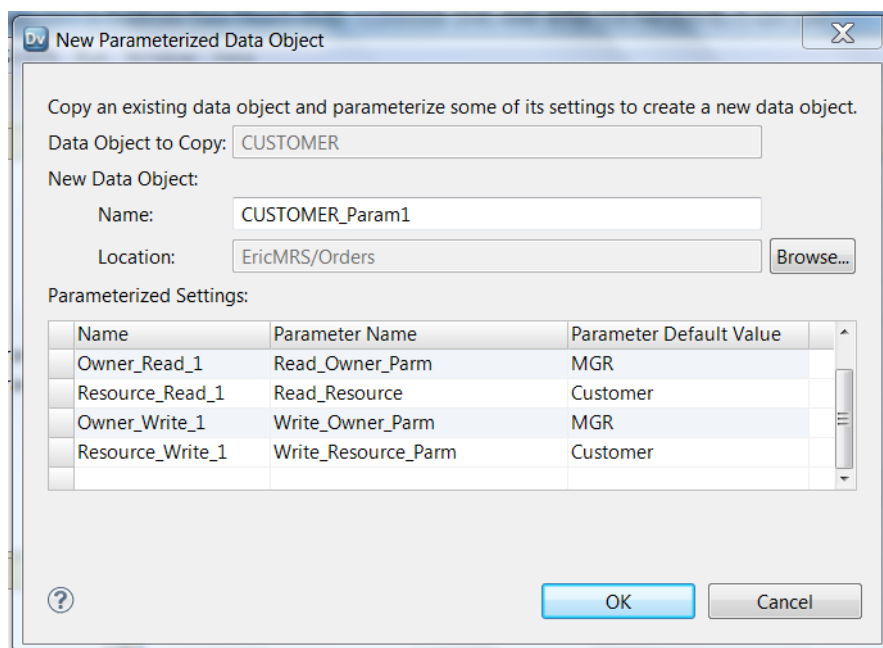




Quando você parametriza um objeto de dados criando um objeto de dados duplicado, a Developer tool cria um conjunto de parâmetros para esse objeto de dados. A Developer tool cria diferentes parâmetros com base em se o objeto de dados é um objeto de dados de arquivo simples ou relacional.

Ao criar o objeto de dados duplicado, configure os valores de parâmetros padrão na caixa de diálogo **Novo Objeto de Dados Parametrizado**.

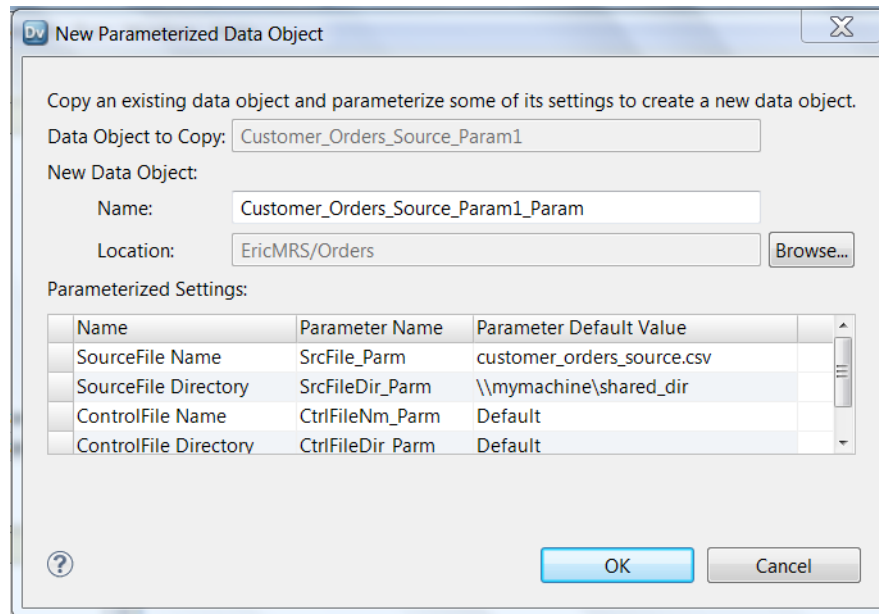
A seguinte imagem mostra o **Novo Objeto de Dados Parametrizado** para um objeto de dados relacionais:



Você pode alterar o nome do objeto de dados. Insira os valores padrão para os parâmetros de proprietário e recurso.

Se o objeto de dados original for parametrizado, a Developer tool copiará os parâmetros desse objeto para o objeto de dados duplicado. Se uma propriedade original não for parametrizada, a Developer tool criará um parâmetro para ela no objeto de dados duplicado. A Developer tool usa o valor da propriedade original como o valor do parâmetro padrão no objeto de dados duplicado. Quando não consegue determinar um valor de propriedade original, ela cria um parâmetro com um valor padrão com base no tipo de parâmetro.

A seguinte imagem mostra a caixa de diálogo **Novo Objeto de Dados Parametrizado** para um objeto de dados de arquivo simples:



Configure valores de parâmetro padrão para o arquivo de origem e seu respectivo diretório. Se um arquivo simples tiver um arquivo de controle, configure o nome e o diretório desse arquivo de controle.

Depois de configurar os valores padrão, a Developer tool cria o objeto de dados duplicado. O nome do objeto de dados duplicado aparece na guia **Objeto de Dados** da transformação de Pesquisa. O objeto de dados duplicado aparece no **Navegador de Objetos**.

Para alterar os valores de parâmetros para o objeto de dados depois de criá-lo, abra o objeto de dados físicos no **Navegador de Objetos**. Clique na guia **Parâmetros**.

## Seletores de Portas

É possível criar uma condição de pesquisa quando a transformação de Pesquisa contém portas geradas. Você pode fazer referência a uma porta dinâmica ou a um seletor de portas em uma condição de pesquisa. Também pode usar um parâmetro de expressão para parametrizar a expressão de pesquisa completa.

Quando a porta dinâmica contém várias portas geradas, é possível definir um seletor de portas para filtrar as portas geradas na condição de pesquisa. A origem de pesquisa pode mudar em um mapeamento dinâmico. Você pode configurar um seletor de portas para filtrar as portas que devem ser usadas para a coluna de pesquisa. O seletor de portas da origem de pesquisa deve conter o mesmo número de portas que o seletor de portas das colunas de entrada.

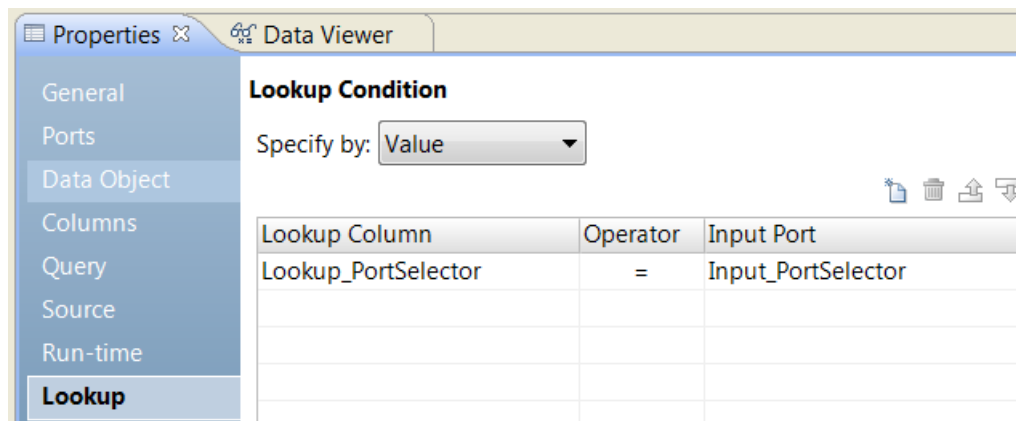
Por exemplo, Lookup\_PortSelector contém as seguintes portas:

```
C_CustKey
C_OrderKey
```

Input\_PortSelector contém as seguintes portas de entrada:

```
CustomerID_IN
OrderID_IN
```

A seguinte imagem mostra uma condição de pesquisa que contém seletores de portas:



A condição de pesquisa se expande para a seguinte expressão:

```
C_CustKey = CustomerID_IN AND C_OrderKey = OrderID_IN
```

Quando a condição de pesquisa contém várias portas, é possível configurar um operador. Por exemplo, você pode mudar o operador para maior que (>). A condição de pesquisa se expande para a seguinte expressão:

```
C_CustKey > CustomerID_IN AND C_OrderKey > OrderID_IN
```

Você pode criar uma condição de pesquisa que contenha uma porta dinâmica:

```
Lookup_PortSelector = Dynamic_Input_Port
```

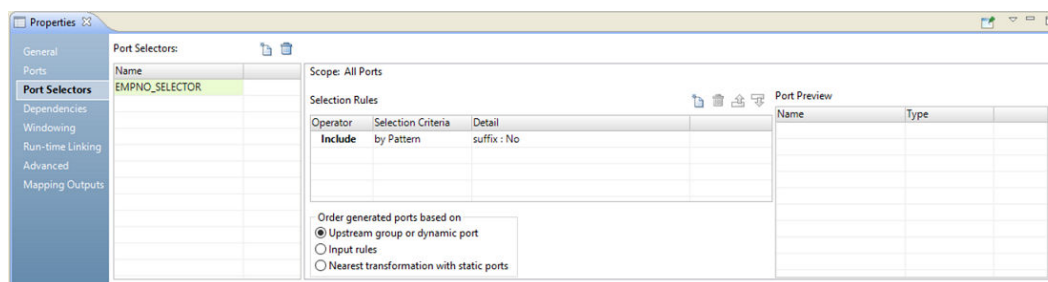
A porta dinâmica deve conter o mesmo número de portas que o seletor de portas.

## Configuração do seletor de portas

Ao configurar um seletor de portas, você define regras de seleção para determinar quais portas geradas devem ser incluídas. As regras de seleção são semelhantes às regras de entrada que você pode configurar para portas dinâmicas.

Um seletor de portas pode incluir portas estáticas ou portas geradas. Configure um seletor de portas na guia **Seletor de Portas**.

A seguinte imagem mostra a guia **Seletor de Portas**:



Configure as seguintes propriedades para um seletor de portas:

### Nome

Identifica o seletor de portas. Você pode criar vários seletores de portas em uma transformação e fazer referência a eles em expressões.

### Escopo

Identifica um grupo de portas ao qual o seletor de portas se aplica. Você deve escolher o escopo ao criar um seletor de portas para uma transformação de Associador ou Pesquisa. Essas transformações têm vários grupos de entrada. A transformação de Associador tem um escopo Mestre ou Detalhado. A transformação de Pesquisa tem um escopo de Importação ou Pesquisa. A transformação de Expressão tem um grupo de entrada. O escopo é sempre Todas as Portas.

### Regras de Seleção

Determina as portas a serem incluídas no seletor de portas. Quando você cria as regras de seleção, o painel **Visualização de Portas** mostra as portas qualificadas a partir das portas de entrada atuais. Essas portas podem mudar. Configure as regras de seleção para acomodar portas de diferentes origens.

## Regras de Seleção

As regras de seleção associadas a um seletor de portas determinam as portas a serem incluídas no seletor de portas.

Quando você cria as regras de seleção, o painel **Visualização de Portas** mostra as portas qualificadas a partir das portas de entrada atuais. Essas portas podem mudar. Configure as regras de seleção para acomodar portas de diferentes origens.

Crie regras de seleção com base nos critérios a seguir:

#### Operador

Inclui ou exclui as portas que são retornadas pelas portas de seleção. O padrão é incluir. Você deve incluir portas antes de poder excluí-las.

#### Critérios de Seleção

O tipo de regra de seleção que você deseja criar. Você pode criar uma regra com base no nome da coluna, tipo de porta, padrão ou definição de tipo de dados complex. Para incluir portas com base no nome da coluna, procure nomes específicos ou procure um padrão de caracteres no nome.

#### Detalhe

Os valores a serem aplicados aos critérios de seleção. Se os critérios de seleção forem por nome de coluna, configure a cadeia ou o nome a ser pesquisado. Se os critérios de seleção forem por tipo de porta, selecione os tipos de porta a serem incluídos.

A seguinte tabela descreve os critérios de seleção e como especificar os detalhes dos critérios:

Critérios de Seleção	Descrição	Detalhe
Tudo	Inclui todas as portas.	Nenhum detalhe é necessário.
Nome	Filtra portas com base no nome da porta.	Selecione os nomes de portas de uma lista de valores ou use um parâmetro do tipo Porta ou Lista de Portas.
Tipo	Filtra portas com base no tipo de dados de cada uma.	Selecione tipos de dados de uma lista.

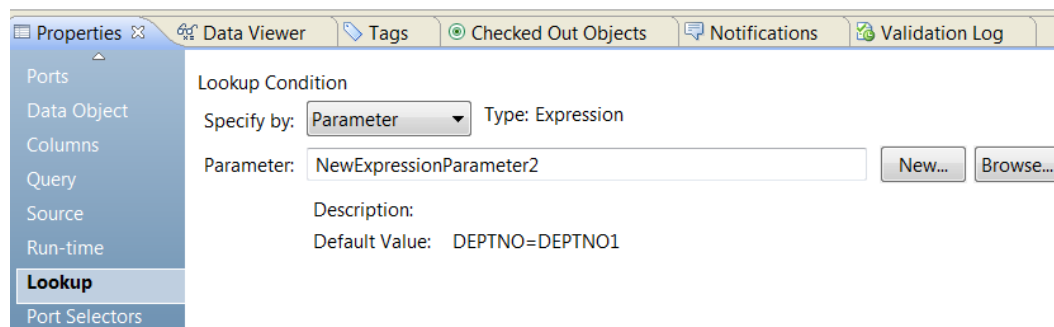
Critérios de Seleção	Descrição	Detalhe
Padrão	Filtra portas por uma cadeia de caracteres no nome ou por uma expressão regular.	Escolha o prefixo, o sufixo ou a expressão regular como o tipo de padrão para o nome da porta. Em seguida, insira um valor para o padrão ou use um parâmetro do tipo Cadeia.
Definição de Tipo de Dados Complex	Filtra portas por uma definição de tipo de dados complex.	Escolha prefixo, sufixo ou expressão regular como o tipo de padrão para a definição de tipo de dados complex. Em seguida, insira um valor para o padrão ou use um parâmetro do tipo Cadeia.

## Parametrizar a Condição de Pesquisa

É possível configurar um parâmetro de expressão que define a condição de pesquisa. Um parâmetro de expressão contém uma expressão completa que você cria em um editor de expressão. Você pode definir um parâmetro de mapeamento para substituir o parâmetro de expressão em tempo de execução.

Se você especificar uma condição de pesquisa usando um parâmetro, poderá procurar um parâmetro de expressão ou criar um parâmetro.

A seguinte imagem mostra onde configurar um parâmetro de expressão para a condição de Pesquisa:



Para criar um parâmetro, clique em **Novo**. Defina um nome para o parâmetro e edite o valor padrão. O valor padrão do parâmetro de expressão é a expressão completa para definir a condição de pesquisa. Você pode usar portas geradas, portas dinâmicas e seletores de portas na expressão.

**Nota:** Quando você cria a expressão, a coluna de pesquisa é sempre o primeiro valor, enquanto a coluna de entrada é o segundo valor.

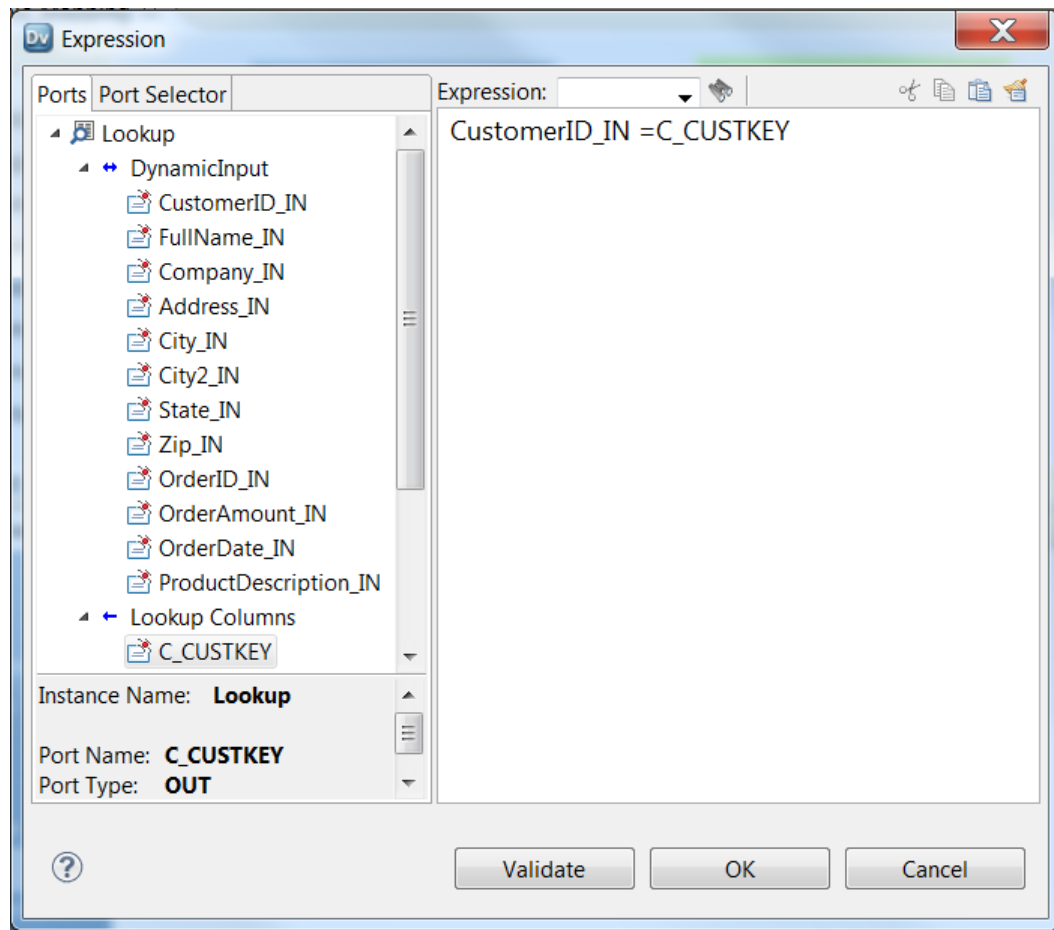
Por exemplo, você cria a seguinte condição de pesquisa em um parâmetro de expressão:

```
CustomerID_IN = C_CUSTKEY
```

CustomerID\_IN é a coluna de pesquisa.

C\_CUSTKEY é a coluna de entrada.

A seguinte imagem mostra a expressão de pesquisa no editor de expressões:



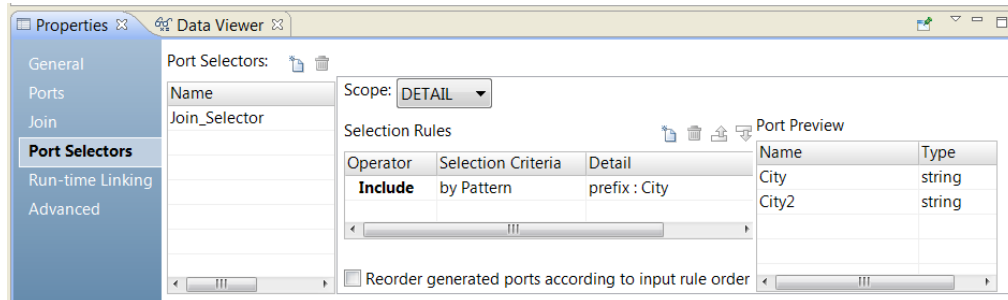
## Criando um Seletor de Portas

Crie um seletor de portas para determinar quais portas devem ser usadas em uma expressão dinâmica, uma condição de pesquisa ou uma condição de associador.

1. Clique na guia **Seletores de Portas**.
2. Na área **Seletores de Portas**, clique em **Novo**.  
A Developer tool cria um seletor de portas com uma regra de seleção padrão que inclui todas as portas.
3. Na área **Seletores de Portas**, mude o nome do seletor de portas para um nome exclusivo.
4. Se estiver trabalhando na transformação de Associador ou na transformação de Pesquisa, escolha o escopo.  
As portas disponíveis mudam de acordo com o grupo de portas que você escolhe.
5. Na área **Regras de Seleção**, selecione um **Operador**.
  - Incluir. Crie uma regra que inclua portas para o seletor de portas. Você deve incluir portas antes de poder excluí-las.
  - Excluir. Crie uma regra que exclua portas específicas do seletor de portas.
6. Escolha os **Critérios de Seleção**.
  - Por Nome. Selecione portas específicas por nome. Você pode selecionar os nomes das portas em uma lista de portas no escopo.

- Por Tipo. Selecione portas por tipo. Você pode selecionar um ou mais tipos de dados.
- Por Padrão. Selecione portas por um padrão de caracteres no nome da porta. É possível pesquisar com caracteres específicos ou criar uma expressão regular.

A seguinte imagem mostra a guia Seletor de Portas:



7. Clique na coluna **Detalhes**.

A caixa de diálogo **Detalhes da Regra de Entrada** é exibida.

8. Selecione os valores pelos quais filtrar as portas.

- Por Nome. Opte por criar uma lista de portas por valor ou de acordo com um parâmetro. Clique em **Escolher** para selecionar as portas na lista.
- Por Tipo. Selecione um ou mais tipos de dados de uma lista. A área **Visualização de Portas** mostra portas dos tipos que você seleciona.
- Por Padrão. Opte por pesquisar o prefixo ou sufixo do nome da porta em busca de um padrão específico de caracteres. Ou, opte por criar uma expressão regular com a qual pesquisar. Configure um parâmetro ou configure o padrão com base no qual pesquisar.

A área **Visualização de Portas** mostra as portas no seletor de portas à medida que você configura as regras.

9. Para reordenar as portas no seletor de portas, selecione **Reordenar portas geradas de acordo com a ordem das regras de entrada**.

# Propriedades de Tempo de Execução

Defina as propriedades de tempo de execução para ativar e configurar o armazenamento em cache de pesquisa. Você deve adicionar a transformação de Pesquisa a um mapeamento que possa configurar as propriedades de pesquisa em tempo de execução.

A seguinte tabela descreve propriedades de tempo de execução para transformações de Pesquisa que executam pesquisas de arquivo simples, de tabela de referência ou relacionais:

Propriedade	Descrição
Armazenamento em Cache Ativado	<p>Indica se o Serviço de Integração armazena valores de pesquisa no cache.</p> <p>Quando você ativa o armazenamento da pesquisa em cache, o Serviço de Integração consulta a origem da pesquisa uma vez, armazena os valores no cache e pesquisa os valores no cache. O armazenamento em cache dos valores de pesquisa pode melhorar o desempenho em tabelas de pesquisa grandes.</p> <p>Quando você desativa o cache, cada vez que uma linha é transferida para a transformação, o Serviço de Integração emite uma instrução de seleção para a origem da pesquisa, para os valores de pesquisa.</p> <p>O Serviço de Integração sempre armazena pesquisas de arquivo simples no cache.</p>
Pesquisar Tamanho do Cache de Dados	<p>Quantidade de memória que o Serviço de Integração de Dados aloca para o cache de dados para a transformação no início da execução do mapeamento. Selecione Automático para que o Serviço de Integração de Dados calcule automaticamente os requisitos de memória no tempo de execução. Insira um valor específico em bytes quando você ajustar o tamanho do cache. O padrão é Automático.</p>
Pesquisar Tamanho do Cache de Índice	<p>Quantidade de memória que o Serviço de Integração de Dados aloca para o cache de índice para a transformação no início da execução do mapeamento. Selecione Automático para que o Serviço de Integração de Dados calcule automaticamente os requisitos de memória no tempo de execução. Insira um valor específico em bytes quando você ajustar o tamanho do cache. O padrão é Automático.</p>
Prefixo do Nome do Arquivo de Cache	<p>Prefixo do arquivo de cache. Você pode especificar o prefixo de nome do arquivo de cache para um cache de pesquisa persistente.</p>
Cache de Pesquisa Pré-Criado	<p>Permite que o Serviço de Integração crie o cache de pesquisa antes que a transformação de Pesquisa receba os dados. O Serviço de Integração pode criar múltiplos arquivos de cache de pesquisa ao mesmo tempo para melhorar o desempenho.</p> <p>Configure uma das seguintes opções:</p> <ul style="list-style-type: none"><li>- Automático. O Serviço de Integração determina o valor.</li><li>- Sempre permitido. Permite que o Serviço de Integração crie o cache de pesquisa antes que a transformação de Pesquisa receba os dados. O Serviço de Integração pode criar múltiplos arquivos de cache de pesquisa ao mesmo tempo para melhorar o desempenho.</li><li>- Nunca permitido. O Serviço de Integração não pode criar o cache de pesquisa antes que a transformação Pesquisa receba a primeira linha.</li></ul>
Nome do Diretório do Cache de Pesquisa	<p>Diretório usado para criar os arquivos de cache de pesquisa quando você configura a transformação de Pesquisa para armazenar a origem da pesquisa no cache.</p> <p>O padrão é o parâmetro do sistema CacheDir. Você pode configurar outro parâmetro do sistema ou parâmetro definido pelo usuário para essa propriedade.</p>
Rearmazenar em Cache da Origem de Pesquisa	<p>Recria o cache de pesquisa para sincronizar o cache persistente com a tabela de pesquisa. Rearmazena a pesquisa do banco de dados no cache quando a transformação de Pesquisa tem um cache de pesquisa persistente e a tabela de pesquisa é alterada ocasionalmente.</p>



## TÓPICOS RELACIONADOS:

- [“Tamanho do Cache” na página 72](#)

# Propriedades avançadas

Configure o cache de pesquisa persistente e a conexão a um banco de dados relacional nas propriedades avançadas. As propriedades exibidas se baseiam no tipo de origem da pesquisa.

A seguinte tabela descreve propriedades avançadas para cada tipo de origem de pesquisa:

Propriedade	Tipo de Origem de Pesquisa	Descrição
Cache de Pesquisa Persistente	Arquivo simples, tabela de referência, pesquisa relacional	Indica se o Serviço de Integração usa um cache de pesquisa persistente, que consiste em pelo menos dois arquivos de cache. Se uma transformação de Pesquisa for configurada para um cache de pesquisa persistente, e arquivos de cache de pesquisa não existirem, o Serviço de Integração criará esses arquivos.
Comparação de Cadeias com Distinção entre Maiúsculas e Minúsculas	Arquivo simples	O Serviço de Integração usa comparações de cadeias com distinção entre maiúsculas e minúsculas ao realizar pesquisas em colunas de cadeia.
Ordenamento Nulo	Arquivo simples	Determina como o Serviço de Integração ordena valores nulos. Você pode optar por classificar valores nulos como altos ou baixos. Por padrão, o Serviço de Integração classifica valores nulos como altos. Isso substitui a configuração do Serviço de Integração para tratar nulos em operadores de comparação como altos, baixos ou nulos. Para pesquisas relacionais, a ordenação de nulos baseia-se no valor padrão do banco de dados.
Nível de Rastreamento	Arquivo simples, objeto de dados lógicos, tabela de referência, pesquisa relacional	Define a quantidade de detalhes que aparecem no log dessa transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.
Atualizar ou Inserir	Tabela de referência, pesquisa relacional	Aplicável somente ao cache de pesquisa dinâmica. O Serviço de Integração atualizará a linha no cache se o tipo de linha que entrar na transformação de Pesquisa for Atualizar, se a linha existir no cache de índice e se os dados do cache forem diferentes da linha existente. O Serviço de Integração inserirá a linha no cache se ela for nova.
Inserir ou Atualizar	Tabela de referência, pesquisa relacional	Aplicável somente ao cache de pesquisa dinâmica. O Serviço de Integração inserirá a linha no cache se o tipo de linha que entrar na transformação de Pesquisa for Inserir e se ela for nova. Se a linha existir no cache de índice, mas o cache de dados for diferente da linha atual, o Serviço de Integração atualizará essa linha no cache de dados.
Valor Antigo de Saída na Atualização	Tabela de referência, pesquisa relacional	O Serviço de Integração gera o valor que existia no cache antes da atualização da linha. Caso contrário, o Serviço de Integração gera o valor atualizado que ele grava no cache.

Propriedade	Tipo de Origem de Pesquisa	Descrição
Atualizar Condição de Cache Dinâmico	Tabela de referência, pesquisa relacional	Aplicável somente ao cache de pesquisa dinâmica. Uma expressão que indica se o cache dinâmico deve ou não ser atualizado. O Serviço de Integração atualiza o cache quando a condição é verdadeira e o dados existem no cache. O padrão é True.
Conexão	Tabela de referência, pesquisa relacional	Conexão com o banco de dados relacional que contém a origem da pesquisa relacional. Você pode usar um parâmetro para a conexão. Para pesquisas de tabela de referência, esse campo é somente leitura.
Entrada Classificada	Arquivo simples	Indica que os dados de entrada são pré-classificados por grupos.
Formato de Data e Hora	Arquivo simples	Defina um formato de data e hora e a largura do campo. Os formatos de milissegundos, microssegundos ou nanossegundos têm a largura de campo de 29. Se você não selecionar um formato de data e hora para uma porta, poderá inserir qualquer formato de data e hora. O padrão é YYYY-MM-DD HH24:MI:SS. O formato de Data e Hora não altera o tamanho da porta. Esse campo é somente leitura.
Separador de Milhares	Arquivo simples	O valor é Nenhum. Esse campo é somente leitura.
Separador Decimal	Arquivo simples	O valor é um ponto. Esse campo é somente leitura.

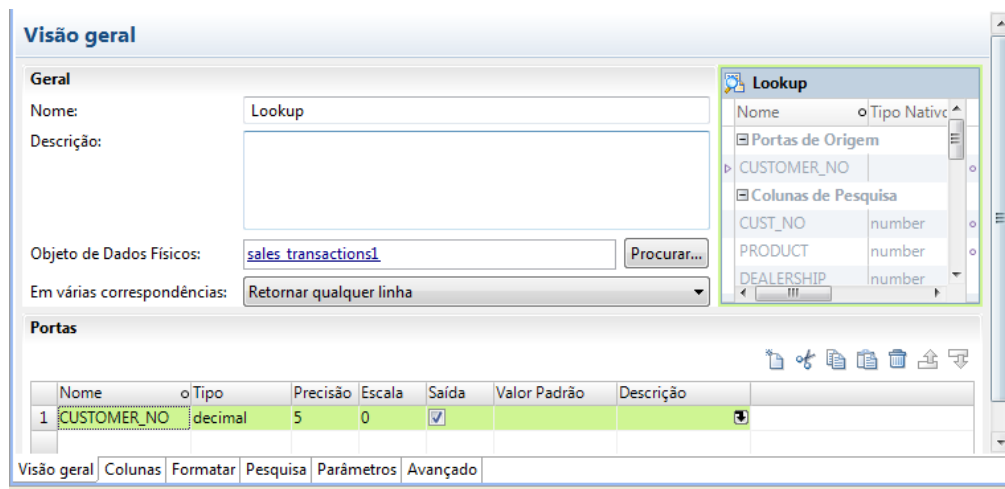
## Criando uma Transformação de Pesquisa Reutilizável

Crie uma transformação de Pesquisa para pesquisar dados em um arquivo simples, um objeto de dados lógicos, uma tabela de referência ou um objeto de dados relacionais.

1. Selecione um projeto ou pasta na exibição **Object Explorer**.
2. Clique em **Arquivo > Novo > Transformação**.
3. Navegue até o assistente de Pesquisa.
4. Selecione **Pesquisa de Objeto de Dados de Arquivo Simples**, **Pesquisa de Objeto de Dados Lógicos**, **Pesquisa de Tabela de Referência** ou **Pesquisa de Objeto de Dados Relacionais**.
5. Clique em **Avançar**.  
A caixa de diálogo **Transformação de Nova Pesquisa** é exibida.
6. Selecione um objeto de dados físicos ou uma tabela de referência na ferramenta Developer.
7. Insira um nome para a transformação.
8. Para **Em várias correspondências**, determine quais linhas são retornadas pela transformação de Pesquisa quando ela encontra várias linhas que correspondem à condição de pesquisa.
9. Clique em **Concluir**.  
A transformação de Pesquisa aparece no editor.

10. Na seção **Portas** da exibição **Visão Geral**, adicione portas de saída à transformação.

A figura a seguir mostra a porta de saída CUSTOMER\_NO em uma transformação de Pesquisa:



11. Na guia **Tempo de Execução** da exibição **Propriedades**, selecione **Armazenamento da Pesquisa em Cache Ativado** para ativar o cache de pesquisa.  
**Nota:** Você deve adicionar a transformação de Pesquisa a um mapeamento para poder configurar propriedades de pesquisa em tempo de execução.
12. Na guia **Pesquisa** da exibição **Propriedades**, adicione uma ou mais condições de pesquisa.
13. Na guia **Avançado** da exibição **Propriedades**, configure as o nível de rastreamento, as propriedades do cache de pesquisa dinâmica e a conexão em tempo real.
14. Salve a transformação.

## Criando uma Transformação de Pesquisa Não Reutilizável

Crie uma transformação de Pesquisa não reutilizável em um mapeamento ou mapplet.

1. Em um mapeamento ou mapplet, arraste uma transformação de Pesquisa da paleta de Transformação para o editor.  
A caixa de diálogo **Novo** é exibida.
2. Selecione **Pesquisa de Objeto de Dados de Arquivo Simples**, **Pesquisa de Objeto de Dados Lógicos**, **Pesquisa de Tabela de Referência** ou **Pesquisa de Objeto de Dados Relacionais**.
3. Clique em **Avançar**.  
A caixa de diálogo **Transformação de Nova Pesquisa** é exibida.
4. Selecione um objeto de dados físicos ou uma tabela de referência na ferramenta Developer.
5. Insira um nome para a transformação.
6. Para **Em várias correspondências**, determine quais linhas são retornadas pela transformação de Pesquisa quando ela encontra várias linhas que correspondem à condição de pesquisa.
7. Clique em **Concluir**.

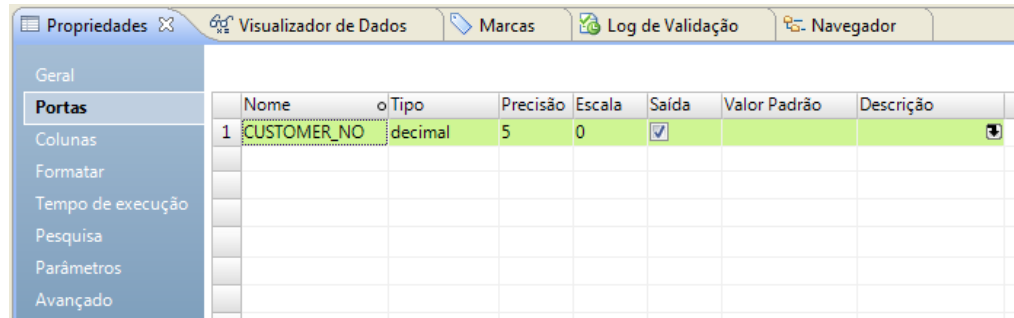
A transformação de Pesquisa aparece no editor.

8. Selecione a transformação de Pesquisa no editor.

A barra de ferramentas é exibida acima da transformação.

9. Na guia **Portas** da exibição **Propriedades**, adicione portas de saída à transformação.

A figura a seguir mostra a porta de saída CUSTOMER\_NO em uma transformação de Pesquisa:



10. Na guia **Tempo de Execução** da exibição **Propriedades**, selecione **Armazenamento da Pesquisa em Cache Ativado** para ativar o cache de pesquisa.

**Nota:** Você deve adicionar a transformação de Pesquisa a um mapeamento para poder configurar propriedades de pesquisa em tempo de execução.

11. Na guia **Pesquisa** da exibição **Propriedades**, adicione uma ou mais condições de pesquisa.
12. Na guia **Avançado** da exibição **Propriedades**, configure as o nível de rastreamento, as propriedades do cache de pesquisa dinâmica e a conexão em tempo real.
13. Salve a transformação.

## Criando uma Transformação de Pesquisa Não Conectada

Crie uma transformação de Pesquisa não conectada quando quiser realizar uma pesquisa a partir de uma expressão. Você pode criar uma transformação de Pesquisa não conectada reutilizável ou não reutilizável em um arquivo simples, em uma tabela de referência ou em um objeto de dados relacionais.

1. Selecione um projeto ou uma pasta na exibição **Object Explorer**.
2. Clique em **Arquivo > Novo > Transformação**.
3. Navegue até o assistente de Pesquisa.
4. Selecione **Pesquisa de Objeto de Dados de Arquivo Simples**, **Pesquisa de Tabela de Referência** ou **Pesquisa de Objeto de Dados Relacionais**.
5. Clique em **Avançar**.

A caixa de diálogo **Nova Pesquisa** é exibida.

6. Selecione um objeto de dados físicos ou uma tabela de referência na ferramenta Developer.
7. Insira um nome para a transformação.
8. Para **Em várias correspondências**, determine qual linha é retornada pela transformação de Pesquisa quando ela encontra várias linhas que correspondem à condição de pesquisa. Não escolha **Retornar Tudo** para uma pesquisa não conectada.

9. Clique em **Concluir**.  
A transformação de Pesquisa aparece no editor.
10. Na seção **Portas** da exibição **Visão Geral**, adicione portas à transformação.  
Crie uma porta de entrada para cada argumento na expressão: LKP. Crie uma porta de entrada para cada condição de pesquisa que você criar. Você pode usar uma porta de entrada em várias condições.
11. Na seção **Portas** da exibição **Visão Geral**, configure uma porta como a porta de retorno.
12. Na exibição **Pesquisa**, adicione uma ou mais condições de pesquisa para comparar os valores de entrada da transformação com os valores no cache ou na origem da pesquisa.  
Quando a condição é verdadeira, a pesquisa retorna um valor na porta de retorno. Se a condição de pesquisa é falsa, a pesquisa retorna NULL.
13. Crie uma expressão :LKP para uma porta em uma transformação que permita expressões, como uma transformação de Agregador, de Expressão ou de Estratégia de Atualização.
14. Quando você criar um mapeamento, adicione a transformação de Pesquisa não conectada ao mapeamento no editor, mas não conecte as portas às outras transformações no mapeamento.

## Exemplo de Pesquisa Não Conectada

Uma loja de varejo na Califórnia adiciona um imposto estadual sobre vendas a cada preço dos itens que vende para clientes dentro do estado. O valor do imposto baseia-se no condado em que o cliente reside. Para recuperar o imposto sobre vendas, você cria uma transformação de Pesquisa que recebe um nome de condado e retorna um valor de imposto sobre vendas referente a esse condado. Se o condado não cobra imposto sobre vendas, a transformação de Pesquisa retorna NULL. Chame a pesquisa de uma transformação de Expressão.

Conclua as seguintes etapas para configurar uma pesquisa não conectada de impostos sobre vendas por condado:

1. Importe um objeto de dados físicos de arquivo simples que contenha os valores dos impostos sobre vendas por condado.
2. Crie a transformação de Pesquisa não conectada.
3. Adicione portas de entrada à transformação de Pesquisa.
4. Defina a porta de retorno.
5. Crie a condição de pesquisa.
6. Chame a pesquisa de uma transformação de Expressão.

### **Etapas 1. Importar a origem da pesquisa de impostos sobre vendas no repositório do Modelo**

O arquivo de impostos sobre vendas deverá estar no repositório do Modelo antes de você criar a transformação de Pesquisa. Para este cenário, o arquivo de impostos sobre vendas contém dois campos: Sales\_County e County\_SalesTax. O condado é uma string que contém um nome de condado. County\_SalesTax é um campo decimal que contém uma alíquota de imposto para o condado. O arquivo de impostos sobre vendas é a origem da pesquisa.

### **Etapas 2. Criar a Transformação de Pesquisa Não Conectada**

Crie uma transformação de Pesquisa de arquivo simples reutilizável com o objeto de dados de arquivo simples de impostos sobre vendas. Para este cenário, o nome da transformação é Sales\_Tax\_Lookup. Selecione **Retornar a Primeira Linha** em várias correspondências.

### Etapa 3. Definir as Portas da Transformação de Pesquisa

Definir as portas da transformação de Pesquisa na guia **Portas** de **Propriedades**.

Tipo de Porta	Nome	Tipo	Comprimento	Escala
Entrada	In_County	String	25	
Saída	SalesTax	Decimal	3	3

### Etapa 4. Configurar a Porta de Retorno da Transformação de Pesquisa

A porta de retorno é o campo no arquivo simples que a pesquisa recupera. Na guia **Colunas**, a coluna **County\_SalesTax** é a porta de retorno.

Quando a pesquisa é verdadeira, o Serviço de Integração localiza o condado na origem de arquivo simples. O Serviço de Integração retorna um valor de imposto sobre vendas na porta de retorno. Se o Serviço de Integração não encontrar o condado, o resultado da pesquisa será falso, e o Serviço de Integração retornará NULL na porta de retorno.

### Etapa 5. Definir a Condição de Pesquisa

Na exibição **Pesquisa**, defina a condição de pesquisa para comparar o valor de entrada com os valores na origem da pesquisa.

Para adicionar a condição de pesquisa, clique na **Coluna de Pesquisa**.

A condição de pesquisa tem a seguinte sintaxe:

```
SALES_COUNTY = IN_COUNTY
```

### Etapa 6. Criar uma Transformação de Expressão

Crie uma transformação de Expressão que recebe registros de vendas de um arquivo simples. A transformação de Expressão recebe um número do cliente, o valor da venda e a o condado onde a venda foi efetuada. Ela retorna o número do cliente, o valor da venda e um imposto sobre vendas.

A transformação de Expressão tem as seguintes portas:

Tipo de Porta	Nome	Tipo	Comprimento	Precisão	Valor padrão
Entrada	County	String	25	10	
Passagem	Customer	String	10		
Passagem	SalesAmt	Decimal	10	2	
Saída	SalesTax	Decimal	10	2	0

A porta **SalesTax** contém uma expressão: **LKP**. A expressão chama a transformação de Pesquisa **Sales\_Tax** e passa o nome do condado como parâmetro. A transformação **Sales\_Tax\_Lookup** retorna a alíquota do imposto sobre vendas à expressão. A transformação de Expressão multiplica a alíquota do imposto pelo valor da venda.

Insira a seguinte expressão para a porta **SalesTax**:

```
(:LKP.Sales_Tax_Lookup(County) * SalesAmt)
```

A porta **SalesTax** contém o resultado da expressão. Se a pesquisa falha, a transformação de Pesquisa retorna NULL e a porta **SalesTax** contém valores nulos.

Você pode adicionar uma expressão para verificar se há valores nulos na porta **SalesTax**. Se **SalesTax** é NULL, você pode configurar a porta **SalesTax** para retornar zero. Adicione o seguinte texto à expressão de pesquisa para verificar a existência de valores nulos e retornar zero:

```
IIF(ISNULL(:LKP.Sales_Tax_Lookup(County) * SalesAmt),0, SalesTax)
```

# Transformação de Pesquisa em um ambiente não nativo

O processamento da transformação de Pesquisa em um ambiente não nativo depende do mecanismo que executa a transformação.

Considere o suporte para os seguintes mecanismos de tempo de execução não nativos:

- Mecanismo Blaze. Suportado com restrições.
- Mecanismo Spark. Com suporte com restrições em mapeamentos de fluxo e em lote.
- Mecanismo Databricks Spark. Suportado com restrições.

## Transformação de Pesquisa no mecanismo Blaze

Algumas regras de processamento para o mecanismo Blaze são diferentes das regras de processamento do Serviço de Integração de Dados.

A validação do mapeamento falha nas seguintes situações:

- O cache é configurado para ser compartilhado, nomeado, persistente, dinâmico ou não armazenado em cache. O cache deve ser um cache estático.

Se você adicionar um objeto de dados que use o Sqoop como uma transformação de Pesquisa em um mapeamento, o Serviço de Integração de Dados não executará o mapeamento por meio do Sqoop. Ele executará o mapeamento por meio do JDBC.

## Transformação de Pesquisa no mecanismo Spark

Algumas regras de processamento para o mecanismo Spark são diferentes das regras de processamento do Serviço de Integração de Dados.

### Validação de Mapeamento

A validação do mapeamento falha nas seguintes situações:

- A diferenciação de maiúsculas e minúsculas está desativada.
- A condição de pesquisa na transformação de Pesquisa contém o tipo de dados binary.
- O cache é configurado para ser compartilhado, nomeado, persistente, dinâmico ou não armazenado em cache. O cache deve ser um cache estático.

O mapeamento falha na seguinte situação:

- A transformação é desconectada e usada com uma transformação de Associador ou Java.

### Várias Correspondências

Quando você escolhe retornar o primeiro, o último ou qualquer valor em várias correspondências, a transformação de Pesquisa retorna qualquer valor.

Se você configurar a transformação para relatar um erro em várias correspondências, o mecanismo Spark desativará as linhas duplicadas e não incluirá as linhas nos registros.

**Nota:** Se uma consulta do HBase não resultar em uma correspondência, ela gerará uma linha com valores nulos para todas as colunas. Você pode adicionar uma transformação de Filtro após a transformação de Pesquisa para filtrar as linhas nulas.

## Transformação de Pesquisa em um mapeamento de fluxo

Os mapeamentos de fluxo têm regras de processamento adicionais que não se aplicam a mapeamentos em lote.

### Validação de Mapeamento

A validação do mapeamento falha nas seguintes situações:

- A pesquisa é um objeto de dados.
- Uma transformação de Agregador está no mesmo pipeline de fluxo que uma transformação de Pesquisa passiva configurada com uma condição de pesquisa de desigualdade.
- Uma transformação de Classificação está no mesmo pipeline de fluxo que uma transformação de Pesquisa passiva configurada com uma condição de pesquisa de desigualdade.
- Um pipeline contém mais de uma transformação de Pesquisa passiva configurada com uma condição de desigualdade.

O mapeamento falha nas seguintes situações:

- A transformação está desconectada.

### Diretrizes gerais

Considere as seguintes diretrizes gerais:

- Usar um tipo de dados flutuante para pesquisar dados pode retornar resultados inesperados.
- Use uma transformação de Pesquisa para pesquisar por dados em um arquivo simples, HDFS, Hive, relacional e HBase.
- Para evitar a união cruzada de DataFrames, configure a transformação de Pesquisa para ignorar os valores nulos correspondentes.

### Pesquisas no HBase

Para usar uma transformação de Pesquisa em tabelas HBase não armazenadas em cache, siga as seguintes etapas:

1. Crie um objeto de dados HBase. Quando você adicionar uma tabela HBase como o recurso para um objeto de dados HBase, inclua a coluna ROW ID.
2. Crie uma operação de leitura de dados HBase e importe-a para o mapeamento de fluxo.
3. Quando você importa a operação de dados para o mapeamento, selecione a opção **Pesquisa**.
4. Na guia Pesquisa, configure as seguintes opções:
  - Coluna Pesquisa. Especifique uma condição de igualdade em ROW ID
  - Operador. Specify =
5. Verifique se o formato de qualquer valor de data nas tabelas do HBase é de um formato de data Java válido. Especifique esse formato na propriedade **Formato de Data/Hora** da guia **Propriedades Avançadas** da operação de leitura de objeto de dados.

**Nota:** Se uma consulta do HBase não resultar em uma correspondência, ela gerará uma linha com valores nulos para todas as colunas. Você pode adicionar uma transformação de Filtro após a transformação de Pesquisa para filtrar as linhas nulas.

A validação do mapeamento falha nas seguintes situações:

- A condição não contém um ROW ID.
- A transformação contém uma condição de desigualdade.
- A transformação contém várias condições.



- Uma coluna de entrada é de um tipo de data.

## Transformação de Pesquisa no mecanismo Databricks Spark

Algumas regras de processamento para o mecanismo Databricks Spark são diferentes das regras de processamento do Serviço de Integração de Dados.

### Validação de Mapeamento

A validação do mapeamento falha nas seguintes situações:

- A diferenciação de maiúsculas e minúsculas está desativada.
- A condição de pesquisa na transformação de Pesquisa contém o tipo de dados binary.
- O cache é configurado para ser compartilhado, nomeado, persistente, dinâmico ou não armazenado em cache. O cache deve ser um cache estático.
- A fonte de pesquisa não é o Microsoft Azure SQL Data Warehouse.

O mapeamento falha na seguinte situação:

- A transformação é desconectada e usada com uma transformação de Associador.

### Várias Correspondências

Quando você escolhe retornar o primeiro, o último ou qualquer valor em várias correspondências, a transformação de Pesquisa retorna qualquer valor.

Se você configurar a transformação para relatar um erro em várias correspondências, o mecanismo Spark desativará as linhas duplicadas e não incluirá as linhas nos registros.

**Nota:** Se uma consulta do HBase não resultar em uma correspondência, ela gerará uma linha com valores nulos para todas as colunas. Você pode adicionar uma transformação de Filtro após a transformação de Pesquisa para filtrar as linhas nulas.

## CAPÍTULO 26

# Caches de Pesquisa

Este capítulo inclui os seguintes tópicos:

- [Visão geral de Caches de Pesquisa, 438](#)
- [Tipos de Cache de Pesquisa, 439](#)
- [Pesquisa sem Cache, 440](#)
- [Cache de Pesquisa Estático, 440](#)
- [Cache de Pesquisa Persistente, 441](#)
- [Cache de Pesquisa Dinâmica, 442](#)
- [Cache de Pesquisa Compartilhado, 442](#)
- [Comparação de Caches, 444](#)
- [Particionamento do Cache para Pesquisas, 444](#)

## Visão geral de Caches de Pesquisa

Você pode configurar uma transformação de Pesquisa para armazenar em cache uma origem de pesquisa de arquivo relacional ou simples. Ative o armazenamento de pesquisa em cache em uma grande tabela ou arquivo de pesquisa para melhorar o desempenho da pesquisa.

O Serviço de Integração cria um cache na memória ao processar a primeira linha dos dados em uma transformação de Pesquisa em cache. O Serviço de Integração cria caches à medida que as linhas de origem entram na transformação de Pesquisa. Ele aloca memória para o cache com base na quantidade configurada na transformação. O Serviço de Integração armazena valores de condição no cache de índice e valores de saída no cache de dados. O Serviço de Integração consulta o cache a cada linha que entra na transformação.

Se os dados não couberem no cache de memória, o Serviço de Integração armazenará os valores excedentes em arquivos de cache. O Serviço de Integração cria os arquivos de cache no diretório de cache. Por padrão, o Serviço de Integração cria os arquivos de cache no diretório especificado no parâmetro do sistema CacheDir. Quando o mapeamento é concluído, o Integration Service libera a memória cache e exclui os arquivos do cache, a menos que você configure a transformação de Pesquisa para usar um cache persistente.

Se você usar uma pesquisa de arquivo simples, o Serviço de Integração armazena a origem da pesquisa em cache. Se você configurar uma pesquisa de arquivo simples na entrada classificada, o Serviço de Integração não poderá armazenar a pesquisa se as colunas de condição não estiverem agrupadas. Se as colunas estiverem agrupadas, mas não classificadas, o Serviço de Integração processa a pesquisa como se você não houvesse configurado a entrada classificada.

Quando você não configura a transformação de Pesquisa para armazenamento em cache, o Serviço de Integração consulta a origem de cada linha de entrada da pesquisa. Não importa se você armazena ou não a origem da pesquisa: o resultado da consulta e do processamento da Pesquisa é o mesmo. No entanto, você pode melhorar o desempenho da pesquisa em uma grande origem de pesquisa se ativar o armazenamento da pesquisa em cache.

## Tipos de Cache de Pesquisa

Você pode configurar tipos diferentes de caches de pesquisa. Por exemplo, você pode configurar um cache compartilhado se quiser compartilhar o cache entre várias transformações de Pesquisa no mesmo mapeamento.

Você pode configurar os seguintes tipos de caches de pesquisa:

### **Cache estático**

Um cache estático não é alterado enquanto o Serviço de Integração processa a pesquisa. O Serviço de Integração recria um cache estático sempre que processa a pesquisa. Por padrão, o Serviço de Integração cria um cache estático quando você ativa o cache para uma transformação de Pesquisa. O Serviço de Integração cria o cache quando processa a primeira solicitação de pesquisa. Ele procura valores no cache para cada linha que entrar na transformação de Pesquisa. Quando a condição de pesquisa é verdadeira, o Serviço de Integração retorna um valor do cache de pesquisa.

Use um cache estático pelos seguintes motivos:

- A origem da pesquisa não muda durante a execução do mapeamento.
- A pesquisa é uma pesquisa desconectada. Você deve usar um cache estático para uma pesquisa desconectada.
- Para melhorar o desempenho. Como o Serviço de Integração não atualiza o cache enquanto processa a transformação de Pesquisa, ele processa uma transformação de Pesquisa com um cache estático mais rápido que uma transformação de Pesquisa com um cache dinâmico.
- Você deseja que o Serviço de Integração retorne o valor padrão para transformações conectadas ou NULL para transformações desconectadas quando a condição de pesquisa é falsa.

### **Cache Persistente**

Um cache persistente não é alterado sempre que o Serviço de Integração processa a pesquisa. O Serviço de Integração salva os arquivos de cache de pesquisa e os reutiliza na próxima vez que processar uma transformação de Pesquisa configurada para usar o cache. Use um cache persistente quando a origem de pesquisa não for alterada.

Se necessário, você pode configurar a transformação de Pesquisa para recriar o cache de pesquisa persistente.

### **Cache Dinâmico**

Um cache de pesquisa dinâmica é alterado enquanto o Serviço de Integração processa a pesquisa. O Serviço de Integração cria um cache de pesquisa dinâmica quando processa a primeira solicitação de pesquisa. Quando processa cada linha, o Serviço de Integração insere ou atualiza dinamicamente os dados no cache de pesquisa e transmite os dados para o destino. O cache dinâmico é sincronizado com o destino.

Use um cache dinâmico quando você deseja atualizar um destino com base nos registros novos e alterados. Você também pode usar um cache dinâmico quando o mapeamento exigir uma pesquisa em dados de destino, mas a conexão para o destino é lenta.

### Cache Compartilhado

Um cache compartilhado pode ser usado por várias transformações de Pesquisa no mesmo mapeamento. Use um cache compartilhado para melhorar o desempenho do mapeamento. Em vez de gerar um cache de pesquisa separado para cada transformação de Pesquisa, o Serviço de Integração gera um cache.

## Pesquisa sem Cache

Uma pesquisa sem cache acontece quando o Serviço de Integração não armazena a origem da pesquisa em cache. Por padrão, o Serviço de Integração não usa um cache de pesquisa para uma transformação de Pesquisa.

O Serviço de Integração processa uma pesquisa sem cache do mesmo modo que processa outra com cache, mas consulta a origem da pesquisa em vez de compilar e consultar o cache.

Quando a condição de pesquisa é verdadeira, o Serviço de Integração retorna os valores da origem de pesquisa. Quando o Serviço de Integração processa uma transformação de Pesquisa conectada, ele retorna os valores representados pelas portas de pesquisa/saída. Quando processa uma transformação de Pesquisa não conectada, o Serviço de Integração retorna o valor representado pela porta de retorno.

Se a condição não for verdadeira, o Serviço de Integração retorna valores NULL ou padrão. Quando o Serviço de Integração processa uma transformação de Pesquisa conectada, ele retorna o valor padrão da porta de saída quando a condição não é atendida. Quando processa uma transformação de Pesquisa não conectada, o Serviço de Integração retorna NULL quando a condição não é atendida.

## Cache de Pesquisa Estático

Um cache de pesquisa estático é um cache que o Serviço de Integração não atualiza quando processa a transformação de Pesquisa. Por padrão, o Serviço de Integração cria um cache de pesquisa estático quando você configura uma transformação de Pesquisa para cache.

O Serviço de Integração cria o cache quando processa a primeira solicitação de pesquisa. Ele consulta o cache com base na condição de pesquisa para cada linha que passa pela transformação.

Quando a condição de pesquisa é verdadeira, o Serviço de Integração retorna os valores do cache de pesquisa estático. Quando o Serviço de Integração processa uma transformação de Pesquisa conectada, ele retorna os valores representados pelas portas de pesquisa/saída. Quando o Serviço de Integração processa uma transformação de Pesquisa não conectada, ele retorna o valor representado pela porta de retorno.

Se a condição não for verdadeira, o Serviço de Integração retorna valores NULL ou padrão. Quando o Serviço de Integração processa uma transformação de Pesquisa conectada, ele retorna o valor padrão da porta de saída quando a condição não é atendida. Quando o Serviço de Integração processa uma transformação de Pesquisa não conectada, ele retorna NULL quando a condição não é atendida.

# Cache de Pesquisa Persistente

Um cache de pesquisa persistente é um cache que o Serviço de Integração reutiliza para várias execuções do mesmo mapeamento. Use um cache de pesquisa persistente quando a origem de pesquisa não mudar entre execuções de mapeamento.

Por padrão, o Serviço de Integração usa cache não persistente quando você ativa o cache de pesquisa em uma transformação de Pesquisa. O Serviço de Integração exclui os arquivos de cache quando o mapeamento é concluído. Na próxima vez em que você executar o mapeamento, o Serviço de Integração compilará o cache de memória da origem da pesquisa.

Se você configurar a transformação de Pesquisa para usar um cache de pesquisa persistente, o Serviço de Integração salvará e reutilizará arquivos de cache para várias execuções de mapeamento. Um cache persistente elimina o tempo necessário para ler a tabela de pesquisa e reconstruir o cache de pesquisa.

Na primeira vez em que o Serviço de Integração executa um mapeamento com um cache de pesquisa persistente, ele salva os arquivos de cache no disco. Na próxima vez que o Integration Service executar o mapeamento, ele compilará o cache de memória dos arquivos de cache.

Você poderá configurar o Serviço de Integração para recriar o cache de pesquisa persistente se a origem de pesquisa original for alterada. Quando você recompilar o cache, o Serviço de Integração criará novos arquivos de cache e gravará uma mensagem no log do Serviço de Integração.

## Recriando um Cache de Pesquisa Persistente

Você pode configurar o Serviço de Integração para recriar o cache de pesquisa persistente. Em alguns casos, o Serviço de Integração recompilará o cache de pesquisa persistente mesmo que você não o configure para isso.

Considere as seguintes regras e diretrizes ao recompilar o cache de pesquisa persistente:

- Recrie um cache de pesquisa persistente se a origem de pesquisa tiver sido alterada desde a última vez que o Serviço de Integração criou o cache.
- Você pode recompilar o cache quando o mapeamento contiver uma ou mais transformações de Pesquisa que compartilham um cache.
- Se a tabela de pesquisa não mudar entre as sessões de mapeamento, configure a transformação de Pesquisa para usar um cache de pesquisa persistente. O Serviço de Integração salva e reutiliza arquivos de cache, eliminando o tempo necessário para ler a tabela de pesquisa.
- Se você configurar transformações de Pesquisa subsequentes para reconstruir o cache de pesquisa, o Serviço de Integração compartilhará o cache ao invés de recriá-lo quando processar a transformação de Pesquisa subsequente.
- Se um mapeamento contiver duas pesquisas persistentes e você configurar a segunda transformação de Pesquisa para recriar o cache, o Serviço de Integração recriará o cache de pesquisa persistente de ambas.

O Serviço de Integração recria o cache de pesquisa persistente nas seguintes situações:

- O Serviço de Integração não pode localizar os arquivos de cache.
- O Serviço de Integração não pode reutilizar o cache. Neste caso, ele reconstrói o cache de pesquisa ou o mapeamento falha.
- Você ativa ou desativa a alta precisão para o Serviço de Integração.
- Você edita a transformação de Pesquisa ou o mapeamento.

**Nota:** O Serviço de Integração não reconstruirá o cache se você editar a descrição da transformação.

- Você altera o número de partições.
- Você altera a conexão de banco de dados ou a localização do arquivo usado para acessar a origem da pesquisa.
- Você altera a ordem de classificação no modo Unicode.
- Você altera a página de código do Serviço de Integração.

## Cache de Pesquisa Dinâmica

Um cache dinâmico é um cache que o Serviço de Integração atualiza quando processa cada linha. Use um cache de pesquisa dinâmica para manter o cache sincronizado com o destino.

Você pode usar um cache dinâmico com uma pesquisa relacional e uma pesquisa de arquivo simples. O Serviço de Integração cria o cache quando processa a primeira solicitação de pesquisa. Ele consulta o cache com base na condição de pesquisa para cada linha que passa pela transformação de Pesquisa. O Serviço de Integração atualiza o cache de pesquisa quando processa cada linha.

Com base nos resultados da consulta de pesquisa, no tipo de linha e nas propriedades da transformação de Pesquisa, o Serviço de Integração insere ou atualiza linhas no cache ou não faz nenhuma alteração nele.

## Cache de Pesquisa Compartilhado

Um cache de pesquisa compartilhado é um cache de pesquisa estático compartilhado por várias transformações de Pesquisa em um mapeamento. Use um cache de pesquisa compartilhado para reduzir a quantidade de tempo necessário para criar o cache.

Por padrão, o Serviço de Integração compartilha o cache para transformações de Pesquisa em um mapeamento que tem estruturas de memória cache compatíveis. Por exemplo, se o mapeamento contiver duas instâncias da mesma transformação de Pesquisa reutilizável em um mapeamento e você utilizar as mesmas portas de saída para ambas, as transformações de Pesquisa compartilharão o cache de pesquisa por padrão.

O Serviço de Integração constrói o cache quando processa a primeira transformação de Pesquisa. Ele usa o mesmo cache para processar transformações de Pesquisa subsequentes que compartilham o cache. Quando o Serviço de Integração compartilha um cache de pesquisa, ele grava uma mensagem no log do Serviço de Integração.

O Serviço de Integração aloca memória de cache de dados e memória cache de índice para a primeira transformação de Pesquisa. Ele não aloca mais memória para transformações de Pesquisa subsequentes que compartilham o cache de pesquisa.

Se a transformação ou a estrutura do cache não permitir compartilhamento, o Serviço de Integração criará um novo cache.

## Regras e Diretrizes para Compartilhar um Cache de Pesquisa

Considere as seguintes regras e diretrizes quando compartilhar um cache de pesquisa:

- Você pode compartilhar um ou mais caches estáticos com uma pesquisa dinâmica. Se uma pesquisa dinâmica compartilha o cache com uma pesquisa estática no mesmo mapeamento, a pesquisa estática reutiliza o cache criado pela pesquisa dinâmica.
- Você não pode compartilhar um cache entre pesquisas dinâmicas.
- Se você configurar várias transformações de Pesquisa para reconstruir o cache de pesquisa persistente, o Serviço de Integração cria o cache para a primeira transformação de Pesquisa e compartilha o cache de pesquisa persistente para as transformações de Pesquisa subsequentes.
- Se você não configurar a primeira transformação de Pesquisa para reconstruir um cache de pesquisa persistente, mas configurar uma transformação de Pesquisa subsequente para reconstruir o cache, as transformações não poderão compartilhar o cache. O Serviço de Integração cria o cache ao processar cada transformação de Pesquisa.
- As portas de pesquisa/saída da transformação de Pesquisa subsequente devem corresponder ou ser um subconjunto das portas da transformação de Pesquisa que o Serviço de Integração utiliza para criar o cache. A ordem das portas não precisa corresponder.
- As transformações de Pesquisa que compartilham o cache devem ter as seguintes características:
  - As transformações de Pesquisa devem usar as mesmas portas na condição de pesquisa.
  - As transformações de Pesquisa devem usar a mesma substituição SQL, se esta for usada.
  - O cache de pesquisa deve ser ativado em todas as transformações de Pesquisa.
  - As transformações de Pesquisa devem usar o mesmo tipo de origem da pesquisa.
  - Todas as transformações de Pesquisa relacionais devem usar a mesma conexão de banco de dados.
  - As transformações de Pesquisa devem usar o mesmo nome de tabela de pesquisa.
  - A estrutura do cache de todas as transformações de Pesquisa deve ser compatível.

## Comparação de Caches

O Serviço de Integração funciona de forma diferente com base no tipo de cache de pesquisa que você configura.

A seguinte tabela compara as transformações Pesquisa a uma pesquisa sem cache, um cache estático e um cache dinâmico:

Sem Cache	Cache Estático	Cache Dinâmico
O Serviço de Integração não insere ou atualiza o cache.	O Serviço de Integração não insere ou atualiza o cache.	O Serviço de Integração pode inserir ou atualizar linhas no cache, à medida que ele passa linhas para o destino.
Você pode usar uma pesquisa relacional.	Você pode usar uma pesquisa relacional ou de arquivo simples.	Você pode usar uma pesquisa relacional ou de arquivo simples.
Quando a condição é verdadeira, o Serviço de Integração retorna um valor a partir da tabela ou cache de pesquisa. Quando a condição não é verdadeira, o Serviço de Integração retorna o valor padrão para as transformações conectadas e NULL para as desconectadas.	Quando a condição é verdadeira, o Serviço de Integração retorna um valor a partir da tabela ou cache de pesquisa. Quando a condição não é verdadeira, o Serviço de Integração retorna o valor padrão para as transformações conectadas e NULL para as desconectadas.	Quando a condição é verdadeira, o Serviço de Integração atualiza as linhas no cache ou deixa o cache inalterado, com base no tipo de linha. Isto indica que a linha está no cache e na tabela de destino. Você pode transferir linhas atualizadas para um destino. Quando a condição não é verdadeira, o Serviço de Integração insere linhas no cache ou deixa o cache inalterado, com base no tipo de linha. Isto indica que a linha não está no cache ou destino. Você pode transferir linhas inseridas para uma tabela de destino.

## Particionamento do Cache para Pesquisas

O particionamento do cache cria um cache separado para cada partição que processa uma transformação de Agregador, Associador, Classificação ou Pesquisa. O particionamento do cache aumenta o desempenho do mapeamento porque cada partição consulta um cache separado em paralelo.

Quando o Serviço de Integração cria partições para um mapeamento, pode usar o particionamento do cache para transformações de Pesquisa particionadas.

O Serviço de Integração usará o particionamento do cache para as transformações de Pesquisa conectadas quando as seguintes condições forem verdadeiras:

- A condição de pesquisa contém apenas operadores de igualdade.
- Quando a transformação de Pesquisa conectada pesquisa dados em uma tabela relacional e o banco de dados está configurado para comparação que faz distinção entre maiúsculas e minúsculas.

O Serviço de Integração não usa particionamento do cache para transformações de Pesquisa desconectadas.

Quando o Serviço de Integração não usa particionamento do cache para uma transformação de Pesquisa, todas as partições da transformação de Pesquisa compartilham o mesmo cache. Cada partição consulta o mesmo cache em série.



## CAPÍTULO 27

# Cache de Pesquisa Dinâmica

Este capítulo inclui os seguintes tópicos:

- [Visão geral do Cache de Pesquisa Dinâmica, 445](#)
- [Usa para um Cache de Pesquisa Dinâmica, 446](#)
- [Propriedades do Cache de Pesquisa Dinâmica, 446](#)
- [Cache de Pesquisa Dinâmica e Valores de Saída, 448](#)
- [Valores de Transformação Pesquisa, 448](#)
- [Substituição SQL e Cache de Pesquisa Dinâmica, 451](#)
- [Configuração de Mapeamento para um Cache de Pesquisa Dinâmica, 452](#)
- [Atualizações de Cache de Pesquisa Dinâmica Condicional, 454](#)
- [Atualização de cache dinâmico com resultados de expressões, 455](#)
- [Exemplo de Cache de Pesquisa Dinâmica, 457](#)
- [Regras e Diretrizes para Cache de Pesquisa Dinâmica, 458](#)

## Visão geral do Cache de Pesquisa Dinâmica

Use um cache de pesquisa dinâmica para manter o cache sincronizado com o destino. Você pode usar um cache dinâmico com uma pesquisa relacional ou de arquivo simples.

O Serviço de Integração cria o cache de pesquisa dinâmica quando processa a primeira solicitação de pesquisa. Ele consulta o cache com base na condição de pesquisa para cada linha que passa pela transformação. O Serviço de Integração atualiza o cache de pesquisa quando processa cada linha.

Com base nos resultados da consulta de pesquisa, o tipo de linha e as propriedades da transformação de Pesquisa, o Serviço de Integração realiza uma das ações a seguir no cache de pesquisa dinâmica ao ler uma linha da origem:

### **Insere a linha no cache**

O Serviço de Integração insere a linha quando ela não está no cache e você configurou a transformação de Pesquisa para inserir linhas no cache. Você pode configurar a transformação para inserir linhas no cache com base nas portas de entrada ou nas IDs de sequência geradas. O Serviço de Integração sinaliza a linha como inserção.

### **Atualiza a linha no cache**

O Serviço de Integração atualiza a linha quando ela existe no cache e você configurou a transformação de Pesquisa para atualizar linhas no cache. O Serviço de Integração atualiza a linha no cache com base nas portas de entrada. O Serviço de Integração sinaliza a linha como uma linha de atualização.

### **Não faz alterações no cache**

O Serviço de Integração não faz alterações quando a linha existe no cache e você configurou a transformação de Pesquisa para inserir somente linhas novas. Ou, a linha não está no cache e você especificou para somente atualizar as linhas existentes. Ou, a linha está no cache, mas com base na condição da pesquisa, nada é alterado. O Serviço de Integração sinaliza a linha como inalterada.

Com base no valor de NewLookupRow, você também pode configurar uma transformação de Roteador ou de Filtro com a transformação de Pesquisa dinâmica para encaminhar as linhas de inserção ou atualização à tabela de destino. Você pode encaminhar linhas inalteradas para outra tabela de destino ou arquivo simples, ou pode descartá-las.

## Usa para um Cache de Pesquisa Dinâmica

Você pode configurar uma transformação de Pesquisa com um cache de pesquisa dinâmica para atualizar o cache com base nas alterações da origem da pesquisa.

Você pode usar um cache de pesquisa dinâmica pelos seguintes motivos:

### **Atualizar uma tabela mestre de clientes com informações novas e atualizadas de clientes.**

Por exemplo, você pode usar uma transformação de Pesquisa para realizar uma pesquisa na tabela de clientes para determinar se um cliente existe no destino. O cache representa a tabela do cliente. A transformação de Pesquisa insere e atualiza linhas no cache, à medida que transfere as linhas para o destino.

### **Use um arquivo simples exportado como origem de pesquisa em vez de uma tabela relacional.**

Se a conexão com o banco de dados estiver lenta, você pode exportar o conteúdo da tabela relacional para um arquivo simples e usar o arquivo como origem de pesquisa. Por exemplo, você pode precisar usar esse método se uma conexão ODBC para um banco de dados estiver lenta. Você pode configurar a tabela de banco de dados como destino relacional no mapeamento e transmitir as alterações do cache de pesquisa novamente para a tabela de banco de dados.

## Propriedades do Cache de Pesquisa Dinâmica

Configure as propriedades da pesquisa dinâmica para habilitar um cache de pesquisa dinâmica e configurar como o cache é atualizado. Por exemplo, você pode configurar os valores que são inseridos e atualizados no cache dinâmico.

Configure as seguintes propriedades quando você ativar um cache de pesquisa dinâmica:

### **Em Várias Correspondências**

Definir como Relatar Erro.

### **Cache de Pesquisa Dinâmica**

Habilita um cache de pesquisa dinâmica.

Essa opção estará disponível depois que você ativar o cache de pesquisa.

### **Atualizar ou Inserir**

Aplica-se às linhas inseridas na transformação de Pesquisa com o tipo de linha atualizar. Quando ativado, o Serviço de Integração atualiza as linhas existentes no cache e insere uma linha se ela for nova. Quando desativado, o Serviço de Integração não insere linhas novas.

Essa opção estará disponível depois que você ativar o cache dinâmico.

### **Inserir ou Atualizar**

Aplica-se às linhas inseridas na transformação de Pesquisa com o tipo de linha inserir. Quando ativado, o Serviço de Integração insere linhas no cache e atualiza linhas existentes. Quando desativado, o Serviço de Integração não atualiza linhas existentes.

Essa opção estará disponível depois que você ativar o cache dinâmico.

### **Valor Antigo de Saída na Atualização**

A transformação de Pesquisa pode gerar a saída de valores novos ou existentes do cache. Quando ativado, o Serviço de Integração gera a saída dos valores existentes das portas de pesquisa/saída antes de atualizar o valor no cache. Quando o Serviço de Integração atualiza uma linha no cache, ele gera a saída do valor no cache de pesquisa antes de atualizar a linha, com base nos dados de entrada. Quando o Serviço de Integração insere uma linha no cache, ele gera a saída de valores nulos.

Desative a propriedade para que o Serviço de Integração transmita os mesmos valores das portas de pesquisa/saída e entrada/saída. Essa propriedade é ativada por padrão.

Essa opção estará disponível depois que você ativar o cache dinâmico.

### **Atualizar Condição de Cache Dinâmico**

Quando ativado, o Serviço de Integração usa a expressão de condição para determinar se é necessário atualizar o cache dinâmico. O Serviço de Integração atualiza o cache quando a condição é verdadeira e os dados existem no cache.

Crie a expressão usando as portas de pesquisa ou de entrada. A expressão pode conter valores de entrada ou valores no cache de pesquisa. O padrão é verdadeiro.

Essa opção estará disponível depois que você ativar o cache dinâmico.

### **NewLookupRow**

A ferramenta Developer adiciona essa porta a uma transformação de Pesquisa configurada com um cache dinâmico.

A propriedade NewLookupRow pode conter um dos seguintes valores:

- 0 = Nenhuma atualização para o cache.
- 1 = Inserir linha no cache.
- 2 = Atualizar linha no cache.

Para manter o cache de pesquisa e a tabela de destino sincronizados, transfira as linhas para o destino quando o valor NewLookupRow for igual a 1 ou 2.

### **Porta Associada**

O Serviço de Integração usa o valor da porta associada ao atualizar os dados do cache. O Serviço de Integração associa as portas de entrada e as portas de origem de pesquisa especificadas na condição de pesquisa. Você deve configurar as portas associadas para as portas de origem de pesquisa restantes na pesquisa dinâmica. Se você não configurar uma porta associada para todas as portas de origem de pesquisa em uma pesquisa dinâmica, a validação de mapeamento falhará.

Você pode associar uma porta de origem de pesquisa com os seguintes objetos:

Objeto	Descrição
Porta de entrada	Atualiza o cache com base no valor de uma porta de entrada.
Expressão associada	Selecione para inserir uma expressão. O Serviço de Integração atualiza o cache com base no resultado da expressão.
ID da Sequência	Gera uma chave primária para linhas inseridas no cache de pesquisa. Você pode associar um ID da sequência somente com as colunas bigint e int.

#### Ignorar Entradas Nulas para Atualizações

A ferramenta Developer ativa esta propriedade para as portas de pesquisa/saída quando você configura a transformação de Pesquisa para usar um cache dinâmico. Selecione esta propriedade quando você não deseja que o Serviço de Integração atualize a coluna no cache com um valor de entrada nulo.

#### Ignorar na Comparação

O Developer ativa esta propriedade para as portas de pesquisa/saída que não são usadas na condição de pesquisa quando você configura a transformação de Pesquisa para usar um cache dinâmico. O Serviço de Integração compara os valores em todas as portas de pesquisa com os valores nas portas associadas, por padrão. Selecione esta propriedade se desejar que o Serviço de Integração ignore a porta quando for comparar os valores antes de atualizar uma linha. Use esta propriedade para melhorar o desempenho da comparação.

## Cache de Pesquisa Dinâmica e Valores de Saída

Se você habilitar um cache de pesquisa dinâmica, os valores da porta de saída variarão dependendo de como você o configura. O valor de saída da porta de pesquisa/saída depende do fato de haver uma escolha de efetuar a saída de um valor antigo ou novo quando o Serviço de Integração atualiza uma linha.

Você pode configurar a propriedade **Valor Antigo de Saída na Atualização** para especificar um dos seguintes tipos de valores de saída para uma porta de pesquisa/saída:

- Efetuar a saída de valores antigos em atualizações. O Serviço de Integração efetua a saída do valor que existia no cache antes de ter atualizado a linha.
- Saída de novos valores na atualização. O Serviço de Integração efetua a saída do valor de saída que registra no cache. O valor da porta de pesquisa/saída combina com o valor da porta de saída.

## Valores de Transformação Pesquisa

A transformação de Pesquisa contém valores para as portas de entrada, pesquisa e saída. Se você habilitar um cache de pesquisa dinâmica, os valores da porta de saída variarão dependendo de como você o configura.

A transformação de Pesquisa contém os seguintes tipos de valores:

#### Valor de entrada

Valor que o Serviço de Integração transmite para a transformação de Pesquisa.

### Valor de pesquisa

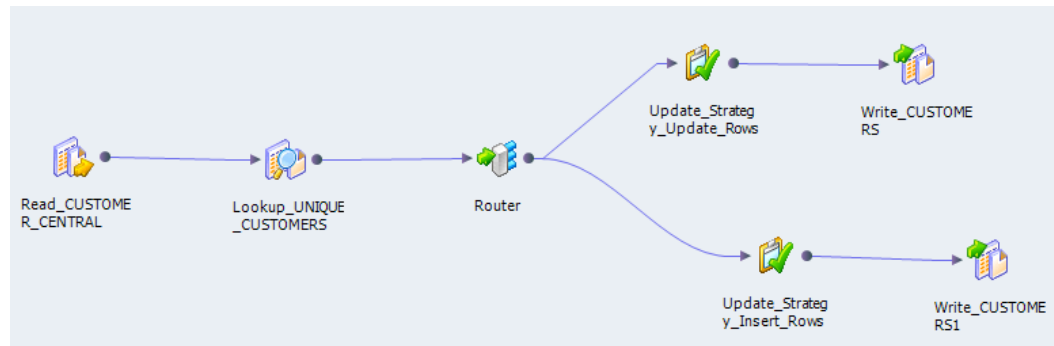
Valor que o Serviço de Integração insere no cache.

### Valor de saída

Valor que o Serviço de Integração transmite da porta de saída da transformação de Pesquisa. O valor de saída da porta de pesquisa/saída depende do fato de haver uma escolha de efetuar a saída de um valor antigo ou novo quando o Serviço de Integração atualiza uma linha.

## Exemplo de Valores de Transformação de Pesquisa

Por exemplo, você cria um mapeamento com os seguintes objetos:



Para a transformação de Pesquisa, você ativa o cache de pesquisa dinâmica e define a seguinte condição de pesquisa:

`IN_CUST_ID = CUST_ID`

Por padrão, o tipo de todas as linhas inseridas na transformação de Pesquisa é inserir. Para efetuar inserções e atualizações no cache e na tabela de destino, selecione a propriedade **Inserir ou Atualizar** na transformação de Pesquisa.

### Valores de Cache Iniciais

Quando você executa o mapeamento, o Serviço de Integração cria o cache de pesquisa da tabela de destino.

A seguinte tabela mostra os valores iniciais do cache de pesquisa:

PK_PRIMARYKEY	CUST_ID	CUST_NAME	ADDRESS
100001	80001	Marion James	100 Main St.
100002	80002	Laura Jones	510 Broadway Ave.
100003	80003	Shelley Lau	220 Burnside Ave.

### Valores de Entrada

A origem contém linhas que existem e outras que não existem na tabela de destino. O Serviço de Integração transmite as linhas de origem para a transformação de Pesquisa.

A seguinte tabela mostra as linhas de origem:

SQ_CUST_ID	SQ_CUST_NAME	SQ_ADDRESS
80001	Marion Atkins	100 Main St.
80002	Laura Gomez	510 Broadway Ave.

<b>SQ_CUST_ID</b>	<b>SQ_CUST_NAME</b>	<b>SQ_ADDRESS</b>
99001	Jon Freeman	555 6th Ave.

## Valores de Pesquisa

O Serviço de Integração procura valores no cache com base na condição de pesquisa. Ele atualiza linhas no cache para as IDs de clientes atuais 80001 e 80002. Ele insere uma linha no cache para a ID de cliente 99001. O Serviço de Integração gera uma nova chave (PK\_PRIMARYKEY) para a nova linha.

A tabela a seguir mostra as linhas e os valores retornados da pesquisa:

<b>PK_PRIMARYKEY</b>	<b>CUST_ID</b>	<b>CUST_NAME</b>	<b>ADDRESS</b>
100001	80001	Marion Atkins	100 Main St.
100002	80002	Laura Gomez	510 Broadway Ave.
100004	99001	Jon Freeman	555 6th Ave.

## Valores de Saída

O Serviço de Integração sinaliza as linhas na transformação de Pesquisa com base nas inserções e atualizações que executa no cache dinâmico. O Serviço de Integração eventualmente transmite as linhas para uma transformação de Roteador que cria uma ramificação para inserir linhas e outra para atualizá-las. Cada ramificação contém uma transformação de Estratégia de Atualização. As transformações de Estratégia de Atualização sinalizam as linhas para inserção ou atualização com base no valor da porta NewLookupRow.

Os valores de saída das portas de pesquisa/saída e entrada/saída dependem do fato de você escolher valores novos ou antigos de saída quando o Serviço de Integração atualizar uma linha. No entanto, os valores de saída da porta NewLookupRow e de qualquer porta de pesquisa/saída que usa a ID da sequência são os mesmos para as linhas novas e atualizadas.

Se você escolher gerar novos valores, as portas de pesquisa/saída gerarão os seguintes valores:

<b>NewLookupRow</b>	<b>PK_PRIMARYKEY</b>	<b>CUST_ID</b>	<b>CUST_NAME</b>	<b>ADDRESS</b>
2	100001	80001	Marion Atkins	100 Main St.
2	100002	80002	Laura Gomez	510 Broadway Ave.
1	100004	99001	Jon Freeman	555 6th Ave.

Se você escolher gerar valores antigos, as portas de pesquisa/saída gerarão os seguintes valores:

<b>NewLookupRow</b>	<b>PK_PRIMARYKEY</b>	<b>CUST_ID</b>	<b>CUST_NAME</b>	<b>ADDRESS</b>
2	100001	80001	Marion James	100 Main St.
2	100002	80002	Laura Jones	510 Broadway Ave.
1	100004	99001	Jon Freeman	555 6th Ave.

Quando o Serviço de Integração atualiza as linhas no cache de pesquisa, ele usa valores da chave primária (PK\_PRIMARYKEY) para as linhas no cache e a tabela de destino.

O Serviço de Integração usa a ID de sequência para gerar uma chave primária para o cliente que não encontra no cache. O Serviço de Integração insere o valor da chave primária no cache de pesquisa e retorna o valor para a porta de pesquisa/saída.

O Serviço de Integração gera os valores das portas de entrada/saída que correspondem aos valores de entrada.

**Nota:** Se o valor de entrada for NULL e você selecionar a propriedade Ignorar Nulo para a porta de entrada associada, o valor de entrada não é igual ao valor de pesquisa ou ao valor gerado como saída da porta de entrada/saída. Quando você seleciona a propriedade Ignorar Nulo, o cache de pesquisa e a tabela de destino podem perder a sincronia se você transferir os valores nulos para o destino. Você deve certificar-se de não transferir valores nulos para o destino.

## Substituição SQL e Cache de Pesquisa Dinâmica

Você pode adicionar uma cláusula WHERE à consulta de pesquisa para filtrar os registros usados para criar o cache e executar uma pesquisa na tabela de banco de dados para uma pesquisa sem cache. No entanto, o Serviço de Integração não usa a cláusula WHERE ao inserir linhas em um cache dinâmico.

Ao adicionar uma cláusula WHERE a uma transformação de Pesquisa usando cache dinâmico, conecte uma transformação Filtro antes da transformação de Pesquisa para filtrar linhas que você não queira inserir na tabela de destino ou do cache. Se você não incluir a transformação Filtro, poderá obter resultados inconsistentes entre o cache e a tabela de destino.

Por exemplo, você configura uma transformação de Pesquisa para realizar uma pesquisa dinâmica na tabela de funcionários EMP, correspondendo linhas por EMP\_ID. Você define a seguinte substituição SQL de pesquisa:

```
SELECT EMP_ID, EMP_STATUS FROM EMP ORDER BY EMP_ID, EMP_STATUS WHERE EMP_STATUS = 4
```

Quando você executa pela primeira vez a , o Serviço de Integração cria o cache de pesquisa da tabela de destino com base na substituição SQL de pesquisa. Todas as linhas no cache corresponderão à condição na cláusula WHERE, EMP\_STATUS = 4.

Por exemplo, o Serviço de Integração lê uma linha de origem que atende à condição de pesquisa especificada, mas o valor de EMP\_STATUS é 2. Embora o destino possa ter a linha onde EMP\_STATUS é 2, o Serviço de Integração não encontra a linha no cache devido à substituição SQL. O Serviço de Integração insere a linha no cache e passa a linha para a tabela de destino. Quando o Serviço de Integração inserir essa linha na tabela de destino, você deverá ter resultados consistentes se a linha já existir. Além disso, nem todas as linhas do cache correspondem à condição da cláusula WHERE na substituição SQL.

Para verificar se você inseriu no cache apenas linhas que correspondam à cláusula WHERE, adicione uma transformação Filtro antes da transformação de Pesquisa e defina a condição de filtro como a condição na cláusula WHERE na Substituição SQL de Pesquisa.

Para o exemplo acima, digite a seguinte condição de filtro na transformação Filtro e a cláusula WHERE na substituição SQL:

```
EMP_STATUS = 4
```

# Configuração de Mapeamento para um Cache de Pesquisa Dinâmica

Se você usar uma pesquisa com um cache dinâmico, deve configurar o mapeamento para atualizar o cache de pesquisa dinâmica e gravar as linhas alteradas no destino.

Conclua as seguintes etapas para configurar um mapeamento com um cache de pesquisa dinâmica:

## **Sinalize as linhas de entrada da transformação de Pesquisa para inserir ou atualizar.**

Por padrão, o tipo de linha de todas as linhas de entrada é INSERT. Adicione uma transformação Estratégia de Atualização antes da transformação de Pesquisa para especificar diferentes tipos de linha para as linhas de entrada.

## **Especifique como o Serviço de Integração trata as linhas de entrada para o cache dinâmico.**

Selecione a opção **Inserir ou Atualizar** ou **Atualizar ou Inserir** para processar as linhas sinalizadas para inserção ou atualização.

## **Crie pipelines de mapeamento separados para as linhas a serem inseridas e atualizadas no destino.**

Adicione uma transformação de Filtro ou de Roteador depois que a transformação de Pesquisa encaminhar as linhas de inserção e atualização para as ramificações de mapeamento separadas. Use o valor de NewLookupRow para determinar a ramificação adequada para cada linha.

## **Configure o tipo de linha para linhas de saída da transformação de Pesquisa.**

Adicione uma transformação de Estratégia de Atualização para sinalizar as linhas para inserção ou atualização.

## Inserir ou Atualizar

Use a propriedade **Inserir ou Atualizar** para atualizar linhas existentes no cache de pesquisa dinâmica quando o tipo de linha é inserir.

Essa propriedade aplica-se somente a linhas que inserem a transformação de Pesquisa com o tipo de linha inserir. Quando uma linha de qualquer outro tipo, como atualizar, acessa a transformação de Pesquisa, a propriedade **Inserir ou Atualizar** não tem efeito sobre como o Serviço de Integração manipula a linha.

Quando você selecionar **Inserir ou Atualizar** e o tipo de linha que insere a transformação de Pesquisa for de inserção, o Serviço de Integração inserirá a linha no cache, caso seja nova. Se a linha existir no cache de índice, mas o cache de dados for diferente da linha atual, o Serviço de Integração atualiza a linha no cache de dados.

Caso você não selecione **Inserir ou Atualizar** e o tipo de linha que entra na transformação de Pesquisa for de inserção, o Serviço de Integração insere a linha no cache se ela for nova e não efetua alterações no cache se a linha já existir.

A seguinte tabela descreve como o Serviço de Integração altera o cache de pesquisa quando o tipo das linhas recebidas na transformação de Pesquisa for inserir:

Opção para Inserir ou Atualizar	Linha encontrada no cache	Cache de dados diferente	Resultado de cache de pesquisa	Valor de NewLookupRow
Eliminado - apenas inserir	Sim	-	Sem alteração	0
Eliminado - apenas inserir	Não	-	Inserir	1



Opção para Inserir ou Atualizar	Linha encontrada no cache	Cache de dados diferente	Resultado de cache de pesquisa	Valor de NewLookupRow
Selecionado	Sim	Sim	Atualizar	2 <sup>1</sup>
Selecionado	Sim	Não	Sem alteração	0
Selecionado	Não	-	Inserir	1

*1. Se você selecionar Ignorar Nulo para todas as portas de pesquisa que não estejam na condição de pesquisa e se todas as portas contiverem valores nulos, o Serviço de Integração não alterará o cache e o valor de NewLookupRow será igual a 0.*

## Atualizar ou Inserir

Use a propriedade **Atualizar ou Inserir** para inserir novas linhas no cache de pesquisa dinâmica quando o tipo de linha é atualizar.

Você pode selecionar a propriedade **Atualizar ou Inserir** na transformação de Pesquisa. Essa propriedade se aplica apenas a linhas recebidas pela transformação de Pesquisa com o tipo de linha atualizar. Quando uma linha de qualquer outro tipo (como inserir) chega à transformação de Pesquisa, essa propriedade não tem efeito sobre como o Serviço de Integração manipula a linha.

Quando você seleciona essa propriedade e o tipo de linha recebido pela transformação de Pesquisa é atualizar, o Serviço de Integração atualizará a linha no cache se ela existir no cache de índice e os dados do cache forem diferentes dos da linha. O Serviço de Integração inserirá a linha no cache se ela for nova.

Se você não selecionar essa propriedade e o tipo de linha recebida pela transformação de Pesquisa for atualizar, o Serviço de Integração atualizará a linha no cache se ela existir e não alterará o cache se a linha for nova.

Se você selecionar **Ignorar Nulo** para todas as portas de pesquisa que não estejam na condição de pesquisa, e se todas as portas contiverem valores nulos, o Serviço de Integração não alterará o cache, e o valor de NewLookupRow será igual a 0.

A tabela a seguir descreve como o Serviço de Integração altera o cache de pesquisa quando o tipo das linhas recebidas na transformação de Pesquisa for atualizar:

Opção Atualizar ou Inserir	Linha encontrada no cache	Cache de dados diferente	Resultado de cache de pesquisa	Valor de NewLookupRow
Eliminado (apenas atualizar)	Sim	Sim	Atualizar	2
Eliminado (apenas atualizar)	Sim	Não	Sem alteração	0
Eliminado (apenas atualizar)	Não	-	Sem alteração	0
Selecionado	Sim	Sim	Atualizar	2
Selecionado	Sim	Não	Sem alteração	0
Selecionado	Não	-	Inserir	1

## Cache de Pesquisa Dinâmica e Destino de Sincronização

Configure as transformações downstream para garantir que o cache de pesquisa dinâmica e o destino sejam sincronizados.

Ao usar um cache de pesquisa dinâmica, o Serviço de Integração registra no cache de pesquisa antes de registrar na tabela de destino. O cache de pesquisa e a tabela de destino podem se tornar não sincronizados se o Serviço de Integração não registrar os dados no destino. Por exemplo, o banco de dados de destino pode rejeitar os dados.

Considere as seguintes orientações para manter o cache de pesquisa sincronizado com a tabela de pesquisa:

- Use uma transformação Roteador para passar linhas ao destino armazenado quando o valor NewLookupRow for igual a um ou dois.
- Use a transformação Roteador para descartar linhas quando o valor NewLookupRow for igual a zero. Ou, promova a saída das linhas para outro destino.
- Use transformações Estratégia de Atualização depois da transformação de Pesquisa para sinalizar linhas para inserção ou atualização no destino.
- Verifique se a transformação de Pesquisa gera para o destino os mesmos valores que o Serviço de Integração grava no cache de pesquisa. Se você quiser transferir valores novos na atualização, conecte apenas portas de pesquisa/saída à tabela de destino, em vez de portas de . Quando quiser transferir valores antigos na atualização, adicione uma transformação Expressão depois da transformação de Pesquisa e antes da transformação Roteador. Adicione portas de saída à transformação Expressão para cada porta na tabela de destino e crie expressões para garantir que valores nulos não sejam transferidos ao destino.
- Selecione Inserir e Atualizar como Atualização quando você define as opções de tabela de destino da estratégia de atualização. Isso garante que o Serviço de Integração irá atualizar as linhas marcadas para atualização e inserir linhas marcadas para inserção.

## Atualizações de Cache de Pesquisa Dinâmica Condicional

Você pode atualizar um cache de pesquisa dinâmica com base nos resultados de uma expressão booleana. O Serviço de Integração atualiza o cache quando a expressão é verdadeira.

Por exemplo, é possível que se tenha um número de produto, quantidade em estoque e uma coluna de carimbo de data/hora em uma tabela de destino. É necessário atualizar a quantidade em estoque com os valores de origem mais atuais. É possível atualizar a quantidade em estoque quando os dados de origem têm um carimbo de data/hora maior que o carimbo de data/hora do cache dinâmico. Crie uma expressão na Transformação Pesquisa que seja semelhante à expressão a seguir:

```
lookup_timestamp < input_timestamp
```

A expressão pode ter as portas de pesquisa e de entrada. É possível acessar variáveis internas, de mapeamento e de parâmetros. É possível incluir funções definidas pelo usuário e consultar transformações desconectadas.

A expressão retornará true, false ou NULL. Se o resultado da expressão for NULL, a expressão é falsa. O Serviço de Integração não atualiza o cache. É possível acrescentar uma marca para valores NULL na expressão se precisar alterar o resultado da expressão para true. O valor padrão da expressão é true.

## Processamento de Cache Dinâmico Condicional

Você pode criar uma condição que determina se o Serviço de Integração atualiza o cache de pesquisa dinâmica. Se a condição for falsa ou NULL, o Serviço de Integração não atualizará o cache de pesquisa dinâmica.

Se a condição for falsa ou nula, independentemente das propriedades da transformação de Pesquisa, o valor NewLookupRow é 0 e os Serviços de Integração não atualizam o cache de pesquisa dinâmica com nenhuma linha de inserção ou atualização.

Se a linha existe no cache e você ativa Inserir ou Atualizar ou Atualizar ou Inserir, o valor NewLookupRow é 1 e o Serviço de Integração atualiza a nova linha no cache.

Se a linha não existe no cache e você ativa Inserir ou Atualizar ou Atualizar ou Inserir, o valor NewLookupRow é 2 e Serviço de Integração insere a nova linha no cache.

## Configurando um Cache de Pesquisa Dinâmica Condicional

Você pode configurar uma expressão que determina se o Serviço de Integração atualiza um cache de pesquisa dinâmica.

1. Crie a Transformação Pesquisa.
2. Na guia **Tempo de Execução** da exibição **Propriedades**, selecione **Armazenamento da Pesquisa em Cache Ativado**.
3. Na guia **Avançado** da exibição **Propriedades**, selecione **Cache de Pesquisa Dinâmica**.
4. Para inserir a condição, clique na seta para baixo da propriedade **Atualizar Condição de Cache Dinâmico**.  
O Editor de Expressão será exibido.
5. Defina uma condição de expressão.  
É possível selecionar portas de entrada, portas de pesquisa e funções para a expressão.
6. Clique em **Validar** para verificar se a expressão é válida.
7. Clique em **OK**.
8. Se aplicável, configure as outras propriedades avançadas que se aplicam ao cache de pesquisa dinâmica.

## Atualização de cache dinâmico com resultados de expressões

A transformação de Pesquisa pode atualizar os valores do cache de pesquisa dinâmica com os resultados de uma expressão.

Por exemplo, um destino da tabela de produto tem uma coluna numérica que contém uma contagem dos pedidos. Cada vez que a transformação de Pesquisa recebe um pedido do produto, ela atualiza o cache dinâmico order\_count com os resultados da seguinte expressão:

```
order_count = order_count + 1
```

A transformação de Pesquisa retorna order\_count.

Você pode configurar como o Serviço de Integração trata o caso quando a expressão é avaliada como nula.

## Valores nulos da expressão

A expressão retorna NULL se um dos valores da expressão for nulo. No entanto, você pode configurar uma expressão para retornar um valor não nulo.

Se a expressão se refere a uma porta de pesquisa, mas os dados de origem são novos, a porta de pesquisa contém um valor padrão. O padrão pode ser NULL. Você pode configurar uma expressão IsNull para verificar valores nulos.

Por exemplo, a seguinte expressão verifica se lookup\_column é NULL:

```
iif (isnull(lookup_column), input_port, user_expression)
```

Se a coluna for nula, então é retornado o valor input\_port. Caso contrário, retorne o valor da expressão.

## Processamento de Expressão

O Serviço de Integração pode inserir e atualizar linhas no cache de pesquisa dinâmica com base nos resultados de expressões. Os resultados da expressão podem variar se o valor de porta de pesquisa é NULL e está incluído na expressão.

Quando você habilitar uma instrução Inserir ou Atualizar, o Serviço de Integração inserirá uma linha com o resultado da expressão se os dados não estiverem no cache. O valor da porta de pesquisa será NULO quando não houver dados no cache. Caso a expressão se refira ao valor da porta de pesquisa, o Serviço de Integração substituirá o valor da porta padrão na expressão. Quando você habilitar Inserir ou Atualizar e houver dados no cache, o Serviço de Integração atualizará o cache com o resultado da expressão.

Quando você habilitar Atualizar ou Inserir, o Serviço de Integração atualizará o cache com o resultado da expressão se houver dados no cache. Quando os dados não estiverem no cache, o Serviço de Integração inserirá uma linha que contenha o resultado da expressão. Se a expressão se referir a um valor de porta de pesquisa, o Serviço de Integração substituirá o valor de porta padrão na expressão.

## Configurando uma Expressão para Atualizações de Cache Dinâmico

Você pode configurar uma expressão para uma atualização de pesquisa de cache dinâmico.

É necessário ativar a Transformação Pesquisa para executar pesquisas dinâmicas antes de poder criar uma expressão condicional.

1. Crie a Transformação Pesquisa.
2. Na guia **Tempo de Execução** da exibição **Propriedades**, selecione **Armazenamento da Pesquisa em Cache Ativado**.
3. Na guia **Avançado** da exibição **Propriedades**, selecione **Cache de Pesquisa Dinâmica**.
4. Se aplicável, configure as outras propriedades avançadas que se aplicam ao cache de pesquisa dinâmica.
5. Para criar uma expressão, selecione a guia **Colunas** na guia **sPropriedades**.
6. Clique na seta da lista suspensa da coluna **Porta Associada** para a porta de pesquisa que você deseja atualizar.
7. Selecione **Expressão Associada** na lista suspensa e clique em **Inserir**.  
O Editor de Expressão será exibido.
8. Defina a expressão.

É possível selecionar portas de entrada, portas de pesquisa e funções para a expressão. O valor de retorno da expressão deve corresponder ao tipo de dados da porta de pesquisa.

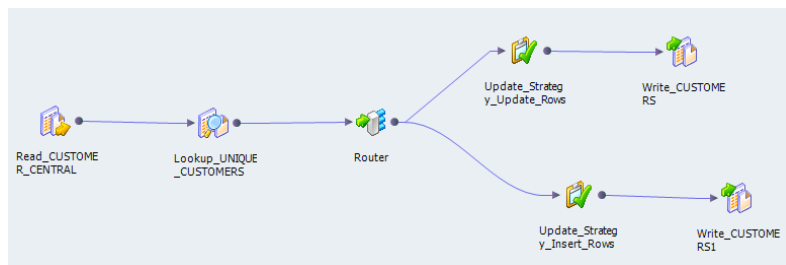
9. Clique em **Validar** para verificar se a expressão é válida.
10. Clique em **OK**.

## Exemplo de Cache de Pesquisa Dinâmica

Você pode usar um cache de pesquisa dinâmica para inserir e atualizar linhas no destino. Quando você usar um cache de pesquisa dinâmica, pode inserir e atualizar as mesmas linhas do cache que você inserir e atualizar no destino.

Por exemplo, você precisa atualizar uma tabela que contém dados do cliente. Os dados de origem contêm linhas de dados do cliente para inserir ou atualizar no destino. Crie um cache dinâmico que represente o destino. Configure uma transformação de Pesquisa para os clientes de pesquisa no cache dinâmico.

A seguinte figura mostra um mapeamento que possui uma transformação de Pesquisa com um cache de pesquisa dinâmica:



A transformação de Roteador é dividida em duas ramificações. A transformação de Roteador transmite as linhas de inserção em uma ramificação e as atualiza em outra. Cada ramificação contém uma transformação de Estratégia de Atualização que grava linhas no destino. As duas ramificações contêm o mesmo destino.

Quando o mapeamento inicia, o Serviço de Integração cria o cache de pesquisa a partir de clientes de tabela de destino. Quando o Serviço de Integração lê uma linha que não está no cache de pesquisa, ele insere a linha no cache.

A transformação de Pesquisa retorna cada linha para uma transformação Roteador. Com base na marcação da linha para inserção ou atualização, a transformação Roteador direciona a linha para uma das transformações Estratégia de Atualização. A transformação Roteador determina se as linhas estão marcadas para inserção ou atualização com base na propriedade NewLookupRow. As transformações Estratégia de Atualização marcam cada linha como inserção ou atualização antes de transferi-la para o destino.

A tabela de destino de clientes é alterada à medida que o mapeamento é executado. O Serviço de Integração insere novas linhas e atualiza as existentes no cache de pesquisa. O Serviço de Integração mantém o cache de pesquisa e a tabela de destino de clientes sincronizados.

Para gerar chaves para o destino, use Sequence-ID na porta associada. O Serviço de Integração usa a ID de sequência como a chave primária para cada nova linha inserida na tabela de destino.

O desempenho da sessão melhora quando você usa um cache de pesquisa dinâmica, pois o cache é criado no banco de dados uma vez.

# Regras e Diretrizes para Cache de Pesquisa Dinâmica

Use as seguintes diretrizes ao usar o cache de pesquisa dinâmica:

- Você deverá definir a propriedade **Em Várias Correspondências** como Relatar Erro quando usar um cache de pesquisa dinâmica. Para redefinir a propriedade, altere a pesquisa dinâmica para uma pesquisa estática, altere a propriedade e, em seguida, altere a pesquisa estática para uma pesquisa dinâmica.
- Não se pode compartilhar o cache entre uma transformação de Pesquisa dinâmica e outra estática no mesmo grupo de ordem de carregamento de destino.
- Você pode ativar um cache de pesquisa dinâmica para uma pesquisa relacional ou de arquivo simples.
- A transformação de Pesquisa precisa ser uma transformação conectada.
- Você pode usar um cache persistente ou não persistente.
- Se o cache não for persistente, o Serviço de Integração recompilará sempre o cache a partir do banco de dados, mesmo se você não ativar **Efetuar Novamente o Cache de Origem da Pesquisa**.
- Só é possível criar uma condição de pesquisa de igualdade. Você não pode procurar um intervalo de dados em um cache dinâmico.
- Você deve associar cada porta de pesquisa que não esteja na condição de pesquisa a uma porta de entrada, uma ID de sequência ou uma expressão associada.
- Use uma transformação Roteador para passar linhas ao destino armazenado quando o valor NewLookupRow for igual a um ou dois.
- Use a transformação Roteador para descartar linhas quando o valor NewLookupRow for igual a zero. Ou você pode gerar a saída de linhas para outro destino.
- Verifique se o Serviço de Integração gera os mesmos valores que grava no cache de pesquisa para o destino. Se você quiser transferir valores novos na atualização, conecte apenas portas de pesquisa e saída à tabela de destino, em vez de as portas de entrada e saída. Quando quiser transferir valores antigos na atualização, adicione uma transformação Expressão depois da transformação de Pesquisa e antes da transformação Roteador. Adicione portas de saída à transformação Expressão para cada porta na tabela de destino e crie expressões para garantir que valores nulos não sejam transferidos ao destino.
- Ao usar uma substituição SQL de pesquisa, mapeie as colunas certas aos destinos adequados para pesquisa.
- Ao adicionar uma cláusula WHERE à substituição SQL de pesquisa, use uma transformação Filtro antes da transformação de Pesquisa. Isso garante que o Serviço de Integração insira ao cache dinâmico e à tabela de destino linhas que coincidam com a cláusula WHERE.
- Quando você configura uma transformação de Pesquisa reutilizável para usar cache dinâmico, não é possível editar a condição nem desativar a propriedade **Cache de Pesquisa Dinâmica** em um mapeamento.
- Use transformações Estratégia de Atualização depois da transformação de Pesquisa para sinalizar as linhas a inserir ou atualizar para o destino.
- Use uma transformação Estratégia de Atualização antes da transformação de Pesquisa para definir algumas ou todas as linhas como atualização, se quiser usar a propriedade Atualizar ou Inserir na transformação de Pesquisa.

## CAPÍTULO 28

# Transformação de Correspondência

Este capítulo inclui os seguintes tópicos:

- [Transformação de Correspondência - Visão Geral, 459](#)
- [Análise de Correspondência, 460](#)
- [Cálculos de Pontuação de Correspondência, 464](#)
- [Análise de Dados Mestre, 467](#)
- [Análise de Correspondência de Identidade e Dados de Indexação Persistente, 469](#)
- [Desempenho do Mapeamento de Correspondência, 470](#)
- [Desempenho de Correspondência na Análise de Identidade, 472](#)
- [Exibições de Transformação de Correspondência, 475](#)
- [Portas de Transformação de Correspondência, 476](#)
- [Mapplets de Correspondência, 481](#)
- [Configurando uma Operação de Análise de Correspondência, 483](#)
- [Transformação de Correspondência em um ambiente não nativo, 484](#)

## Transformação de Correspondência - Visão Geral

A transformação de Correspondência é uma transformação ativa que analisa os níveis de similaridade entre registros. Use a transformação de Correspondência para localizar registros que contenham informações duplicadas em um conjunto de dados ou entre dois conjuntos de dados.

A transformação de Correspondência analisa os valores em uma porta de entrada e gera um conjunto de pontuações numéricas que representam os graus de semelhança entre os valores. É possível selecionar várias portas para determinar os níveis gerais de semelhança entre os registros de entrada. Você pode especificar uma pontuação mínima como valor de limite para identificar os registros que podem conter informações duplicadas.

Você pode usar a transformação de Correspondência nos seguintes projetos de dados:

- Gerenciamento do Relacionamento com Clientes. Por exemplo, uma loja cria uma campanha por e-mail e precisa verificar se há registros duplicados no banco de dados de clientes.
- Fusões e aquisições. Por exemplo, um banco compra outro banco na mesma região, e os dois têm clientes em comum.

- Iniciativas de conformidade regulamentar. Por exemplo, uma empresa opera sob regulamentações do governo ou do setor que insistem que todos os sistemas de dados estejam isentos de registros duplicados.
- Gerenciamento de risco financeiro. Por exemplo, um banco pode querer pesquisar relacionamentos entre titulares de contas.
- Gerenciamento de dados mestre. Por exemplo, uma cadeia de varejo tem um banco de dados mestre de registros de clientes, e cada loja da cadeia envia registros ao banco de dados mestre regularmente.
- Qualquer projeto que deve identificar registros duplicados em um conjunto de dados.

## Análise de Correspondência

Você pode definir diferentes tipos de análises de duplicatas na transformação de Correspondência. As operações de análise de duplicatas que você define dependem do número de fontes de dados no mapeamento e do tipo de informações que as origens contêm.

Considere os seguintes fatores ao configurar uma transformação de Correspondência:

- É possível selecionar uma única coluna de um conjunto de dados ou várias colunas.
- Você pode analisar colunas de uma única fonte de dados ou pode analisar duas fontes de dados.
- Você pode configurar a transformação de Correspondência para analisar os dados brutos nos campos de portas de entrada ou pode configurar a transformação para analisar as informações de identidade nos dados.
- É possível configurar a transformação de Correspondência para gravar diferentes tipos de saída. O tipo de saída selecionado determina o número de registros gravados pela transformação e a ordem desses registros.
- Para melhorar o desempenho, classifique os registros de entrada em grupos antes de realizar a análise de correspondência.

## Análise de Colunas

Quando você configurar uma transformação de Correspondência, selecione uma ou mais colunas para análise.

A transformação de Correspondência analisa as colunas em pares. Quando você seleciona uma única coluna para análise, a transformação cria uma cópia temporária dela, comparando a coluna de origem com a coluna temporária. Quando você seleciona duas colunas para análise, a transformação compara os valores entre as duas colunas que você selecionou. A transformação compara cada valor em uma coluna com todos os valores na outra coluna. A transformação retorna uma pontuação de correspondência para cada par de valores que ela analisa.

Selecione as colunas a serem analisadas quando você configurar uma estratégia na transformação de Correspondência. A estratégia especifica as colunas a serem analisadas e o algoritmo a ser aplicado às colunas. O algoritmo calcula os níveis de similaridade entre cada par de valores. Os diferentes algoritmos na transformação usam diferentes critérios para medir os níveis de similaridade entre os valores. Você pode definir várias estratégias em uma transformação, além de atribuir colunas diferentes a cada estratégia.

### Exemplo de Análise de Coluna

Você deseja comparar os valores em uma coluna de dados de sobrenome. Você cria um mapeamento que inclui uma fonte de dados e uma transformação de Correspondência. Você conecta a porta *Sobrenome* à



transformação de Correspondência. A transformação cria uma cópia temporária dos dados na porta *Sobrenome* quando o mapeamento é executado.

A seguinte imagem mostra um fragmento dos dados de sobrenome:

	A	B
1	<b>Surname</b>	<b>Surname_1</b>
2	Annan	Annan
3	Baker	Baker
4	Barker	Barker
5	Edwards	Edwards
6	Parker	Parker
7	Smith	Smith
8	Smith	Smith
9	Zhang	Zhang

O mapeamento gera um conjunto de pontuações de correspondência que indicam que os seguintes valores podem ser duplicatas:

- Baker, Barker
- Barker, Parker
- Smith, Smith

Ao analisar os dados, você decide que *Baker*, *Barker* e *Parker* não são valores duplicados. Você decide que *Smith* e *Smith* são valores duplicados.

## Análise de Origem Única e Análise de Origem Dupla

É possível configurar a transformação de Correspondência para analisar dados de uma ou duas fontes de dados. Selecione as portas de cada fonte de dados ao definir uma estratégia na transformação.

Ao configurar a transformação para realizar a análise de origem única, você seleciona uma ou mais portas de um único conjunto de dados. Ao configurar a transformação para realizar a análise de origem dupla, você seleciona uma ou mais portas de cada conjunto de dados. As portas devem ser selecionadas em pares. Para cada par de portas selecionado, a transformação compara cada valor em uma porta com cada valor na outra porta. Se você executar a análise de origem única nos dados de uma única coluna, a transformação criará uma cópia temporária da porta que você selecionar.

**Nota:** Ao realizar a análise de correspondência de identidades, você pode comparar uma fonte de dados a um índice persistente de dados de identidade criado em um mapeamento anterior. Use as opções de **Tipo de Correspondência** para especificar uma análise de identidades com um índice persistente.

## Análise de Correspondência de Campos e Análise de Correspondência de Identidade

É possível configurar uma transformação de Correspondência para realizar uma análise de correspondência de campos ou uma análise de correspondência de identidade.

Na análise de correspondência de campos, a transformação de Correspondência analisa os dados de origem que entram na transformação. Você pode realizar a análise de correspondência de campos em qualquer tipo de dados. Na análise de correspondência de identidade, a transformação de Correspondência gera um índice de valores de dados alternativos dos dados de entrada e analisa os dados de índice. Configure a transformação de Correspondência para a análise de correspondência de identidade quando as portas de

entrada contiverem dados de identidade. Uma identidade é um grupo de valores de dados que identifica uma pessoa ou uma organização.

Um conjunto de dados pode representar uma única identidade de formas diferentes. Por exemplo, os seguintes valores de dados representam o nome John Smith:

- John Smith
- Smith, John
- jsmith@email.com
- SMITHJMR

A transformação de Correspondência lê os dados de identidade em um registro e calcula as possíveis versões alternativas da identidade. A transformação cria um índice que inclui as versões atuais e as versões alternativas das identidades. A transformação de Correspondência analisa os valores de índice e não os valores nos registros de entrada.

### Arquivos de Preenchimento de Identidade

As operações de correspondência de identidade fazem a leitura de arquivos de dados de referência, chamados de preenchimentos. Os arquivos de preenchimento definem as possíveis variações nos dados de identidade. Os arquivos não são instalados com os aplicativos Informatica. Você deve adquirir e baixar os arquivos de dados de preenchimento da Informatica.

Instale os arquivos em uma localização que o Serviço do Gerenciamento de Conteúdo possa acessar. Use o Informatica Administrator para definir a localização no Serviço do Gerenciamento de Conteúdo.

## Grupos na Análise de Correspondência

Um mapeamento de correspondência de análise pode demorar a ser executado devido ao número de comparações de dados que a transformação deve realizar. O número de comparações está relacionado ao número de valores de dados nas portas que você seleciona.

A seguinte tabela mostra o número de cálculos que um mapeamento realiza para diferentes números de valores de dados em uma única porta:

Número de valores de dados	Número de comparações
10.000	50 milhões
100.000	5.000 milhões
1 milhão	500.000 milhões

Para reduzir o tempo gasto para executar o mapeamento, atribua os registros de dados de entrada a grupos. Um grupo é um conjunto de registros que contêm valores idênticos em uma porta que você especifica. Quando você executa a análise de correspondência em dados agrupados, a transformação de Correspondência analisa os registros em cada grupo. A transformação não compara os registros em um grupo com os registros em outro grupo. Os grupos reduzem o número total de comparações que a transformação deve realizar sem qualquer perda de precisão na análise de mapeamento.

Considere as seguintes regras e diretrizes ao organizar dados em grupos:

- A porta na qual você agrupa os dados é a porta de chave de grupo. Uma porta de chave de grupo deve conter um intervalo de valores duplicados, como um nome de cidade ou de estado em um conjunto de dados de endereço. Se o mapeamento de dados não contiver uma porta de chave de grupo utilizável, use o Gerador de Chaves para criar a porta dos dados de mapeamento atuais. Conecte a porta de saída da chave de grupo da transformação do Gerador de Chaves para a transformação de Correspondência. Você também pode usar a transformação do Gerador de Chaves para adicionar identificadores de sequência aos dados de mapeamento.
- As operações de correspondência de campo devem especificar uma porta de chave de grupo. Se você configurar a transformação de Correspondência para análise de identidade, não selecione uma porta de chave de grupo. A análise de identidade gera chaves de grupo para os dados de índice de identidade.
- Não especifique uma porta de chave de grupo que você pretenda utilizar na análise de correspondência.
- Ao criar grupos, você deve verificar se eles são de um tamanho válido. Se os grupos forem muito pequenos, a análise de correspondência talvez não encontre todos os dados duplicados no conjunto de dados. Se os grupos forem muito grandes, a análise de correspondência poderá retornar falsas duplicatas. Selecione as chaves de grupo que criam um grupo com tamanho médio de 10.000 registros.
- Os grupos não reorganizam a posição dos registros no conjunto de dados de mapeamento.

## Pares de Correspondência e Clusters

A transformação de Correspondência pode ler e gravar diferentes números de linhas de entrada e linhas de saída, além de poder alterar a sequência das linhas de saída. Você determina o formato de saída dos resultados da análise de correspondência.

A transformação pode gravar linhas nos seguintes formatos:

### Pares correspondentes

A transformação grava uma linha para cada par de registros correspondentes a uma pontuação que se enquadra no limite de correspondência. A transformação grava cada par de registros em uma única linha.

Como um registro pode corresponder mais de um outro registro, ele pode aparecer em mais de uma linha de saída.

### Melhor correspondência

A transformação grava uma linha para cada registro em um conjunto de dados e adiciona o registro mais semelhante de outro conjunto de dados à mesma linha.

### Clusters

A transformação atribui os registros de saída a clusters com base nos níveis de similaridade entre os registros. Um cluster é um conjunto de registros no qual cada registro corresponde a pelo menos um outro registro com uma pontuação que se enquadra no limite de correspondência. A transformação grava cada registro em uma única linha.

Cada registro em um cluster deve corresponder a pelo menos um outro registro nesse mesmo cluster. Portanto, um cluster pode conter pares de registros que não correspondem entre si. Um cluster poderá conter um único registro se este não corresponder a nenhum outro registro.

**Nota:** A opção Clusters na análise de campo corresponde à opção Clusters - Corresponder Tudo em uma análise da identidade. A opção Clusters - Melhor Correspondência em uma análise de identidade combina cálculos de cluster e cálculos de pares correspondentes.

Configure as opções de saída na exibição **Saída de Correspondência** da transformação.

## TÓPICOS RELACIONADOS:

- [“Opções de Saída de Cluster” na página 465](#)

# Cálculos de Pontuação de Correspondência

A pontuação de correspondência é um valor numérico que representa o grau de similaridade entre dois valores de coluna. Um algoritmo calcula uma pontuação de correspondência como um valor decimal no intervalo de 0 a 1. Um algoritmo atribui uma pontuação de 1 quando dois valores de coluna são idênticos.

Quando você seleciona vários pares de coluna para análise, a transformação calcula uma pontuação média com base nas pontuações das colunas selecionadas. Por padrão, a transformação atribui peso igual às pontuações de cada par de colunas. A transformação não infere a importância relativa dos dados da coluna no conjunto de dados.

Você pode editar os valores de peso usados pela transformação para calcular a pontuação de correspondência. Edite os valores de peso quando quiser atribuir prioridade mais alta ou mais baixa às colunas no conjunto de dados.

Você também pode definir as pontuações aplicadas pela transformação quando um valor nulo é encontrado em uma coluna. Por padrão, a transformação trata valores nulos como erros de dados e atribui uma pontuação de correspondência baixa a qualquer par de valores que contenha um valor nulo.

**Nota:** O algoritmo selecionado determina a pontuação de correspondência entre dois valores. O algoritmo gera uma única pontuação para os dois valores. As pontuações de correspondência não dependem do tipo de saída de correspondência ou do tipo de método de pontuação selecionado.

## Pontuações Ponderadas

Quando você seleciona várias colunas para análise de correspondência, a transformação calcula uma pontuação média para cada registro com base nas pontuações das colunas. A pontuação média inclui qualquer valor de peso aplicado aos algoritmos de comparação para cada coluna.

Por padrão, todos os algoritmos usam um valor de peso 0,5. Você poderá aumentar esse valor se as colunas selecionadas tiverem maior probabilidade de conter informações duplicadas. O valor de peso poderá ser diminuído se valores duplicados nas colunas selecionadas tiverem menor probabilidade de indicar informações duplicadas genuínas entre os registros. A transformação de Correspondência usa a pontuação média como a única pontuação de correspondência para cada par de registros.

## Pontuações de Correspondência Nulas

Um algoritmo de correspondência aplica uma pontuação de correspondência predefinida a um par de valores quando um ou ambos os valores são nulos. Você pode editar a pontuação de correspondência aplicada por um algoritmo de correspondência de campo a valores nulos.

### Pontuações de Correspondência Nula e Algoritmos de Correspondência de Campo

Ao configurar um algoritmo de correspondência de campo, verifique os valores de pontuação de correspondência aplicados pelo algoritmo a dados nulos. Um algoritmo de correspondência de campo aplica uma pontuação padrão de 0,5 quando compara dois valores, e um ou ambos os valores são nulos. Uma pontuação de 0,5 indica um baixo nível de similaridade entre os valores de dados.

Considere as seguintes regras e diretrizes ao verificar as pontuações de correspondência nulas:

- Quando o algoritmo analisar as colunas que contêm as chaves primárias ou outros dados críticos, não edite as pontuações padrão. Nesse caso, um valor nulo representa um erro de dados, e as pontuações padrão são apropriadas para os dados.
- Quando o algoritmo analisar colunas que podem opcionalmente conter dados, atualize os valores de pontuação de correspondência nula para o mesmo valor do limite de correspondência. Você cancela o efeito dos valores nulos na análise de correspondência quando define as pontuações de correspondência nula de acordo com o valor de limite de correspondência.

### Pontuações de Correspondência Nula e Algoritmos de Correspondência de Identidade

Um algoritmo de correspondência de identidade aplica uma pontuação de correspondência de 0 quando compara dois valores, e um ou ambos os valores são nulos. A análise de correspondência de identidade atribui um registro com uma pontuação de correspondência nula a um cluster de registro exclusivo e registra um valor de tamanho de cluster de 1. Não é possível editar a pontuação aplicada por um algoritmo de correspondência de identidade a dados nulos.

## Opções de Saída de Cluster

Selecione uma opção de saída de cluster quando quiser organizar registros semelhantes ou idênticos nos dados de saída.

Quando você seleciona uma opção de saída de cluster, a transformação adiciona um valor de ID de cluster a cada registro de saída. É possível classificar os registros pelos valores de ID de cluster. A saída da transformação inclui uma linha para cada registro. Se um registro não corresponder a outro registro com uma pontuação que atenda ao limiar de correspondência, a transformação atribuirá um ID de cluster exclusivo ao registro. Use a exibição **Saída de Correspondência** para selecionar ou atualizar as opções de saída de cluster.

É possível selecionar as seguintes opções de saída de cluster:

#### Clusters

Selecione a opção para atribuir valores de ID de cluster aos registros de saída.

#### Clusters - Melhor Correspondência

Selecione a opção para adicionar o par de registros com a maior pontuação de correspondência a um cluster. Como um registro pode representar a melhor correspondência com mais de um outro registro, mais de um par de registros pode compartilhar um valor de ID de cluster.

#### Clusters - Corresponder Tudo

A opção **Clusters - Corresponder Tudo** funciona da mesma maneira que a opção **Clusters**.

A transformação usa **Clusters - Corresponder Tudo** e **Clusters - Melhor Correspondência** como nomes de opções na análise de correspondência de identidades.

**Nota:** Se um Serviço de Integração de Dados executar várias transformações de Correspondência ao mesmo tempo, ele gerará valores exclusivos de ID de cluster para a saída de cada transformação. Portanto, os valores de ID de cluster para os registros gerados por cada transformação podem ser não consecutivos.

### As Opções Clusters e Clusters - Corresponder Tudo

Selecione a opção **Clusters** na análise de correspondência de campos. Selecione a opção **Clusters - Corresponder Tudo** na análise de correspondência de identidades.

A transformação de Correspondência usa as seguintes regras para criar os clusters:

- Quando dois registros têm uma pontuação de correspondência que atende ao limite de correspondência, a transformação de Correspondência adiciona esses registros a um cluster.
- Quando um registro no conjunto de dados corresponde a qualquer registro no cluster, a transformação adiciona esse registro ao cluster.
- Se um registro em um cluster corresponde a um registro em outro cluster, o processo mescla os clusters.
- A transformação realiza uma varredura contínua dos os resultados de correspondência até que todos os registros pertençam a um cluster.
- Se um registro não corresponder a nenhum outro registro no conjunto de dados, a transformação atribuirá um valor de ID de cluster exclusivo a esse registro.

### A Opção Clusters - Melhor Correspondência

Selecione a opção Clusters - Melhor Correspondência na análise de correspondência de identidades.

A transformação usa as seguintes regras para criar os clusters:

- A transformação identifica o registro que tem a pontuação de correspondência mais alta com o registro atual. Se a pontuação de correspondência atingir o limite, a transformação adicionará o par de registros a um cluster.
- Se um dos registros correspondentes estiver em um cluster, a transformação adicionará o outro registro ao cluster atual.
- A transformação realiza uma varredura contínua dos os resultados da pontuação de correspondência até que todos os registros pertençam a um cluster.
- Um cluster poderá conter um único registro se este não corresponder a nenhum outro registro nos dados.

**Nota:** É possível usar a propriedade **Correspondência** na exibição **Saída de Correspondência** para especificar como a transformação compara uma única fonte de dados a um armazenamento de dados persistente. A propriedade **Correspondência** determina se a transformação procura duplicatas nos dados de origem ou no armazenamento de dados persistente.

### TÓPICOS RELACIONADOS:

- [“Pares de Correspondência e Clusters” na página 463](#)

## Pontuações de Controle e Pontuações de Vínculo na Análise de Cluster

Ao selecionar uma opção de saída de cluster na transformação de Correspondência, você pode adicionar dados de pontuação de vínculo e de controle à saída.

A pontuação de vínculo é a pontuação entre dois registros que os identifica como membros do mesmo cluster. Os vínculos entre os registros determinam a composição do cluster. Qualquer registro pode se vincular a qualquer outro registro no mesmo cluster.

A pontuação de controle é a pontuação entre o registro com o valor de ID de sequência mais alto em um cluster e outro registro no mesmo cluster. As pontuações de controle fornecem um meio de avaliar todos os registros em um cluster em relação a um único registro. Quando você adiciona pontuações de controle à saída de correspondência, o mapeamento é executado mais lentamente, pois a transformação de Correspondência não poderá calcular as pontuações até que todos os clusters estejam completos.

**Nota:** A análise de correspondência gera um único conjunto de pontuações para cada estratégia que você define. A pontuação de controle e a pontuação de vínculo indicam as pontuações de correspondência para diferentes pares de registros em cada cluster. As pontuações de controle e as pontuações de vínculo podem

depende da ordem em que os registros entram na transformação. A pontuação de controle pode ser menor que o limite de correspondência.

### Exemplo de Análise de Cluster

Você configura uma estratégia de correspondência de campos para analisar uma coluna de dados de sobrenome. Você define um limite de correspondência de 0,825 na estratégia. Você seleciona um formato de saída clusterizado e executa o Visualizador de Dados na transformação.

A seguinte tabela mostra os dados exibidos pelo Visualizador de Dados:

Sobrenome	ID da Sequência	ID de Cluster	Tamanho do Cluster	ID do Driver	Pontuação de Driver	ID do Link	Pontuação de Vínculo
SMITH	1	1	2	1 - 6	1	1 - 1	1
SMYTH	2	2	2	1 - 3	0.83333	1 - 2	1
SMYTHE	3	2	2	1 - 3	1	1 - 2	0.83333
SMITT	4	3	1	1 - 4	1	1 - 4	1
SMITS	5	4	1	1 - 5	1	1 - 5	1
SMITH	6	1	2	1 - 6	1	1 - 1	1

O Visualizador de Dados contém as seguintes informações sobre os dados de sobrenome:

- SMITT e SMITS não correspondem a nenhum registro com uma pontuação que se enquadra no limite de correspondência. A transformação de Correspondência determina que os registros são exclusivos no conjunto de dados.  
SMITT e SMITS têm um tamanho de cluster de 1. Para localizar registros exclusivos na saída de cluster, procure clusters que contêm um único registro.
- SMITH e SMITH têm uma pontuação de vínculo de 1. A transformação de Correspondência determina que os registros são idênticos. A transformação adiciona os registros a um único cluster.
- SMYTH e SMYTHE têm uma pontuação de vínculo de 0,83333. A pontuação excede o limite de correspondência. Portanto, a transformação adiciona os registros a um único cluster.

## Análise de Dados Mestre

Ao analisar duas fontes de dados na transformação de Correspondência, você deve identificar uma origem como o conjunto de dados mestre. A transformação compara os valores de dados de cada registro do conjunto de dados especificado com os valores correspondentes em cada registro do segundo conjunto de dados.

Em várias organizações, um conjunto de dados mestre constitui um armazenamento de dados permanente de alta qualidade. Antes de adicionar registros a um conjunto de dados mestre, use a transformação de Correspondência para verificar se os registros não adicionam informações duplicadas aos dados mestre.

## Exemplo de Dados Mestre

Um banco mantém um conjunto de dados mestre de registros de contas de clientes. O banco atualiza o conjunto de dados mestre todos os dias com registros que identificam novas contas de clientes. O banco usa um mapeamento de análise de duplicatas para verificar se os novos registros duplicam as informações de clientes no conjunto de dados mestre. O conjunto de dados mestre e a tabela de novas contas têm uma estrutura em comum, e as tabelas usam o mesmo tipo de banco de dados. Portanto, o banco pode reutilizar o mapeamento de análise de duplicatas sempre que ele precisar atualizar o conjunto de dados mestre.

## Direcionalidade na Análise do Conjunto de Dados Mestre

A transformação de Correspondência compara registros entre dois conjuntos de dados em uma única direção. A transformação compara cada registro no conjunto de dados mestre com todos os registros no segundo conjunto de dados. Ela não compara cada registro no segundo conjunto de dados com todos os registros no conjunto de dados mestre. Portanto, a seleção do conjunto de dados mestre pode afetar os resultados da análise de correspondência.

A seguinte tabela mostra dois conjuntos de dados que você pode comparar em uma análise de correspondência de identidades:

Conjunto de Dados 1	Conjunto de Dados 2
Alex Bell	Alexander Bell
Alexander Graham Bell	Thomas Edison
Alva Edison	Nicola Tesla
Marie Curie	Irene Joliot Curie
Dorothy Crowfoot	Dorothy Hodgkin

Se você selecionar Conjunto de Dados 1 como o conjunto de dados mestre e selecionar a opção de saída **Melhor Correspondência**, a saída incluirá os seguintes registros:

- Alex Bell, Alexander Bell
- Alexander Graham Bell, Alexander Bell

Se você selecionar Conjunto de Dados 2 como o conjunto de dados mestre e selecionar a opção de saída **Melhor Correspondência**, a saída incluirá os seguintes registros:

- Alexander Bell, Alex Bell

Quando Conjunto de Dados 2 é o conjunto de dados mestre, a transformação não pode corresponder Alexander Bell e Alexander Graham Bell, pois Alexander Bell já corresponde a Alex Bell nos dados de saída.

## Reutilização de Mapeamentos

Se você adiciona dados a um conjunto de dados mestre regularmente, configure um mapeamento de análise de duplicatas que possa ser reutilizado. Você poderá reutilizar o mapeamento se as configurações da porta na fonte de dados que você compara com o conjunto de dados mestre não forem alteradas.

Ao executar o mapeamento, verifique se a transformação especifica os dados mestre e os dados mais recentes. Você pode executar o mapeamento sem qualquer outra atualização de configuração.



# Análise de Correspondência de Identidade e Dados de Indexação Persistente

Quando você executa um mapeamento para analisar as informações de identidade, a transformação de Correspondência gera um índice que armazena as versões alternativas das identidades no conjunto de dados. Por padrão, a transformação de Correspondência grava os dados de índice em arquivos temporários. Você pode configurar a transformação para salvar os dados de índice em tabelas de banco de dados.

As tabelas de índice que você gera representam um armazenamento de dados que pode ser reutilizado em mapeamentos subsequentes. Você pode comparar as tabelas de índice com uma fonte de dados e, opcionalmente, atualizá-las com dados de índice da fonte de dados. Como a transformação não gera novamente as tabelas de índice, os mapeamentos subsequentes são executados mais rapidamente. Além disso, as tabelas de índice podem representar um armazenamento de dados confiável de dados de identidade.

## TÓPICOS RELACIONADOS:

- [“Desempenho de Correspondência na Análise de Identidade” na página 472](#)

## Regras e Diretrizes para Dados de Índice Persistentes

Considere as seguintes regras e diretrizes ao configurar uma transformação de Correspondência para analisar um conjunto de dados mestre de informações de identidade:

- Para gerar um índice reutilizável para um conjunto de dados mestre, configure a transformação para gravar os dados de índice nas tabelas de banco de dados. As tabelas de banco de dados constituem um armazenamento persistente de dados de índice.
- Para comparar as identidades em outro conjunto de dados com um armazenamento de dados de índice, configure o conjunto de dados como a fonte de dados do mapeamento. Configure a transformação de Correspondência para ler a fonte de dados e o armazenamento de dados de índice. Selecione as tabelas de índice do esquema padrão da conexão de banco de dados que você especificar.
- A transformação de Correspondência adiciona o valor do identificador de sequência do registro de entrada às linhas dos dados de índice que correspondem ao registro. A porta de entrada *SequenceID* contém os identificadores de sequência. A transformação usa os identificadores de sequência para rastrear os dados de índice pelas diferentes etapas na análise de correspondência. Não desconecte a porta de ID de sequência.
- Quando você conecta uma transformação de Correspondência a um armazenamento de índice, a transformação reutiliza os valores de propriedades de preenchimento, nível de chave, tipo de chave e campo de chave dos valores da transformação que criou esse armazenamento. A transformação também reutiliza a configuração de porta da transformação que criou o armazenamento.

Se as propriedades da transformação não corresponderem, a análise de identidade não poderá comparar os dados de origem do mapeamento e os dados de índice corretamente.

- A transformação de Correspondência usa os dados na porta de entrada que você seleciona como o campo de chave para gerar o índice de identidade. Ela também pode gravar dados de outras portas no índice. Se você desconectar da transformação as portas de dados de campos não relacionados a chaves, todos os dados nas colunas de índice correspondentes serão apagados quando executar o mapeamento. Para preservar os dados de portas de entrada nas tabelas de índice, não desconte as portas de dados de entrada.
- Você pode desativar a análise de correspondência na transformação de Correspondência ao gerar os dados da tabela de índice para um conjunto de dados. Por exemplo, você pode desativar a análise de

correspondência ao criar um armazenamento de índice para um conjunto de dados. Quando você desativa a análise de correspondência, o mapeamento é executado mais rapidamente.

Quando você desativa a análise de correspondência, a transformação de Correspondência pode gerar e exibir códigos de status de persistência e descrições de status de persistência. A transformação não gera nem exibe pontuações de correspondência ou outros dados associados aos resultados da análise de correspondência. Por exemplo, se você configurar a transformação para atribuir registros a clusters e desativar a análise de correspondência, a transformação não gerará nem exibirá valores de ID de cluster.

- Determine se a transformação de Correspondência atualiza o armazenamento de índice com os dados da origem de mapeamento. A transformação de Correspondência usa identificadores de sequência para determinar se as linhas no armazenamento de índice e os dados do mapeamento representam os mesmos registros.

## Desempenho do Mapeamento de Correspondência

É possível visualizar os fatores de dados que determinam o desempenho da transformação de Correspondência antes de executar o mapeamento que contém essa transformação. Você pode verificar se o sistema tem os recursos necessários para executar o mapeamento. Você também pode verificar se configurou a transformação corretamente para medir os níveis de similaridade nos dados de entrada.

Use a opção **Análise de Desempenho de Correspondência** para verificar se o sistema tem os recursos necessários. Use a opção **Análise de Cluster de Correspondência** para verificar se o mapeamento pode medir com precisão os níveis de similaridade dos dados de entrada.

Execute a análise de desempenho de correspondência e a análise de cluster de correspondência em qualquer transformação de Correspondência que leia uma única fonte de dados. Execute a análise de desempenho de correspondência em qualquer transformação de Correspondência que realize a análise de correspondência de campo de origem dupla. Não execute a análise de desempenho de correspondência ou a análise de cluster de correspondência em uma estratégia de correspondência de identidade que se conecte a tabelas de índice.

### Busca Detalhada na Análise de Desempenho de Correspondência

Você pode fazer uma busca detalhada nos dados de análise de correspondência para exibir os pares de registros que atendem ou excedem o limite de correspondência. Clique duas vezes em um registro na exibição **Detalhes** e use o Visualizador de Dados para exibir os registros que correspondem ao registro que você selecionar. O Visualizador de Dados exibe os dados de cada par de registros em uma única linha. A linha contém o identificador de linha de cada registro no par.

### Busca Detalhada na Análise de Cluster de Correspondência

Você pode fazer uma busca detalhada nos dados de análise de cluster para exibir os registros em cada cluster. Clique duas vezes em um cluster na exibição **Detalhes** e exiba os dados no Visualizador de Dados. O Visualizador de Dados exibe um cluster de cada vez. Os dados do cluster incluem as opções de pontuação selecionadas, como pontuação de controle, pontuação de vínculo, identificador de controle ou identificador de vínculo.

### Registro da Transformação de Correspondência em Log

Quando você executa um mapeamento que usa uma transformação de Correspondência, o log da Developer tool controla o número de cálculos de comparação executados pelo mapeamento. Para exibir os dados de log, selecione a opção **Mostrar Log** no Visualizador de Dados.

O mapeamento atualiza o log a cada 100 mil cálculos.

## Exibindo Dados de Análise de Cluster de Correspondência

É possível exibir dados estatísticos sobre os clusters que a transformação pode criar. As estatísticas de cluster resumem o nível de duplicação de registros no conjunto de dados com base na configuração de mapeamento atual.

Para exibir os dados, clique com o botão direito do mouse na transformação de Correspondência na tela de mapeamento e selecione **Análise de Cluster de Correspondência**.

Antes de executar a análise, valide o mapeamento que contém a transformação.

A análise de cluster de correspondência exibe dados para as seguintes propriedades:

Propriedade	Descrição
Origem	O número de linhas de dados de entrada.
Última execução	A data e a hora da análise.
Número total de clusters descobertos	O número de clusters gerados pela análise de correspondência quando o mapeamento é executado.
Tamanho mínimo do cluster	O número de registros em um ou mais clusters que contém o menor número de registros. Se o tamanho mínimo do cluster for 1, o conjunto de dados contém pelo menos um registro exclusivo.
Tamanho máximo do cluster	O número de registros em um ou mais clusters que contém o maior número de registros. Se esse valor exceder significativamente o tamanho médio do cluster, talvez o maior cluster contenha duplicatas falsas.
Número de registros exclusivos	O número de registros no conjunto de dados que não correspondem a outros registros com uma pontuação que se enquadra no limite de correspondência.
Número de registros duplicados	O número de registros no conjunto de dados que correspondem a outros registros com uma pontuação que se enquadra no limite de correspondência.
Total de comparações	O número de operações de comparação realizadas pelo mapeamento.
Tamanho médio do cluster	O número médio de registros em um cluster.

## Exibindo Dados de Análise de Desempenho de Correspondência

Você pode exibir dados estatísticos sobre os grupos de registros lidos pelo mapeamento como dados de entrada.

Para exibir os dados, clique com o botão direito do mouse na transformação de Correspondência na tela de mapeamento e selecione **Análise de Desempenho de Correspondência**.

Antes de executar a análise, valide o mapeamento que contém a transformação.

A análise de desempenho de correspondência exibe dados para as seguintes propriedades:

Propriedade	Descrição
Origem	O número de linhas de dados de entrada.
Última execução	A data e a hora da análise.
Número total de grupos descobertos	O número de grupos definidos para o conjunto de dados, com base no valor de chave de grupo selecionado.
Taxa de transferência (registros por minuto)	Valor variável que estima a velocidade da análise de correspondência. Você define esse valor. Use esse valor para estimar o tempo necessário para executar a análise de correspondência.
Tempo estimado para corresponder registros	O tempo necessário para analisar todos os registros do conjunto de dados, com base na configuração da transformação de Correspondência.
Número total de pares gerados	O número de comparações que a transformação deve realizar, com base no número de linhas de dados de entrada e no número de grupos.
Tamanho mínimo do grupo	Valor variável que indica o número mínimo de registros que um grupo pode conter. Você define esse valor. Use esse valor para verificar se o mapeamento criará grupos de um tamanho utilizável. <b>Nota:</b> O valor mínimo de tamanho do grupo não determina o tamanho dos grupos criados quando o mapeamento é executado.
Número de grupos abaixo do limite mínimo	O número de grupos que contêm menos registros que o valor mínimo de tamanho do grupo. Se vários grupos estiverem abaixo do tamanho mínimo, talvez seja necessário editar a transformação e selecionar uma chave de grupo diferente.
Tamanho máximo do grupo	Valor variável que indica o número máximo de registros que um grupo pode conter. Você define esse valor para a análise de desempenho. Use esse valor para verificar se o mapeamento criará grupos de um tamanho utilizável. <b>Nota:</b> O valor não determina o tamanho dos grupos criados quando o mapeamento é executado.
Número de grupos acima do limite máximo	O número de grupos que contêm mais registros que o valor máximo de tamanho do grupo. Se vários grupos estiverem acima do tamanho máximo, talvez seja necessário editar a transformação e selecionar uma chave de grupo diferente.

## Desempenho de Correspondência na Análise de Identidade

Para melhorar o desempenho de mapeamento quando você realizar a análise de identidade em dois conjuntos de dados, configure a transformação de Correspondência para ler dados de índice de identidade de tabelas de banco de dados. Execute um mapeamento para criar as tabelas de índice do conjunto de

dados mestre. Execute o mapeamento novamente para comparar os dados de índice com outra fonte de dados.

Use as opções na exibição **Tipo de Correspondência** para identificar as tabelas de banco de dados que armazenam os dados de índice. Use as mesmas opções para selecionar as tabelas de índice quando você configurar a transformação para comparar os dados de índice com os dados de outra fonte.

Para gravar os dados de índice em tabelas de banco de dados, realize as seguintes tarefas:

1. Crie um mapeamento que leia uma fonte de dados de informações de identidade.
2. Configure uma transformação de Correspondência no mapeamento para gravar os dados de índice em um banco de dados.
3. Execute o mapeamento para gerar os dados de índice. Os dados de índice representam um armazenamento de dados que você pode reutilizar.

Para ler os dados de índice de tabelas de banco de dados, realize as seguintes tarefas:

1. Crie um mapeamento que leia outra fonte de dados de identidade.
2. Configure uma transformação de Correspondência no mapeamento para ler os dados de índice do banco de dados que você especificou anteriormente.

Quando a fonte de dados de mapeamento e os dados de índice compartilham uma estrutura em comum, você pode reutilizar o mapeamento que gerou os dados de índice.

3. Execute o mapeamento para comparar a fonte de dados com os dados de índice.

O mapeamento gera dados de índice para a fonte de dados. O mapeamento não precisa gerar dados de índice para o conjunto de dados maior. Portanto, ele é executado de maneira mais rápida do que um mapeamento de origem dupla que gera dados de índice para ambos os conjuntos de dados.

## TÓPICOS RELACIONADOS:

- [“Análise de Correspondência de Identidade e Dados de Indexação Persistente” na página 469](#)

## Criando um Armazenamento de Dados para Dados de Índice de Identidade

Configure um mapeamento que leia uma fonte de dados contendo informações de identidade. Use uma transformação de Correspondência para gravar os dados de índice em um banco de dados.

1. Crie um mapeamento e adicione a fonte de dados à tela de mapeamentos.
2. Adicione uma transformação de Correspondência à tela de mapeamentos.
3. Na fonte de dados, selecione as portas que contêm as informações de identidade.
  - Conecte as portas de informações de identidade com a transformação de Correspondência.
  - Conecte a porta que contém os valores do identificador de sequência da transformação de correspondência.
4. Na transformação de Correspondência, selecione a exibição **Tipo de Correspondência**.
5. Defina o tipo de correspondência como **Correspondência de Identidade com ID de Registro Persistente**.
6. Configure as opções a seguir para criar o armazenamento de dados de índice.
  - Defina o método de persistência como **Atualizar o banco de dados com novos IDs**.
  - Verifique o valor do processo de correspondência. O valor padrão é **Ativar**. Se você desativar o processo de correspondência, o mapeamento criará as tabelas do índice de identidade, mas não executará a análise de correspondência nos dados.

- No menu Conexão de BD, selecione um banco de dados para as tabelas de índice.
  - No menu Armazenamento Persistente, selecione **Criar Novo**. Insira um nome para o índice na caixa de diálogo **Criar Tabelas de Armazenamento**.
7. Realize qualquer outro etapa de configuração necessária para a transformação de Correspondência. Por exemplo, configure uma estratégia de transformação.
  8. Adicione um objeto de dados de destino ao mapeamento.
  9. Conecte as portas de saída da transformação de Correspondência com o objeto de dados de destino.
  10. Execute o mapeamento.

O mapeamento grava os dados de índice da fonte de dados nas tabelas de banco de dados que você especificar.

## Usando o Armazenamento de Dados de Índice na Análise de Origem Única

Configure um mapeamento que leia uma fonte de dados de informações de identidade. Use uma transformação de Correspondência para comparar a fonte de dados com o armazenamento de dados de índice do conjunto de dados mestre.

Antes de configurar o mapeamento, verifique se a fonte de dados contém uma coluna de valores de identificador de sequência.

Para poupar tempo, você pode copiar ou reutilizar o mapeamento que criou o armazenamento de dados de índice.

1. Abra o mapeamento que gerou o armazenamento de dados de índice.  
Como alternativa, abra uma cópia desse mapeamento.
2. Verifique a fonte de dados no mapeamento.  
Se necessário, substitua a fonte de dados por uma fonte que contenha os dados atuais.  
**Nota:** Se você excluir a fonte de dados, também excluirá as conexões de portas na transformação de Correspondência.
3. Identifique as portas de fonte de dados que contêm as informações de identidade.
  - Conecte as portas de informações de identidade com a transformação de Correspondência.
  - Conecte a porta que contém os valores do identificador de sequência da transformação de correspondência.  
As portas de entrada e a ordem das portas de entrada devem corresponder às portas de entrada na transformação que criou as tabelas de índice.
4. Na transformação de Correspondência, selecione a exibição **Tipo de Correspondência**.
5. Defina o tipo de correspondência como **Correspondência de Identidade com ID de Registro Persistente**.
6. Verifique os valores de preenchimento, nível de chave, tipo de chave e campo de chave.
7. Configure as opções para identificar o armazenamento de dados de índice:
  - Defina o método de persistência. Por exemplo, selecione **Não Atualizar o Banco de Dados** para preservar os dados atuais nas tabelas de índice.
  - Defina o processo de correspondência como **Ativar**.
  - No menu Conexão de BD, selecione o banco de dados que contém as tabelas de índice.
  - No menu Armazenamento de Persistência, navegue até as tabelas que contêm os dados de índice.
8. Configure as seguintes propriedades na exibição Saída de Correspondência:

- **Corresponder.** Identifica os registros a serem analisados quando a transformação lê dados de índice de tabelas do banco de dados.
  - **Saída.** Filtra os registros que a transformação grava como saída.
9. Realize qualquer outro etapa de configuração necessária para a transformação de Correspondência. Por exemplo, configure uma estratégia de transformação.
  10. Verifique o destino de dados no mapeamento. Se necessário, substitua os dados de destino por outro objeto de destino.
  11. Conecte as portas de saída da transformação de Correspondência com o objeto de dados de destino.
  12. Execute o mapeamento.

O mapeamento compara os registros de fonte de dados com o armazenamento de dados de índice. A transformação grava os dados de índice da fonte de dados no armazenamento de dados.

## Exibições de Transformação de Correspondência

A transformação de Correspondência organiza as opções que você pode configurar em uma série de exibições. Em uma transformação reutilizável, as exibições aparecem como guias na tela da Developer tool. Para abrir uma exibição, clique em uma guia. Em uma transformação não reutilizável, as exibições aparecem como uma lista na tela. Para abrir uma exibição, clique em um item na lista.

Ao configurar uma operação de análise de correspondência, você pode configurar as seguintes exibições:

### **Geral**

Use a exibição Geral para atualizar o nome e a descrição de uma transformação de Correspondência não reutilizável. A descrição é opcional.

### **Portas**

Use a exibição Portas para verificar as portas de entrada e as portas de saída em uma transformação de Correspondência não reutilizável.

### **Visão geral**

Use a exibição Visão Geral para atualizar o nome e a descrição de uma transformação reutilizável. A descrição é opcional. Você também pode usar a exibição Visão Geral para criar as portas de entrada e as portas de saída em uma transformação reutilizável.

### **Tipo de Correspondência**

Use a exibição Tipo de Correspondência para selecionar o tipo de análise de duplicatas realizado pela transformação. É possível selecionar uma análise de correspondência de campos ou uma análise de correspondência de identidade. Você pode especificar uma única fonte de dados ou duas fontes de dados.

### **Estratégias**

Use a exibição Estratégias para definir uma ou mais estratégias de análise dos dados de entrada. Selecione duas colunas de dados em cada estratégia e atribua um algoritmo de análise de correspondência a essas colunas.

### **Saída de Correspondência**

Use a exibição de saída de correspondência para especificar a estrutura e o formato dos dados de saída.

## Parâmetros

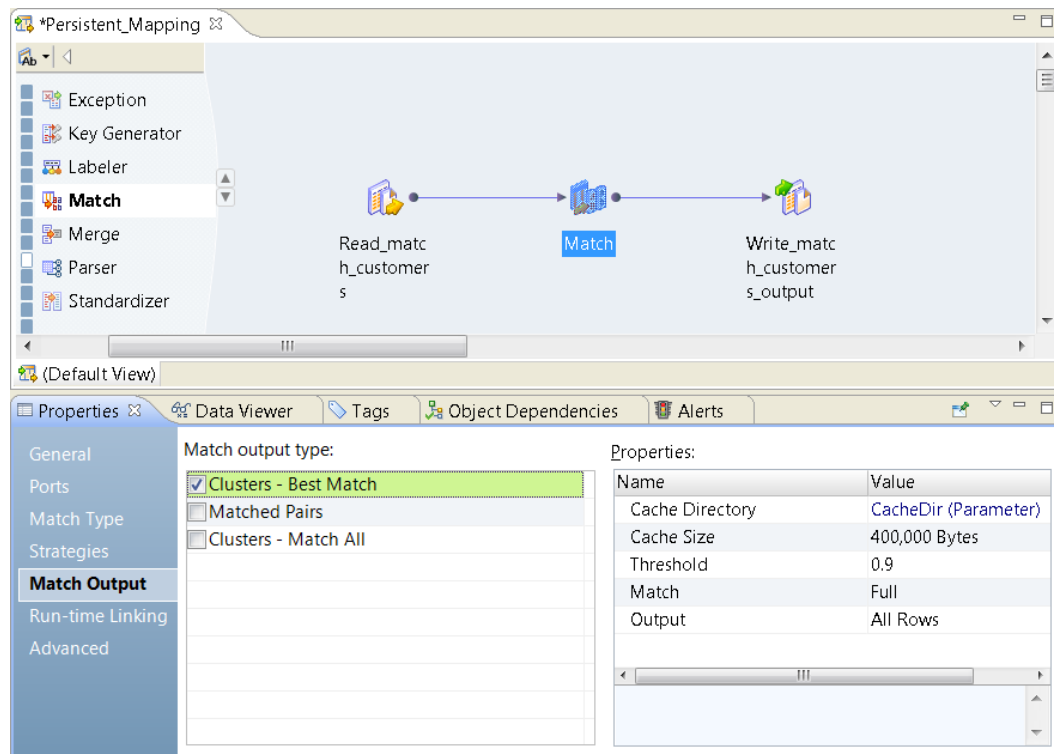
Use a exibição Parâmetros para definir qualquer parâmetro que o Serviço de Integração de Dados pode aplicar à transformação quando você executar um mapeamento que contém essa transformação.

## Avançado

Use a exibição Avançado para especificar as seguintes propriedades:

- O nível de detalhes em qualquer mensagem de log gravada pelo mapeamento para a transformação.
- O número de processos usados pelas operações de correspondência de identidade quando você executa um mapeamento que contém a transformação.
- Se a transformação transmite ou não dados de registros idênticos diretamente para as portas de saída. Você pode filtrar registros idênticos ao configurar a transformação para gravar uma saída clusterizada.

A seguinte imagem mostra as exibições em uma transformação de Correspondência que você configura para análise de identidade com um armazenamento de índice persistente:



# Portas de Transformação de Correspondência

A transformação de Correspondência inclui um conjunto de portas de entrada e portas de saída predefinidas que contêm metadados para as operações de análise de correspondência que você define. A transformação seleciona ou desmarca as portas quando você configura as opções de tipo de correspondência e saída de correspondência.

Ao configurar a transformação, analise as portas de metadados. Ao adicionar a transformação a um mapeamento, certifique-se de conectar as portas de metadados às portas corretas nos objetos de mapeamento upstream e downstream.



## Portas de Entrada da Transformação de Correspondência

As portas de entrada predefinidas contêm os metadados que a transformação precisa usar para a análise de correspondência.

Após criar uma transformação de Correspondência, você pode configurar as seguintes portas de entrada:

### **SequenceId**

Identificador exclusivo para cada registro no conjunto de dados de origem do mapeamento. Cada registro em um conjunto de dados de entrada deve incluir um identificador de sequência exclusivo. Se um conjunto de dados contiver identificadores de sequência duplicados, a transformação de Correspondência não poderá identificar registros duplicados corretamente. Use a transformação de Gerador de Chaves para criar identificadores exclusivos caso não exista nenhum nos dados.

Quando você cria um armazenamento de dados de índice para dados de identidade, a transformação de Correspondência adiciona o identificador de sequência de cada registro ao armazenamento de dados. Quando você configurar a transformação para comparar uma fonte de dados com o armazenamento de dados de índice, a transformação pode encontrar um identificador de sequência comum em ambos os conjuntos de dados. A transformação poderá analisar os identificadores de sequência se eles forem exclusivos nos respectivos conjuntos de dados.

### **GroupKey**

Valor de chave que identifica o grupo ao qual o registro pertence.

**Nota:** Para melhorar a velocidade de mapeamento, configure a porta de entrada GroupKey e a porta de saída que se conecta a ela com o mesmo valor de precisão.

## Portas de Saída da Transformação de Correspondência

As portas de saída predefinidas contêm metadados sobre a análise que é realizada pela transformação.

Após criar uma transformação de Correspondência, você pode configurar as seguintes portas de saída:

### **GroupKey**

Valor de chave que identifica o grupo ao qual o registro pertence.

Transformações downstream, como a transformação de Associação, podem ler o valor de chave de grupo.

### **ClusterId**

O identificador do cluster ao qual o registro pertence. Usado na saída de cluster.

### **ClusterSize**

O número de registros no cluster aos quais um registro pertence. Quando um cluster contém um registro exclusivo, o tamanho do cluster é 1. Usado na saída de cluster.

### **RowId e RowId1**

Um identificador de linha exclusivo para o registro. A transformação de Correspondência usa o identificador de linha para identificar a linha durante as operações de análise de correspondência. O identificador pode não corresponder ao número da linha nos dados de entrada.

### **DriverId**

O identificador de linha do registro de controle em um cluster. Usado na saída de cluster. O registro de controle é o registro no cluster com o valor mais alto na porta de entrada SequenceID.

**DriverScore**

A transformação atribui uma pontuação de controle na saída de pares correspondentes e na saída de cluster. Em um par correspondente, a pontuação de controle é a pontuação de correspondência entre o par de registros. Em um cluster, a pontuação de controle é a pontuação de correspondência entre o registro atual e o registro de controle no cluster.

**LinkId**

O identificador de linha do registro que foi correspondido com o registro atual e o vinculou ao cluster. Usado na saída de cluster.

**LinkScore**

A pontuação de correspondência entre dois registros que resulta na criação de um cluster ou na adição de um registro a um cluster. A porta LinkID identifica o registro com o qual o registro atual compartilha a pontuação de vínculo. Usado na saída de cluster.

**PersistenceStatus**

Um código de oito caracteres que representa os resultados da análise de correspondência em um registro de entrada. Usado na análise de identidade de origem única quando a transformação compara os dados de origem com um armazenamento de dados de índice.

A transformação preenche os três primeiros caracteres do código. A transformação pode retornar diferentes caracteres em cada posição. A transformação retorna 0 para as posições de quatro a oito.

Quando você configura a transformação para gerar saída uma saída de pares correspondentes, a transformação cria uma porta PersistenceStatus e uma porta PersistenceStatus1.

**PersistenceStatusDesc**

Uma descrição de texto dos valores de códigos de status de persistência. Usado na análise de identidade de origem única quando a transformação compara os dados de origem com um armazenamento de dados de índice.

Quando você configura a transformação para gerar saída uma saída de pares correspondentes, a transformação cria uma porta PersistenceStatusDesc e uma porta PersistenceStatusDesc1.

## Códigos de Status de Persistência e Descrições de Status de Persistência

Os códigos de status de persistência e as descrições de status de persistência descrevem o relacionamento entre os diferentes tipos de dados de índice que a transformação de Correspondência analisa. A transformação gera os códigos e as descrições de status quando você a configura para ler um armazenamento de dados da identidade persistentes.

A transformação grava o código de status de persistência na porta PersistenceStatus. O código contém oito caracteres. A transformação preenche as três primeiras posições na cadeia com os valores de código. A transformação retorna 0 para as posições de quatro a oito.

A transformação grava a descrição de status de persistência na porta PersistenceStatusDesc. A descrição contém três cadeias de texto separadas por vírgula que descrevem os valores nas três primeiras posições no código de status de persistência.

A transformação usa os valores do identificador de sequência dos registros de dados de origem para comparar os dados de índice dos dois conjuntos de dados.

A seguinte tabela descreve os tipos de informações que a transformação grava em cada posição na descrição e no código de status:

Posição	Descrição
1	Identifica o conjunto de dados que contém o registro.
2	Indica o status de duplicado do registro. A transformação procura identificadores de sequência comuns entre os dados de entrada da transformação e o armazenamento de dados de índice.
3	Descreve as ações que a transformação realiza nos dados.
4-8	O código de status contém 0 em cada posição. A descrição do status não contém texto para a posição.

## Valores do Código de Status e Valores de Descrição de Status

Os códigos e as descrições de status de persistência descrevem o relacionamento entre os registros de entrada da transformação e os registros que o armazenamento de dados representa. A transformação usa os valores de identificador de sequência para identificar os registros e determinar o relacionamento entre os registros nos conjuntos de dados.

Os códigos e as descrições de status de persistência têm uma estrutura comum. Os códigos e as descrições de status contêm as mesmas informações em cada posição na cadeia de dados de saída.

### Status do Conjunto de Dados

O primeiro valor do código e da descrição de status identifica o conjunto de dados que contém o registro.

A seguinte tabela descreve os códigos e as descrições de status que a transformação pode retornar na primeira posição:

Código de Status	Descrição do Status
S	Armazenamento. O registro atual foi originado no armazenamento de dados de índice.
I	Entrada. O registro atual foi originado nos dados de entrada da transformação.

### Status do Registro Duplicado

O segundo valor no código e na descrição de status descreve o relacionamento entre os dados de índice da transformação e o armazenamento de dados persistente.

A seguinte tabela descreve os códigos e as descrições de status que a transformação pode retornar na segunda posição:

Código de Status	Descrição do Status
A	Ausente. O armazenamento de dados de índice não contém dados do registro atual.
E	Existe. O registro atual está presente no armazenamento de dados de índice e nos de dados de entrada da transformação.
I	Inválido(a). A transformação não consegue analisar o registro atual. Por exemplo, a transformação não pode gerar dados de índice do registro porque o campo de chave na guia Tipo de Correspondência não é compatível com os dados de registro.
N	Novo. O registro está presente na fonte de dados.
0	[Dash] O registro está presente no armazenamento de dados de índice.

### Status do Armazenamento de Dados

O terceiro valor no código de status e na descrição do status descreve qualquer ação que a transformação realiza nas tabelas de dados de índice.

A seguinte tabela descreve os códigos e as descrições de status que a transformação pode retornar na terceira posição:

Código de Status	Descrição do Status
A	Adicionados. A transformação adiciona os dados de índice do registro de entrada atual ao armazenamento de dados persistentes. Os dados de entrada da transformação e os dados de índice persistente têm diferentes identificadores de sequência.
I	Ignorado. A transformação não adiciona dados de índice do registro de entrada atual ao armazenamento de dados persistentes.
N	A transformação retorna uma das seguintes descrições: <ul style="list-style-type: none"> <li>- Sem alteração. O registro é originado no armazenamento de dados persistentes, e a transformação não realiza nenhuma ação.</li> <li>- Não é adicionado. A transformação não atualiza o armazenamento de dados persistentes com os dados do registro de entrada atual devido à política de correspondência que você definiu.</li> </ul>

Código de Status	Descrição do Status
R	Removido. A transformação remove os dados de índice do registro do armazenamento de dados de índice.
U	Atualizado. A transformação atualiza as linhas no armazenamento de dados persistentes com os dados de índice do registro de entrada de transformação. Os dados de entrada da transformação e os dados de índice persistente têm identificadores de sequência comuns.

### Exemplo de Descrição de Status de Persistência

O código de status de persistência INA00000 tem a seguinte descrição de status de persistência:

*Entrada, Novos, Adicionados*

O código e a descrição de status contêm as seguintes informações sobre o registro:

- O registro foi originado nos dados de entrada de transformação.
- O armazenamento de dados persistentes não contém uma cópia do registro.
- A transformação adiciona os dados de índice do registro ao armazenamento de dados persistentes.

## Portas de Saída e Seleção de Saída de Correspondência

As opções de saída de correspondência selecionadas determinam as portas de saída na transformação. Por exemplo, a transformação cria uma porta ClusterId e uma porta ClusterSize quando você seleciona um tipo de saída clusterizado.

Selecione o tipo de saída de transformação que você precisa e examine as portas na transformação.

Se você atualizar o tipo de saída de correspondência, verifique a configuração da porta de saída na transformação depois de fazer isso. Se você usar a transformação em um mapeamento, talvez seja necessário reconectar as portas de saída a objetos downstream no mapeamento.

## Mapplets de Correspondência

Um mapplet de correspondência é um tipo de mapplet que você pode criar e inserir em transformações de Correspondência.

Você cria mapplets de correspondência salvando a configuração de uma transformação de Correspondência como um mapplet de correspondência. Ao criar um mapplet de correspondência, você converte as configurações de transformação de Correspondência em transformações de Comparação e de Média Ponderada.

Depois de criar um mapplet de correspondência, você poderá adicionar transformações para personalizar a análise de correspondência. Por exemplo, você pode adicionar uma transformação de Expressão para avaliar as pontuações de vínculo de duas estratégias e escolher a pontuação mais alta.

Diferentemente das transformações de Correspondência, os mapplets de correspondência são passivos, o que significa que você pode usá-los como regras dentro da ferramenta Analyst. Use mapplets de

correspondência na ferramenta Analyst para corresponder registros como parte de processos de criação de perfil de dados.

A transformação de Correspondência pode ler apenas mapplets de correspondência criados dentro de uma transformação de Correspondência.

## Criando um Mapplet de Correspondência

Crie um mapplet de correspondência para definir uma operação de análise de correspondência que usa várias transformações.

1. Abra uma transformação de Correspondência no editor e selecione a exibição **Estratégias**.
2. Selecione **Usar Regra de Correspondência**.
3. No campo **Nome**, selecione **Criar novo**.  
A janela **Novo Mapplet** é aberta.
4. Na janela **Novo Mapplet**, insira um nome para o mapplet e escolha um local para salvar o mapplet.
5. Opcionalmente, selecione **Reutilizar Estratégias da transformação de Correspondência** para copiar as entradas, as estratégias e os pesos da transformação de Correspondência atual para o mapplet de correspondência.

**Nota:** A Informatica recomenda o uso desta configuração para criar mapplets de correspondência rapidamente que replicam a funcionalidade de correspondência atualmente definida em transformações de Correspondência.

6. Clique em **Concluir**.  
O mapplet de correspondência é aberto no editor.
7. Opcionalmente, crie operações de correspondência adicionando e configurando transformações de Comparação e de Média Ponderada no mapplet de correspondência.
8. Clique em **Arquivo > Salvar** para salvar o mapplet.
9. Feche o mapplet e selecione o editor que contém a transformação de Correspondência. Verifique se o mapplet criado aparece no campo **Nome**.
10. Opcionalmente, configure os campos de correspondência no mapplet clicando no botão **Campos de Correspondência**.  
A janela **Configurar Regra de Correspondência** é aberta.
11. Clique duas vezes nos campos das colunas **Campos de Entrada** e **Entradas Disponíveis** para atribuir portas de entrada a corresponder entradas.
12. Clique em **Arquivo > Salvar** para salvar a transformação.

## Usando um Mapplet de Correspondência

Você pode selecionar e configurar um mapplet de correspondência definido anteriormente na transformação de Correspondência.

1. Abra uma transformação de Correspondência no editor e selecione a exibição **Estratégias**.
2. Selecione **Usar Regra de Correspondência**.
3. No campo **Nome**, selecione **Usar Existente**.  
A janela **Configurar Regra de Correspondência** é aberta.
4. Clique em **Procurar** para localizar um mapplet de correspondência no repositório.

**Importante:** Só é possível selecionar mapplets criados pela transformação de Correspondência.

- A janela **Selecionar Maplet de Correspondência** é aberta.
5. Selecione um maplet de correspondência e clique em **OK**.
  6. Clique duas vezes nos campos das colunas **Campos de Entrada** e **Entradas Disponíveis** para atribuir portas de entrada a corresponder entradas.
  7. Clique em **OK**.
- A janela **Configurar Regra de Correspondência** é fechada.
8. Clique em **Arquivo > Salvar** para salvar a transformação de Correspondência.

## Configurando uma Operação de Análise de Correspondência

Para configurar uma operação de correspondência, conecte os dados de origem à transformação de Correspondência e edite as propriedades nas exibições da transformação.

1. Crie uma transformação de Correspondência e conecte os dados de origem à transformação.
2. Selecione a exibição **Tipo de Correspondência** e escolha um tipo de correspondência.
3. Configure as propriedades para o tipo de operação de correspondência selecionado.  
Se você selecionou um tipo de correspondência de origem dupla, configure a propriedade **Conjunto de Dados Mestre**.
4. Selecione a exibição **Estratégias** e escolha **Definir estratégias de correspondência**.
5. Clique em **Novo**.  
O assistente de **Nova Estratégia de Correspondência** é aberto.
6. Escolha uma estratégia de correspondência e clique em **Avançar**.
7. Opcionalmente, edite as configurações de peso e de correspondência nula. Clique em **Avançar**.
8. Clique duas vezes nas células da coluna Disponível para selecionar as portas de entrada para análise.  
Clique em **Avançar** para configurar outra estratégia ou clique em **Concluir** para sair do assistente.  
**Nota:** Para editar a configuração da estratégia, clique na seta na célula dessa estratégia, na exibição **Estratégias**.
9. Selecione a exibição **Saída de Correspondência**.  
Escolha um tipo de saída de correspondência e configure as propriedades.

**Nota:** Você também pode configurar estratégias de correspondência selecionando ou criando um maplet de correspondência na exibição **Estratégias**. Um maplet de correspondência é um tipo de maplet que você pode inserir em uma transformação de Correspondência.

# Transformação de Correspondência em um ambiente não nativo

O processamento da transformação de Correspondência no ambiente não nativo depende do mecanismo que executa a transformação.

Considere o suporte para os seguintes mecanismos de tempo de execução não nativos:

- Mecanismo Blaze. Suportado com restrições.
- Mecanismo Spark. Com suporte com restrições em mapeamentos em lote. Sem suporte em mapeamentos de fluxo.
- Mecanismo Databricks Spark. Sem suporte.

## Transformação de Correspondência no mecanismo Blaze

A validação de mapeamento falha quando a transformação de Correspondência é configurada para gravar dados de índice de identidade em tabelas de banco de dados.

Uma transformação de Correspondência gera valores de ID de cluster de maneira diferente em ambientes nativos e não nativos. Em um ambiente não nativo, a transformação anexa um valor de ID de grupo à ID do cluster.

## Transformação de Correspondência no mecanismo Spark

A validação de mapeamento falha quando a transformação de Correspondência é configurada para gravar dados de índice de identidade em tabelas de banco de dados.

Uma transformação de Correspondência gera valores de ID de cluster de maneira diferente em ambientes nativos e não nativos. Em um ambiente não nativo, a transformação anexa um valor de ID de grupo à ID do cluster.



## CAPÍTULO 29

# Transformações de Correspondência na Análise de Campos

Este capítulo inclui os seguintes tópicos:

- [Análise de Correspondência de Campos, 485](#)
- [Fluxo de Processos para Análise de Correspondência de Campos, 486](#)
- [Opções de Tipo de Correspondência de Campos, 486](#)
- [Estratégias de Correspondência de Campos, 487](#)
- [Opções de Saída de Correspondência de Campos, 490](#)
- [Propriedades Avançadas de Correspondência de Campos, 492](#)
- [Exemplo de Análise de Correspondência de Campos, 493](#)

## Análise de Correspondência de Campos

Realize a análise de correspondência de campos para localizar registros semelhantes ou duplicados em um conjunto de dados ou entre dois conjuntos de dados.

Ao configurar a transformação de Correspondência para análise de correspondência de campos, você define as opções nas seguintes exibições:

- Tipo de Correspondência
- Estratégias
- Saída de Correspondência

Opcionalmente, defina as opções nas exibições Parâmetros e Avançado.

# Fluxo de Processos para Análise de Correspondência de Campos

O seguinte fluxo de processos resume as etapas realizadas para configurar uma transformação de Correspondência para a análise de correspondência de campos. É possível definir um processo que utiliza apenas a transformação de Correspondência ou que utiliza a transformação de Correspondência junto com outras transformações.

**Nota:** Ao adicionar uma transformação de Correspondência a um mapeamento na análise de correspondência de campos, adicione uma transformação de Gerador de Chaves upstream a esse mapeamento.

Para preparar os dados para a transformação de Correspondência, realize as seguintes etapas:

1. Organize os registros de dados de origem em grupos.  
Use uma transformação de Gerador de Chaves para atribuir um valor de chave de grupo a cada registro. As atribuições de grupo reduzem o número de cálculos que a transformação de Correspondência deve realizar.
2. Verifique se os registros de fonte de dados contêm valores exclusivos de identificador de sequência. É possível usar uma transformação de Gerador de Chaves para criar esses valores.

Realize as seguintes etapas na transformação de Correspondência:

1. Especifique a análise de campos como o tipo de correspondência e defina o número de fontes de dados.  
Se você configurar a transformação para analisar dois conjuntos de dados, selecione um conjunto de dados mestre.  
Use a exibição **Tipo de Correspondência** para definir o tipo e o número de fontes de dados.
2. Defina uma estratégia de análise de correspondência. Selecione um algoritmo e atribua um par de colunas a esse algoritmo.  
Use a exibição **Estratégias** para definir a estratégia.
3. Especifique o método usado pela transformação para gerar resultados de análise de correspondência.
4. Defina o valor de limite de correspondência. O limite de correspondência é a pontuação mínima que pode identificar dois registros como duplicatas um do outro.  
Use a exibição **Saída de Correspondência** para selecionar o método de saída e o limite de correspondência.

**Nota:** É possível definir o limite de correspondência em uma transformação de Correspondência ou em uma transformação de Média Ponderada. Use a transformação de Média Ponderada se criar um mapplet de correspondência.

## Opções de Tipo de Correspondência de Campos

A exibição Tipo de Correspondência contém uma única opção que se aplica à análise de correspondência de origem dupla. Essa opção identifica o conjunto de dados mestre. A exibição Tipo de Correspondência não contém uma opção para a análise de correspondência de origem única.

Ao analisar dois conjuntos de dados, você deve selecionar um deles como um conjunto de dados mestre. Se nenhum conjunto de dados representar um conjunto de dados mestre no projeto ou na organização, selecione o conjunto de dados maior como o conjunto de dados mestre.

Use a opção **Conjunto de Dados Mestre** para especificar o conjunto de dados mestre.

## Estratégias de Correspondência de Campos

A exibição Estratégias lista as estratégias que você define para os dados de entrada.

Essas estratégias determinam como a transformação mede as semelhanças e as diferenças entre os registros de fonte de dados.

### Algoritmos de Correspondência de Campos

A transformação de Correspondência inclui algoritmos que comparam valores de dados entre duas colunas. Cada algoritmo calcula o grau de diferença entre valores de dados de maneira diferente.

Selecione um algoritmo que pode medir a diferença de tipos de dados que você espera encontrar nas colunas selecionadas.

#### Bigram

Use o algoritmo Bigram para comparar strings de texto longo, como endereços postais inseridos em um único campo.

O algoritmo Bigram calcula uma pontuação de correspondência para duas strings de dados com base na ocorrência de caracteres consecutivos em ambas as strings. O algoritmo procura pares de caracteres consecutivos que são comuns para ambas as strings. Ele divide o número de pares correspondentes nessas strings pelo número total de pares de caracteres.

#### Exemplo de Bigram

Considere as seguintes strings:

- larder
- lerder

Essas strings geram os seguintes grupos de Bigram:

```
l a, a r, r d, d e, e r  
l e, e r, r d, d e, e r
```

Observe que a segunda ocorrência da string "e r" dentro da string "lerder" não é correspondente, pois não há uma segunda ocorrência correspondente de "e r" na string "larder".

Para calcular a pontuação de correspondência do Bigram, a transformação divide o número de pares correspondentes (6) pelo número total de pares em ambas as strings (10). Nesse exemplo, as strings são 60% semelhantes, e a pontuação de correspondência é 0,60.

#### Distância de Hamming

Use o algoritmo Distância de Hamming quando a posição dos caracteres de dados for um fator crítico, por exemplo, em campos numéricos ou de código como números de telefone, códigos postais ou códigos de produto.

O algoritmo Distância de Hamming calcula uma pontuação de correspondência para duas strings de dados computando o número de posições nas quais os caracteres diferem entre as strings de dados. Para strings

de comprimento diferente, cada caractere adicional na string mais longa é contado como uma diferença entre as strings.

### Exemplo de Distância de Hamming

Considere as seguintes strings:

- Morlow
- Marlowes

Os caracteres realçados indicam as posições que o algoritmo de Hamming identifica como diferentes.

Para calcular a pontuação de correspondência de Hamming, a transformação divide o número de caracteres correspondentes (5) pelo comprimento da string mais longa (8). Nesse exemplo, as strings são 62,5% semelhantes, e a pontuação de correspondência é 0,625.

### Distância de Edição

Use o algoritmo Distância de Edição para comparar palavras ou strings de texto curto, como nomes.

O algoritmo Distância de Edição calcula o "custo" mínimo de transformar uma string em outra por meio da inserção, exclusão ou substituição de caracteres.

### Exemplo de Distância de Edição

Considere as seguintes strings:

- Levenston
- Levenshtein

Os caracteres realçados indicam as operações necessárias para transformar uma string em outra.

O algoritmo Distância de Edição divide o número de caracteres inalterados (8) pelo comprimento da string mais longa (11). Nesse exemplo, as strings são 72,7% semelhantes, e a pontuação de correspondência é 0,727.

### Distância de Jaro

Use o algoritmo Distância de Jaro para comparar duas strings quando a similaridade dos caracteres iniciais nas strings for uma prioridade.

A pontuação de correspondência Distância de Jaro reflete o grau de semelhança entre os quatro primeiros caracteres de ambas as strings e o número de transposições de caracteres identificadas. A transformação pesa a importância da correspondência entre os quatro primeiros caracteres usando o valor que você insere na propriedade Penalidade.

### Propriedades da Distância de Jaro

Ao configurar um algoritmo Distância de Jaro, é possível definir as seguintes propriedades:

#### Penalidade

Determina a penalidade da pontuação de correspondência se os quatro primeiros caracteres em duas strings comparadas não forem idênticos. A transformação subtrai o valor total da penalidade para uma incompatibilidade de primeiro caractere e subtrai frações da penalidade com base na posição dos outros caracteres incompatíveis. O valor da penalidade padrão é 0,20.

#### Distinção Entre Maiúsculas e Minúsculas

Determina se o algoritmo Distância de Jaro faz distinção entre maiúsculas e minúsculas ao comparar caracteres.

## Exemplo de Distância de Jaro

Considere as seguintes strings:

- 391859
- 813995

Se você usar o valor de Penalidade padrão de 0,20 para analisar essas strings, o algoritmo Distância de Jaro retornará uma pontuação de correspondência de 0,513. Essa pontuação de correspondência indica que as strings são 51,3% semelhantes.

## Distância de Hamming Invertida

Use o algoritmo Distância de Hamming Invertido para calcular a porcentagem de posições de caractere que diferem entre as duas strings quando se faz a leitura da direita para a esquerda.

O algoritmo Distância de Hamming calcula uma pontuação de correspondência para duas strings de dados computando o número de posições nas quais os caracteres diferem entre as strings de dados. Para strings de comprimento diferente, o algoritmo conta cada caractere adicional na string mais longa como uma diferença entre as strings.

## Exemplo de Distância de Hamming Invertido

Considere as seguintes strings, que usam alinhamento da direita para a esquerda para imitar o algoritmo Hamming Invertido:

- 1-999-9999
- **011-01**-999-9991

Os caracteres realçados indicam as posições que o algoritmo Distância de Hamming Invertido identifica como diferentes.

Para calcular a pontuação de correspondência do Hamming Invertido, a transformação divide o número de caracteres correspondentes (9) pelo comprimento da string mais longa (15). Nesse exemplo, a pontuação de correspondência é 0,6, indicando que as strings são 60% semelhantes.

# Propriedades da Estratégia de Correspondência de Campos

Abra o assistente de **Estratégia** na exibição **Estratégias** e configure as propriedades para cada estratégia de correspondência de campos.

Ao configurar uma estratégia de correspondência de campos, você pode definir as seguintes propriedades:

### Nome

Identifica a estratégia pelo nome.

### Peso

Determina a prioridade relativa atribuída à pontuação de correspondência quando a pontuação geral para o registro é calculada. O padrão é 0.5.

### Nulo de Campo Único

Define a pontuação de correspondência que o algoritmo aplica a um par de valores de dados quando um valor é nulo. O padrão é 0.5.

### Ambos os Campos Nulos

Define a pontuação de correspondência que o algoritmo aplica a um par de valores de dados quando ambos os valores são nulos. O padrão é 0.5.

**Nota:** Um algoritmo de correspondência não calcular uma pontuação de correspondência quando um ou ambos os valores de colunas com correspondência são nulos. O algoritmo aplica as pontuações definidas nas propriedades de correspondência nula. Não é possível limpar as propriedades de correspondência nula.

## Opções de Saída de Correspondência de Campos

Configure as opções de **Saída de Correspondência** para definir o formato de saída da análise de correspondência de campos.

Você configura as opções na área **Tipo de Saída de Correspondência** e na área **Propriedades**.

### Tipos de Saída de Correspondência

A exibição Saída de Correspondência inclui opções que especificam o formato dos dados de saída. É possível configurar a transformação para gravar registros em clusters ou em pares correspondentes.

Selecione um dos seguintes tipos de saída de correspondência:

#### Melhor Correspondência

Grava cada registro no conjunto de dados mestre com o registro que representa a melhor correspondência no segundo conjunto de dados. A operação de correspondência seleciona o registro no segundo conjunto de dados com a pontuação de correspondência mais alta do registro principal. Se dois ou mais registros retornarem a pontuação mais alta, a operação de correspondência selecionará o primeiro registro do segundo conjunto de dados. A opção Melhor Correspondência grava cada par de registros em uma única linha.

Você pode selecionar **Melhor Correspondência** ao configurar a transformação para a análise de origem dupla.

#### Clusters

Grava clusters que contêm conjuntos dos registros que se vinculam uns aos outros com pontuações de correspondência que atendam ao limite de correspondência. Cada registro deve corresponder a pelo menos um outro registro no cluster com uma pontuação que atenda ao limite.

Você pode selecionar **Clusters** ao configurar a transformação para análises de origem única e de origem dupla.

#### Pares Correspondentes

Grava todos os pares de registros que são correspondidos entre si com uma pontuação que atende ao limite de correspondência. A transformação grava cada par em uma única linha e adiciona a pontuação de correspondência de cada par a cada linha. Se um registro corresponder a mais de um outro registro, a transformação gravará uma linha para cada par de registros.

Você pode selecionar **Pares Correspondentes** ao configurar a transformação para análises de origem única e de origem dupla.

## Propriedades da Saída de Correspondência

A exibição Saída de Correspondência inclui propriedades que especificam o comportamento da memória cache, o limite de pontuação de correspondência e as pontuações de correspondência exibidas na saída de transformação.

Você também pode usar as propriedades da saída de correspondência para especificar como a transformação adiciona valores de pontuação de correspondência aos registros de saída.

Depois de selecionar um tipo de saída de correspondência, configure as seguintes propriedades:

### Diretório de Cache

Especifica o diretório no qual o Serviço de Integração de Dados grava os dados temporários durante a análise de correspondência de campos. O Serviço de Integração de Dados grava os arquivos temporários no diretório quando o volume de dados que a análise de correspondência gera é maior do que a memória disponível do sistema. O Serviço de Integração de Dados exclui os arquivos temporários depois que o mapeamento é executado.

Você pode inserir um caminho de diretório na propriedade ou usar um parâmetro para identificar o diretório. Especifique um caminho local na máquina do Serviço de Integração de Dados. O Serviço de Integração de Dados deve ser capaz de gravar no diretório. O valor padrão é o parâmetro do sistema CacheDir.

### Tamanho do Cache

Determina a quantidade de memória do sistema que o Serviço de Integração de Dados atribui à análise de correspondência de campos. O valor padrão é 400.000 bytes.

Antes de classificar os dados, o Serviço de Integração de Dados aloca a quantidade de memória especificada por você. Se a análise de correspondência gerar uma quantidade maior de dados, o Serviço de Integração de Dados gravará o excesso de dados no diretório de cache. Se a análise de correspondência exigir mais memória do que a memória do sistema e o armazenamento de arquivos podem fornecer, o mapeamento falhará.

**Nota:** Se você inserir um valor igual ou maior que 65536, a transformação lerá o valor em bytes. Se você inserir um valor menor, a transformação lerá o valor em megabytes.

### Limite

Define a pontuação de correspondência mínima que identifica dois registros como possíveis duplicatas um do outro.

É possível atribuir um parâmetro ao valor de limite. Defina um valor decimal no intervalo de 0 a 1.

### Método de Pontuação

Determina os valores de pontuação de correspondência que aparecem na saída da transformação. Selecione um método de pontuação para saídas de cluster.

A seguinte tabela descreve as opções de método de pontuação:

Opção de Método de Pontuação	Descrição
Ambos	Adiciona a pontuação de link e de driver a cada registro no cluster.
Pontuação de Link	Adiciona a pontuação de link a cada registro no cluster. Opção padrão.

Opção de Método de Pontuação	Descrição
Pontuação de Driver	Adiciona a pontuação de driver a cada registro no cluster.
Nenhuma	Não adiciona uma pontuação de correspondência a qualquer registro no cluster.

**Nota:** Se você adicionar a pontuação de driver aos registros, aumente o tempo de execução do mapeamento. O mapeamento aguarda a conclusão de todos os clusters antes de adicionar os valores de pontuação de driver aos registros.

## Propriedades Avançadas de Correspondência de Campos

A transformação inclui propriedades avançadas que determinam o número de instâncias de execução, como a transformação analisa linhas idênticas e o nível de rastreamento de dados de log.

Você pode configurar as seguintes propriedades avançadas:

### Instâncias de Execução

Determina o número de segmentos que a transformação usa em tempo de execução.

A transformação de Correspondência usa uma única instância de execução na análise de correspondência de campos. Você pode editar o número de instâncias de execução ao configurar a transformação para a análise de correspondência de identidade.

### Correspondência Exata do Filtro

Determina se a transformação aplica o algoritmo de comparação em uma estratégia de correspondência a pares de registros idênticos nos dados de entrada.

Quando a transformação encontra um par de registros idênticos, o algoritmo não precisa analisar os níveis de similaridade entre os registros. A transformação pode transmitir os registros diretamente para a fase de saída sem análise adicional. Para configurar a transformação para transmitir os registros idênticos diretamente para a fase de saída, selecione Filtrar Correspondência Exata. Quando os dados de entrada contêm muitas linhas idênticas, o algoritmo de comparação realiza menos cálculos e o mapeamento é executado mais rapidamente.

Selecione a opção quando os dados de entrada contiverem muitas linhas idênticas. Não selecione a opção se os dados de entrada não contiverem muitas linhas idênticas porque a transformação poderá ser executada mais lentamente.

**Nota:** A saída de transformação contém os mesmos dados de registro quando você seleciona ou desmarca a opção. A transformação pode atribuir diferentes pontuações de link e de driver para os registros de saída quando você seleciona e desmarca a opção.

### Nível de Rastreamento

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.



# Exemplo de Análise de Correspondência de Campos

Você é um administrador de dados em uma agência bancária. Você recebeu um conjunto de registros de contas de clientes que abriram contas bancárias nos últimos sete dias. Você deseja verificar se o conjunto de dados não contém registros duplicados. Você cria um mapeamento para pesquisar os registros em busca de dados duplicados.

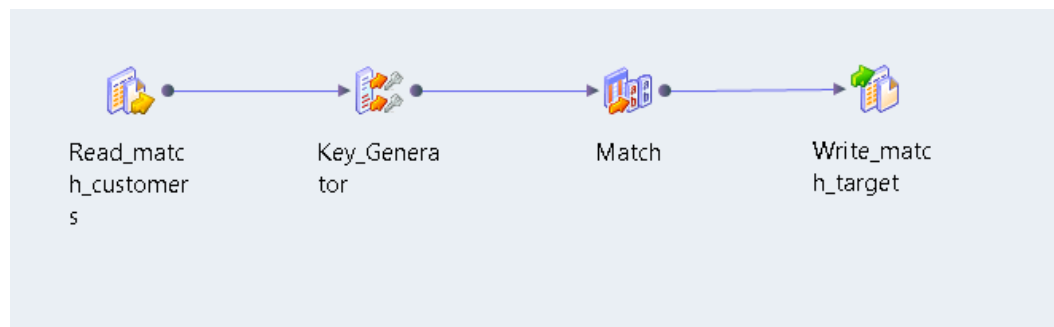
## Criar o Mapeamento

Crie um mapeamento que procure dados duplicados em vários campos.

O mapeamento realiza as seguintes tarefas:

- Lê uma fonte de dados.
- Adiciona um valor de chave de grupo e um valor de identificador de sequência aos registros de origem.
- Analisa os dados de campo nos registros.
- Grava os resultados em um destino de dados.

A seguinte imagem mostra o mapeamento na ferramenta Developer:



O mapeamento criado contém os seguintes objetos:

Nome do objeto	Descrição
Read_Match_Customers	Fonte de dados. Contém os detalhes dos titulares das contas, incluindo nomes, endereços e números de conta.
Key_Generator	Transformação de Gerador de Chaves. Adiciona valores de chave de grupo e valores de identificador de sequência aos dados de origem.
Correspondência	Transformação de Correspondência. Analisa os níveis de duplicação nos dados das contas.
Write_match_target	Destino de dados. Contém os resultados da análise de campos.

## Amostra de Dados de Entrada

O conjunto de dados contém o número da conta, o nome, o endereço e o empregador de cada cliente. Você cria uma fonte de dados a partir do conjunto de dados no repositório do Modelo. Você adiciona a fonte de dados ao mapeamento.

O seguinte fragmento de dados apresenta uma amostra dos dados de contas dos clientes:

CustomerID	Sobrenome	Cidade	Estado	CEP
15954467	JONES	SCARSDALE	NY	10583
10110907	JONES	MINNEAPOLIS	MN	55437
19131127	JONES	INDIANAPOLIS	EM	46240
10112097	JONES	HOUSTON	TX	77036
19133807	JONES	PLANTATION	FL	33324
10112447	JONES	SCARSDALE	NY	10583
15952487	JONES	HOUSTON	TX	77002
10112027	JONES	OAKLAND	CA	94623

## Configuração da Transformação de Gerador de Chaves

Ao configurar a transformação de Gerador de Chaves, conecte as portas de fonte de dados que você deseja analisar. Especifique a porta que contém os dados da chave de grupo. Se os registros não incluírem um identificador exclusivo, use a porta de ID de sequência para adicionar identificadores exclusivos ao registro.

Ao especificar a porta da chave de grupo, considere as seguintes diretrizes:

- Selecione uma porta contendo valores que se repetem regularmente nos dados da porta. Preferencialmente, selecione uma porta que crie grupos semelhantes em tamanho.
- Selecione uma porta que não é relevante para a análise de duplicatas.

No exemplo atual, selecione a porta Cidade como a chave de grupo. Se um nome de conta aparecer mais de uma vez em uma cidade, talvez as contas contenham dados duplicados. Se um nome de conta aparecer mais de uma vez em diferentes cidades, é improvável que as contas sejam duplicatas.

**Dica:** Execute um perfil de coluna na fonte de dados antes de selecionar a porta da chave de grupo. Os resultados do perfil podem indicar quantas vezes cada valor aparece em uma porta.

## Configuração da Transformação de Correspondência

Adicione uma transformação de Correspondência não reutilizável ao mapeamento para realizar a análise de campos.

Conclua as seguintes tarefas para configurar a transformação de Correspondência:

1. Selecione o tipo de análise de correspondência a ser realizada.
2. Conecte as portas de entrada à transformação.
3. Configure estratégias para comparar os dados de registros.

4. Selecione o tipo dos dados de saída de correspondência criados pela transformação.
5. Conecte as portas de saída a um destino de dados.

## Selecionar o Tipo de Operação de Correspondência

Use as opções na exibição **Tipo de Correspondência** para selecionar a operação de correspondência. Para comparar os dados de contas, configure a transformação para realizar uma análise de campos de origem única.

## Conectar as Portas de Entrada

Conecte as portas de dados de contas de clientes à transformação de Correspondência.

A transformação de Correspondência usa portas de entrada predefinidas para determinar a ordem na qual ela processa os registros. A transformação usa um identificador de sequência para rastrear os registros das portas de entrada até os pares correspondentes ou clusters que ela grava como saída. A transformação usa uma chave de grupo para classificar os registros processados.

Conecte as seguintes portas predefinidas na transformação de Correspondência às portas de saída predefinidas na transformação de Gerador de Chaves:

- SequenceID
- GroupKey

## Configurar as Estratégias de Análise de Campos

Use as opções na exibição **Estratégias** para configurar as estratégias. As estratégias determinam os tipos de análise que a transformação realiza nos dados de registros.

Crie as seguintes estratégias:

- Crie uma estratégia que use o algoritmo Distância de Edição para analisar os números de identificação dos clientes. Selecione as portas CustomerID\_1 e CustomerID\_2.
- Crie uma estratégia que use o algoritmo Distância de Jaro para analisar os dados de sobrenome. Selecione as portas Lastname\_1 e Lastname\_2.

**Nota:** O algoritmo Distância de Jaro aplica uma penalidade adicional a strings semelhantes que começam com caracteres diferentes. Portanto, o algoritmo Distância de Jaro pode aplicar uma alta pontuação de correspondência a PATTON e PATTEN, mas aplica uma pontuação de correspondência menor a BAYLOR e TAYLOR.

- Crie uma estratégia que use o algoritmo Distância de Hamming Invertida para analisar os dados de CEP. Selecione as portas Zip\_1 e Zip\_2.

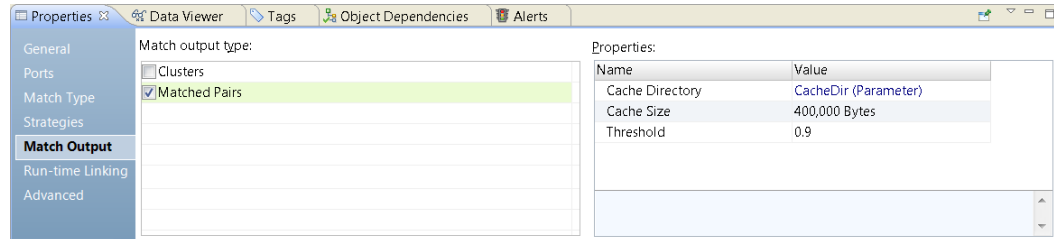
## TÓPICOS RELACIONADOS:

- [“Algoritmos de Correspondência de Campos” na página 487](#)

## Selecionar o Tipo de Saída de Correspondência

Use as opções na exibição **Saída de Correspondência** para definir o formato de saída dos resultados da análise de correspondência.

A seguinte imagem mostra a exibição Saída de Correspondência para a análise de campos de origem única:



Você configura a transformação de forma a organizar os registros de saída em pares correspondentes. Como a transformação retorna registros em pares, os resultados da análise não incluem registros exclusivos.

## Conectar as Portas de Saída

Conecte as portas de saída da transformação de Correspondência ao destino de dados no mapeamento. Selecione as portas contendo os dados de registros que você deseja gravar no destino de dados.

A transformação inclui portas predefinidas que identificam os registros em cada par correspondentes. Os nomes das portas predefinidas são **RowId** e **RowId1**. Cada valor de ID de linha identifica um registro nos dados de saída.

Os valores de ID de linha correspondem às portas que você seleciona em uma estratégia de correspondência. Ao configurar uma estratégia, você seleciona nomes de portas com o sufixo **\_1** ou **\_2**. O valor de RowId identifica o registro que contém uma porta com o sufixo **\_1**. O valor de RowId1 identifica o registro que contém uma porta com o sufixo **\_2**.

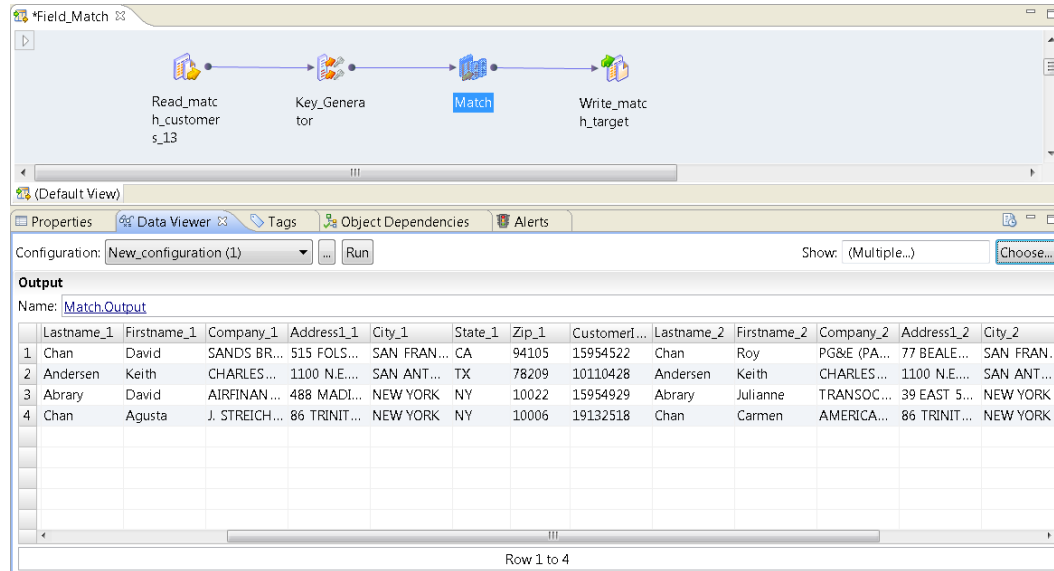
É possível usar outras portas de saída para analisar os relacionamentos entre os registros. Os valores de porta de vínculo e os valores de porta de driver indicam a extensão da semelhança entre os registros em cada cluster.

No exemplo atual, você conecta todas as portas ao destino de dados. Para exibir os dados de saída nessas portas, execute o Visualizador de Dados.

## Executar o Visualizador de Dados

Execute o Visualizador de Dados para analisar os resultados da análise de correspondência. Por padrão, o Visualizador de Dados mostra todas as portas de saída na transformação de Correspondência. Ao executar o mapeamento, você atualiza o destino de dados com os dados das portas de saída.

A seguinte imagem mostra os dados de saída no Visualizador de Dados:



	Lastname_1	Firstname_1	Company_1	Address1_1	City_1	State_1	Zip_1	Customer1...	Lastname_2	Firstname_2	Company_2	Address1_2	City_2
1	Chan	David	SANDS BR...	515 FOLS...	SAN FRAN...	CA	94105	15954522	Chan	Roy	PG&E (PA...	77 BEALE...	SAN FRAN...
2	Andersen	Keith	CHARLES...	1100 N.E....	SAN ANT...	TX	78209	10110428	Andersen	Keith	CHARLES...	1100 N.E....	SAN ANT...
3	Abrary	David	AIRFINAN...	488 MADL...	NEW YORK	NY	10022	15954929	Abrary	Julianne	TRANSOC...	39 EAST 5...	NEW YORK
4	Chan	Agusta	J. STREICH...	86 TRINIT...	NEW YORK	NY	10006	19132518	Chan	Carmen	AMERICA...	86 TRINIT...	NEW YORK

O Visualizador de Dados verifica se os dados de contas de clientes incluem um ou mais registros duplicados.

Considere os seguintes dados no Visualizador de Dados:

- A transformação determina que os registros para Augusta Chan e Carmen Chan podem conter as mesmas informações, pois contêm os mesmos dados de sobrenome e endereço. Ao analisar os registros, você decide que eles são exclusivos no conjunto de dados. No entanto, você percebe que esses registros compartilham um de valor de ID de cliente comum. Como a coluna de ID de cliente é uma chave primária no conjunto de dados, você entra em contato com o escritório de Nova York. O escritório de Nova York resolve o erro.
- A transformação determina que os registros para Keith Anderson podem conter as mesmas informações. Ao analisar os registros, você confirma que os dois registros representam a mesma conta. No entanto, você percebe que os registros têm diferentes valores de ID de cliente. Como uma conta de cliente deve ter um único valor de ID, você entra em contato com o escritório de San Antonio. O escritório de San Antonio resolve o erro.

## Conclusão

Os resultados da análise de correspondência indicam que os dados de contas de clientes incluem pelo menos um par de registros duplicados. Você entra em contato com a agência local das contas que precisam ser verificadas. Você confirma que os outros registros do conjunto de dados identificam exclusivamente uma conta de cliente.

## CAPÍTULO 30

# Transformações de Correspondência na Análise de Identidade

Este capítulo inclui os seguintes tópicos:

- [Análise de Correspondência de Identidade, 498](#)
- [Fluxo de Processos para Análise de Correspondência de Identidade, 499](#)
- [Propriedades do Tipo de Correspondência de Identidade, 500](#)
- [Estratégias de Correspondência de Identidade, 503](#)
- [Opções de Saída de Correspondência de Identidade, 506](#)
- [Propriedades Avançadas de Correspondência de Identidade, 509](#)
- [Estudo de Caso do Índice Persistente, 510](#)
- [Exemplo de Análise de Correspondência de Identidade, 512](#)

## Análise de Correspondência de Identidade

Realize a análise de correspondência de identidade para localizar identidades semelhantes ou duplicadas em um conjunto de dados ou entre dois conjuntos de dados.

Identidades semelhantes em dois ou mais registros podem indicar a duplicação de registros. Ou, as identidades semelhantes podem indicar uma conexão entre os registros, como uma identidade de família compartilhada ou uma identidade de empregador compartilhada.

Ao configurar a transformação de Correspondência para análise de correspondência de identidade, você define as opções nas seguintes exibições:

- Tipo de Correspondência
- Estratégias
- Saída de Correspondência

Opcionalmente, defina as opções nas exibições Parâmetros e Avançado.

Ao configurar uma transformação de Correspondência para a análise de correspondência de identidade, você deve conectar portas de entrada a todos os campos primários obrigatórios em uma estratégia de identidade. A maioria das estratégias de identidade contém campos primários obrigatórios. Algumas

estratégias também contêm campos secundários obrigatórios. Conecte uma porta de entrada a pelo menos um campo obrigatório secundário.

## Fluxo de Processos para Análise de Correspondência de Identidade

O seguinte fluxo de processos resume as etapas realizadas para configurar uma transformação de Correspondência para a análise de correspondência de identidade. É possível definir um processo que utiliza apenas a transformação de Correspondência ou que utiliza a transformação de Correspondência junto com outras transformações.

Antes de se conectar à transformação de Correspondência para objetos de dados upstream, verifique se os registros contêm valores de identificador de sequência exclusivos. É possível usar uma transformação de Gerador de Chaves para criar esses valores. Ao realizar a análise de correspondência de identidade, você tem a opção de organizar os dados de entrada em grupos.

Realize as seguintes etapas na transformação de Correspondência:

1. Especifique a análise de identidade como o tipo de correspondência e defina o número de fontes de dados.  
  
Se você configurar a transformação para analisar dois conjuntos de dados, selecione um conjunto de dados mestre.  
  
Use a exibição **Tipo de Correspondência** para definir o tipo e o número de fontes de dados.
2. Identifique a localização para armazenar os dados de índice. A transformação pode gravar os dados de índice em arquivos temporários ou salvar esses dados em tabelas de banco de dados.  
  
Use a exibição **Tipo de Correspondência** para especificar o armazenamento de dados de índice.
3. Defina uma estratégia de análise de correspondência. Selecione uma população e um algoritmo de comparação, e atribua um par de colunas ao algoritmo.  
  
O preenchimento indica os pares de colunas a serem selecionados.  
  
Use a exibição **Estratégias** para definir a estratégia.
4. Especifique o método usado pela transformação para gerar resultados de análise de correspondência.
5. Defina o valor de limite de correspondência. O limite de correspondência é a pontuação mínima que pode identificar dois registros como duplicatas um do outro.  
  
Use a exibição **Saída de Correspondência** para selecionar o método de saída e o limite de correspondência.  
  
**Nota:** É possível definir o limite de correspondência em uma transformação de Correspondência ou em uma transformação de Média Ponderada. Use a transformação de Média Ponderada se criar um maplet de correspondência.

# Propriedades do Tipo de Correspondência de Identidade

Use a exibição Tipo de Correspondência para especificar o tipo de análise de que a transformação de Correspondência executa e configurar as propriedades que definem a análise. Você pode especificar a análise de origem única ou de origem dupla. Você também pode especificar um armazenamento de dados persistentes para os dados de índice de identidade.

As propriedades que você configura dependem do tipo de análise que você selecionar. Muitas das opções são comuns a todos os tipos de análise.

## Propriedades Comuns

As seguintes propriedades são comuns a todos os tipos de análise de identidade:

### Preenchimento

Identifica o arquivo de preenchimento usado pela transformação. O arquivo de preenchimento contém os algoritmos de criação de chaves que geram as chaves de índice.

### Nível de Chave

Determina o número de chaves geradas pelos algoritmos de identidade. A configuração padrão é Padrão. A configuração Limitada resulta em um número menor de chaves, maior precisão e mais tempo de processamento. A configuração Estendida resulta em um número maior de chaves, menor precisão e tempo de processamento mais curto.

### Tipo de Chave

Descreve o tipo de informações contida no campo de chave. A análise de identidade pode gerar chaves para nomes de pessoas, organizações e endereços. Selecione o tipo de chave que melhor descreve a coluna especificada na propriedade **Campo de Chave**.

### Nível de Pesquisa

Indica o equilíbrio entre a profundidade e a velocidade da pesquisa que a transformação aplica na análise de correspondência. A profundidade da pesquisa é inversamente correlacionada ao número de correspondências retornadas. Por exemplo, a opção **Abrangente** retorna menos correspondências.

### Campo de Chave

Especifica a coluna utilizada pela transformação de Correspondência para gerar os dados de chaves de índice. Verifique se a coluna selecionada contém o tipo de informações especificado na propriedade **Tipo de chave**.

### Diretório de Índice

Identifica o diretório no qual o Serviço de Integração de Dados grava os dados de chave de índice da transformação atual. Por padrão, a propriedade está em branco. Se você não especificar um diretório de índice, o Serviço de Integração de Dados usará a localização definida no Serviço do Gerenciamento de Conteúdo.

Você pode inserir um caminho para o diretório ou usar um parâmetro para identificá-lo. Especifique um caminho local na máquina do Serviço de Integração de Dados. O Serviço de Integração de Dados deve ser capaz de gravar no diretório.

### Diretório de Cache

Identifica o diretório no qual o Serviço de Integração de Dados grava os dados temporários durante a fase de criação do índice da análise de correspondência de identidade. Atualize a propriedade para especificar uma localização para os dados da transformação atual. Por padrão, a propriedade está em



branco. Se você não especificar um diretório de cache, o Serviço de Integração de Dados usará a localização definida no Serviço do Gerenciamento de Conteúdo.

Você pode inserir um caminho para o diretório ou usar um parâmetro para identificá-lo. Especifique um caminho local na máquina do Serviço de Integração de Dados. O Serviço de Integração de Dados deve ser capaz de gravar no diretório.

#### **Tamanho do Cache**

Determina a quantidade de memória do sistema que o Serviço de Integração de Dados atribui à criação de índice de identidade. O valor padrão é 400.000 bytes.

Se a operação de criação de índice gerar uma quantidade de dados maior, o Serviço de Integração de Dados gravará o excesso de dados no diretório de cache. Se a operação exigir mais memória do que a memória do sistema e o armazenamento de arquivos podem fornecer, o mapeamento falhará.

**Nota:** Se você inserir um valor igual ou maior que 65536, a transformação lerá o valor em bytes. Se você inserir um valor menor, a transformação lerá o valor em megabytes.

#### **Propriedades de origem dupla**

Defina a seguinte propriedade, além das propriedades comuns, ao configurar a transformação para a análise de origem dupla:

##### **Conjunto de Dados Mestre**

Identifica a fonte de dados que contém os dados mestre. Especifique um conjunto de dados mestre na análise de origem dupla.

#### **Propriedades do Armazenamento de Dados Persistentes**

Defina as seguintes propriedades, além das propriedades comuns, ao configurar a transformação para usar o armazenamento de dados de índice persistente:

##### **Método de Persistência**

Especifica se a transformação atualiza as tabelas de índice atuais com dados de índice da fonte de dados do mapeamento. Selecione uma das seguintes opções:

- **Atualizar o banco de dados com novas IDs.**  
A transformação adiciona todas as linhas aos dados de índice que não duplicam um identificador de sequência nos dados de índice. A transformação não atualiza as linhas atuais no índice.  
Por padrão, a transformação realiza a análise de correspondência quando você seleciona a opção. Você pode usar a opção Processo de Correspondência para ativar ou desativar a análise de correspondência.
- **Não atualizar o banco de dados.**  
A transformação não atualiza as tabelas de índice com dados de índice da fonte de dados do mapeamento.  
A transformação realiza a análise de correspondência quando você seleciona a opção.
- **Remova os IDs do banco de dados.**  
A transformação excluirá linhas das tabelas de índice se as linhas compartilharem identificadores de sequência com os dados de origem de mapeamento.  
A transformação não realiza a análise de correspondência quando você seleciona a opção.

- Atualize os IDs atuais no banco de dados.  
A transformação substituirá as linhas nas tabelas de índice por linhas dos dados de origem do mapeamento se as linhas compartilharem identificadores de sequência. A transformação não adiciona linhas ao índice.  
Por padrão, a transformação realiza a análise de correspondência quando você seleciona a opção. Você pode usar a opção Processo de Correspondência para ativar ou desativar a análise de correspondência.

O método de persistência padrão é **Atualizar o Banco de Dados com Novos IDs**.

#### Processo de Correspondência

Determina se a transformação atual realiza a análise de identidade.

A opção que você seleciona na propriedade Método de Persistência determina as opções na propriedade Processo de Correspondência.

#### Conexão de DB

Identifica o banco de dados que contém as tabelas de índice.

#### Armazenamento Persistente

Identifica as tabelas de índice no banco de dados que você especifica.

### TÓPICOS RELACIONADOS:

- [“Estudo de Caso do Índice Persistente” na página 510](#)
- [“Parâmetros de Métodos de Persistência” na página 503](#)

## Propriedades do Diretório de Índice e do Diretório de Cache

O diretório de índice e o diretório de cache armazenam os dados temporários que a transformação de Correspondência gera durante a análise de identidade.

O Serviço de Integração de Dados gravará os dados no diretório de índice se você não configurar a transformação para gravar dados de índice nas tabelas de banco de dados. O Serviço de Integração de Dados gravará dados no diretório de cache se a análise de identidade exigir uma quantidade maior de memória do que a especificada no tamanho do cache. Por padrão, as propriedades estão em branco na exibição Tipo de Correspondência. Se você não especificar um diretório de índice ou um diretório de cache, o Serviço de Integração de Dados lerá os caminhos de diretório do Serviço do Gerenciamento de Conteúdo.

Quando você especificar um diretório de índice ou de cache na exibição Tipo de Correspondência, insira um caminho local na máquina do Serviço de Integração de Dados. Você pode inserir um caminho totalmente qualificado ou relativo. Se inserir um caminho relativo, inicie o caminho com um ponto. O caminho é relativo ao diretório `tomcat/bin` na máquina do Serviço de Integração de Dados.

A seguinte tabela mostra um caminho relativo em um diretório de cache ou em uma propriedade do diretório de índice e identifica o caminho totalmente qualificado que a propriedade representa:

Caminho Relativo	Caminho Totalmente Qualificado
<code>./ch</code>	<code>[Diretório_da_instalação_Informatica]/tomcat/bin/ch</code>

Você pode configurar a propriedade do diretório de índice e a propriedade do diretório de cache para identificar o mesmo diretório. Se você especificar o mesmo diretório em cada propriedade, o Serviço de Integração de Dados criará diretórios no diretório especificado. O Serviço de Integração de Dados cria um diretório de `index` para os dados de índice e um diretório `cache` para os dados de cache. Se você especificar

diferentes diretórios em cada propriedade, o Serviço de Integração de Dados gravará os dados nos diretórios especificados.

O Serviço de Integração de Dados exclui os arquivos de índice e de cache dos diretórios após a execução do mapeamento. O Serviço de Integração de Dados não exclui os diretórios. O Serviço de Integração de Dados poderá reutilizar os diretórios em outros mapeamentos se a transformação de Correspondência identificá-los.

### Propriedades do Serviço do Gerenciamento de Conteúdo

O Serviço do Gerenciamento de Conteúdo especifica as seguintes localizações padrão para o diretório de índice e de cache:

- `./identityIndex`. O diretório padrão dos dados de índice de identidade.
- `./identityCache`. O diretório padrão dos dados de cache de identidade.

Se você não definir as propriedades na transformação de Correspondência, o Serviço de Integração de Dados criará os diretórios padrão quando um mapeamento de correspondência de identidade for executado. O Serviço de Integração de Dados cria os diretórios no diretório `tomcat/bin`.

## Parâmetros de Métodos de Persistência

Ao selecionar um índice de dados de persistência na análise de correspondência de identidades, você pode usar um parâmetro para identificar o método de persistência. Use uma cadeia para definir o valor do parâmetro.

A seguinte tabela lista os métodos de persistência que você pode selecionar na guia **Tipo de Correspondência** e os valores de parâmetros que podem ser definidos para os métodos:

Método de Persistência	Parâmetro
Atualizar o banco de dados com novos IDs	ignore
Não atualizar o banco de dados	addNone
Remover IDs do banco de dados	remove
Atualizar os IDs atuais no banco de dados	update

Os valores de parâmetro fazem distinção entre maiúsculas e minúsculas.

### TÓPICOS RELACIONADOS:

- [“Propriedades do Tipo de Correspondência de Identidade” na página 500](#)

## Estratégias de Correspondência de Identidade

A exibição Estratégias lista as estratégias que você define para os dados de identidade. Essas estratégias determinam como a transformação mede as semelhanças e as diferenças entre os dados no índice de identidade.

## Algoritmos de Correspondência de Identidade

A transformação de Correspondência inclui algoritmos de identidade predefinidos que comparam os valores de dados no índice de identidade. Selecione o algoritmo que melhor represente o tipo de dados de identidade no conjunto de dados.

A tabela a seguir descreve os algoritmos e identifica as entradas que você seleciona para cada um:

Algoritmo de Identidade	Descrição
Endereço	Identifica registros que compartilham um endereço. O algoritmo requer a seguinte entrada primária: <ul style="list-style-type: none"><li>- Endereço</li></ul> O algoritmo não requer uma entrada secundária.
Contato	Identifica registros que compartilham um contato no local de uma única organização. O algoritmo requer as seguintes entradas primárias: <ul style="list-style-type: none"><li>- Person_Name</li><li>- Organization_Name</li><li>- Address_Part1</li></ul> O algoritmo não requer uma entrada secundária.
Entidade Corporativa	Identifica registros que compartilham dados de identificação corporativa. Opcionalmente, selecione esse algoritmo para analisar dados de endereços e telefones. O algoritmo requer a seguinte entrada primária: <ul style="list-style-type: none"><li>- Organization_Name</li></ul> O algoritmo não requer uma entrada secundária.
Divisão	Identifica registros que compartilham uma localização de escritório dentro de uma organização. O algoritmo requer as seguintes entradas primárias: <ul style="list-style-type: none"><li>- Organization_Name</li><li>- Address_Part1</li></ul> O algoritmo não requer uma entrada secundária.
Família	Identifica indivíduos que pertencem a uma família. Analisa dados de nome, endereço e número de telefone. O algoritmo requer a seguinte entrada primária: <ul style="list-style-type: none"><li>- Person_Name</li></ul> O algoritmo requer uma das seguintes entradas secundárias: <ul style="list-style-type: none"><li>- Address_Part1</li><li>- Telephone_Number</li></ul>
Campos	Identifica registros que compartilham dados nas portas que você seleciona. O algoritmo não especifica uma entrada obrigatória. Selecione uma ou mais portas que podem conter dados de identidade duplicados.
Membros da Família	Identifica indivíduos que pertencem a uma unidade familiar. Analisa dados de nome e dados de endereço. O algoritmo requer as seguintes entradas primárias: <ul style="list-style-type: none"><li>- Person_Name</li><li>- Address_Part1</li></ul>

Algoritmo de Identidade	Descrição
Indivíduo	Identifica indivíduos duplicados. Analisa dados de nome, data de nascimento e identificação pessoal, como números de previdência social, conta e identificação de veículo. O algoritmo requer a seguinte entrada primária: - Person_Name O algoritmo requer uma das seguintes entradas secundárias: - Data - ID
Organização	Identifica registros que compartilham dados da organização. O algoritmo requer a seguinte entrada primária: - Organization_Name O algoritmo não requer uma entrada secundária.
Nome da Pessoa	Identifica registros que compartilham informações sobre indivíduos. O algoritmo requer a seguinte entrada primária: - Person_Name O algoritmo não requer uma entrada secundária.
Residente	Identifica indivíduos duplicados em um endereço. Opcionalmente, configure essa estratégia para analisar dados de identificação pessoal. O algoritmo requer a seguinte entrada primária: - Person_Name - Address_Part1 O algoritmo não requer uma entrada secundária.
Contato Amplo	Identifica registros que compartilham um contato em uma organização. O algoritmo requer as seguintes entradas primárias: - Person_Name - Organization_Name O algoritmo não requer uma entrada secundária.
Família Ampla	Identifica indivíduos que pertencem à mesma família. O algoritmo requer a seguinte entrada primária: - Address_Part1 O algoritmo não requer uma entrada secundária.

## Propriedades da Estratégia de Correspondência de Identidade

Configure as propriedades para cada estratégia de identidade.

Ao configurar uma estratégia de identidade, você pode configurar as seguintes propriedades da estratégia:

### Preenchimento

Determina o preenchimento a ser aplicado à análise de identidade. Os preenchimentos contêm algoritmos de criação de chaves para localidades e idiomas específicos.

### Nível de correspondência

Indica o equilíbrio entre a qualidade e a velocidade da pesquisa. A velocidade da pesquisa é inversamente proporcional ao número de correspondências retornado. Pesquisas que usam a configuração **Ampla** retornam menos correspondências, enquanto que pesquisas que usam a configuração **Conservadora** retornam mais correspondências.

# Opções de Saída de Correspondência de Identidade

A exibição Saída de Correspondência inclui opções que especificam o formato dos dados de saída. É possível configurar a transformação para gravar registros em clusters ou em pares correspondentes. Também é possível configurar a transformação para incluir ou excluir diferentes categorias de identidade quando você realiza a análise de identidade em um armazenamento de dados de índice persistente.

Você configura as opções na área **Tipo de saída de correspondência** e na área **Propriedades**.

## Tipos de Saída de Correspondência

A exibição Saída de Correspondência inclui opções que especificam o formato dos dados de saída. É possível configurar a transformação para gravar registros em clusters ou em pares correspondentes.

Selecione um dos seguintes tipos de saída de correspondência:

### Melhor Correspondência

Grava cada registro no conjunto de dados mestre com o registro que representa a melhor correspondência no segundo conjunto de dados. A operação de correspondência seleciona o registro no segundo conjunto de dados com a pontuação de correspondência mais alta do registro principal. Se dois ou mais registros retornarem a pontuação mais alta, a operação de correspondência selecionará o primeiro registro do segundo conjunto de dados. A opção Melhor Correspondência grava cada par de registros em uma única linha.

Você pode selecionar **Melhor Correspondência** ao configurar a transformação para a análise de origem dupla.

### Clusters - Melhor Correspondência

Grava clusters que representam a melhor correspondência entre um registro e outro no mesmo conjunto de dados ou entre dois conjuntos de dados. A pontuação de correspondência entre os dois registros deve atender o limite de correspondência. Clusters com melhores correspondências poderão conter mais de dois registros se um registro representar a melhor correspondência com mais de um outro registro.

É possível selecionar **Clusters - Melhor Correspondência** em qualquer tipo de análise de identidade.

**Nota:** O método de armazenamento de dados de índice que você selecionar pode afetar o conteúdo da saída de cluster no modo **Clusters - Melhor Correspondência**. Uma transformação que se conecta a tabelas de índice pode criar clusters diferentes de uma transformação que armazena dados de índice para os mesmos registros em arquivos temporários. O método de armazenamento de dados de índice não afeta as pontuações de correspondência gerados pela transformação para os pares de registros.

### Clusters - Corresponder Tudo

Grava clusters de registros que são correspondidos com uma pontuação que atende ao limite de correspondência. Cada registro deve corresponder a pelo menos um outro registro no cluster.

É possível selecionar **Clusters - Corresponder Tudo** em qualquer tipo de análise de identidade.

### Pares Correspondentes

Grava todos os pares de registros que são correspondidos entre si com uma pontuação que atende ao limite de correspondência. A transformação grava cada par em uma única linha e adiciona a pontuação de correspondência de cada par a cada linha. Se um registro corresponder a mais de um outro registro, a transformação gravará uma linha para cada par de registros.

É possível selecionar **Pares Correspondentes** em qualquer tipo de análise de identidade.

## Propriedades da Saída de Correspondência

A exibição Saída de Correspondência contém propriedades que especificam o comportamento da memória cache e o limite de pontuação de correspondência. Você também pode usar as propriedades para determinar como a transformação seleciona registros de armazenamento de dados para análise e como ela grava esses registros como saída.

Depois de selecionar um tipo de saída de correspondência, configure as seguintes propriedades:

### Diretório de Cache

Especifica o diretório no qual o Serviço de Integração de Dados grava dados temporários durante a análise de correspondência de identidade. O Serviço de Integração de Dados grava os arquivos temporários no diretório quando o volume de dados que a análise de correspondência gera é maior do que a memória disponível do sistema. O Serviço de Integração de Dados exclui os arquivos temporários depois que o mapeamento é executado.

Você pode inserir um caminho de diretório na propriedade ou usar um parâmetro do sistema para identificar o diretório. Especifique um caminho local na máquina do Serviço de Integração de Dados. O Serviço de Integração de Dados deve ser capaz de gravar no diretório. O valor padrão é o parâmetro do sistema CacheDir.

### Tamanho do Cache

Determina a quantidade de memória do sistema que o Serviço de Integração de Dados atribui à análise de correspondência de identidade. O valor padrão é 400.000 bytes.

Se a análise de correspondência gerar uma quantidade maior de dados, o Serviço de Integração de Dados gravará o excesso de dados no diretório de cache. Se a análise de correspondência exigir mais memória do que a memória do sistema e o armazenamento de arquivos podem fornecer, o mapeamento falhará.

**Nota:** Se você inserir um valor igual ou maior que 65536, a transformação lerá o valor em bytes. Se você inserir um valor menor, a transformação lerá o valor em megabytes.

### Correspondência

Identifica os registros a serem analisados quando a transformação lê dados de índice de tabelas do banco de dados. Use as opções na exibição **Tipo de Correspondência** para identificar as tabelas de índice.

Por padrão, a transformação analisa todos os registros na fonte de dados e nas tabelas de banco de dados de índice. Configure a propriedade Correspondência para especificar um subconjunto dos registros para análise duplicada.

### Saída

Filtra os registros que a transformação grava como saída quando você configura essa transformação para ler tabelas de banco de dados de índice. Use as opções na exibição **Tipo de Correspondência** para identificar as tabelas de índice.

Por padrão, a transformação de Correspondência grava todos os registros da fonte de dados e das tabelas de banco de dados de índice como saída. Configure a propriedade **Saída** quando você não precisar examinar todos os registros nos dados de entrada.

### Limite

Define a pontuação de correspondência mínima que identifica dois registros como possíveis duplicatas um do outro.

É possível atribuir um parâmetro ao valor de limite. Defina um valor decimal no intervalo de 0 a 1.

## Configuração da Propriedade Correspondência

Use a propriedade **Correspondência** na exibição **Saída de Correspondência** para especificar como a transformação seleciona dados de entrada para análise. Configure a propriedade quando você configurar a transformação de Correspondência para ler um armazenamento persistente de dados de índice. A propriedade Correspondência refina as opções que você definir na exibição **Tipo de Correspondência**.

Você pode configurar a propriedade Correspondência para realizar os seguintes tipos de análise:

### Comparar os registro de fonte de dados com os registros de dados de índice

Para procurar registros duplicados entre a fonte de dados e as tabelas de dados de índice, selecione **Exclusivo**.

Quando você seleciona a opção Exclusivo, a transformação de Correspondência compara os registros de fonte de dados com o armazenamento de dados de índice. A transformação não analisa os registros dentro da fonte de dados ou do armazenamento de dados.

Selecione a opção **Exclusiva** se você souber que o armazenamento de dados de índice e a fonte de dados não contêm registros duplicados.

### Comparar os registros de fonte de dados com os registros de dados de índice e comparar os registros de fonte de dados entre si

Para procurar duplicatas na fonte de dados e entre a fonte de dados e as tabelas de índice, selecione **Parcial**.

A transformação compara os registros de fonte de dados com o armazenamento de dados de índice. A transformação também compara os registros dentro da fonte de dados entre si.

Selecione **Parcial** se você souber que o armazenamento de dados de índice não contém registros duplicados, mas ainda não realizou uma análise de duplicatas na fonte de dados.

### Comparar todos os registros na fonte de dados e nas tabelas de índice como um único conjunto de dados

Para procurar duplicatas entre a fonte de dados e as tabelas de índice e dentro da fonte de dados e das tabelas de índice, selecione **Completo**. A opção padrão é Completo.

A transformação analisa a fonte de dados e o armazenamento de dados como um único conjunto de dados e compara todos os dados de registros nesse conjunto de dados.

Selecione a opção **Completa** quando não for possível verificar se o conjunto de dados contém ou não registros duplicados.

## Configuração da Propriedade Saída

Use a propriedade **Saída** na exibição **Saída de Correspondência** para filtrar os registros que a transformação grava como saída. Configure essa propriedade quando você especificar tabelas de dados de índice e selecionar um formato de saída em cluster. Filtre os registros para limitar a saída a clusters que contenham um ou mais registros da fonte de dados.

É possível filtrar os dados de saída das seguintes maneiras:

### Gravar cada cluster que inclui um registro proveniente da fonte de dados ou das tabelas de índice

Selecione **Todas as Linhas**. A transformação grava cada cluster que contém pelo menos um registro proveniente da fonte de dados ou do armazenamento de dados de índice. O padrão é Todas as Linhas.

Como um cluster pode conter um único registro, a saída contém todos os registros.

### Gravar cada cluster que inclui um registro proveniente da fonte de dados

Selecione **Linhas Novas e Associadas**. A transformação grava cada cluster que contém pelo menos um registro proveniente da fonte de dados.



Como um cluster pode conter um único registro, a saída contém todos os registros na fonte de dados. Os clusters também podem incluir registros das tabelas de índice.

#### Gravar cada cluster da fonte de dados

**Selecione Novas Linhas Somente.** A transformação grava os clusters que contêm registros provenientes da fonte de dados. A saída não contém registros provenientes das tabelas de índice.

## Propriedades Avançadas de Correspondência de Identidade

A transformação incluirá as propriedades avançadas que determinam o número de instâncias de execução se analisar as linhas idênticas e o nível de rastreamento dos dados de log.

Você pode configurar as seguintes propriedades avançadas:

#### Instâncias de Execução

Especifica o número de segmentos que o Serviço de Integração de Dados tenta criar para a transformação atual em tempo de execução. O Serviço de Integração de Dados considerará o valor Instâncias de Execução se você substituir a propriedade de tempo de execução Paralelismo Máximo no mapeamento que contém a transformação. O valor padrão de Instâncias de Execução é Automático.

O Serviço de Integração de Dados considera vários fatores para determinar o número de segmentos a serem atribuídos à transformação. Os principais fatores são o valor de Instâncias de Execução e os valores no mapeamento e nos serviços de aplicativo associados no domínio.

O Serviço de Integração de Dados que executa o mapeamento lê os seguintes valores para determinar o número de segmentos a serem usados para a transformação:

- O valor máximo de *Paralelismo Máximo* no Serviço de Integração de Dados. O padrão é 1.
- Qualquer valor de *Paralelismo Máximo* que você define no nível de mapeamento. O padrão é Automático.
- O valor de *Instâncias de Execução* na transformação. O padrão é Automático.

Se você substituir o valor de Paralelismo Máximo no nível de mapeamento, o Serviço de Integração de Dados tentará usar o valor mais baixo nas propriedades para determinar o número de segmentos.

Se você usar o valor de padrão de Paralelismo Máximo no nível de mapeamento, o Serviço de Integração de Dados ignorará o valor de Instâncias de Execução.

Considere as seguintes regras e diretrizes quando você definir o número de instâncias de execução:

- Vários usuários podem executar mapeamentos simultâneos em um Serviço de Integração de Dados. Para calcular o número correto de segmentos, divida o número de unidades centrais de processamento que o Serviço de Integração de Dados pode acessar pelo número de mapeamentos simultâneos.
- Quando você usa o valor padrão de Instâncias de Execução e os valores padrão de Paralelismo Máximo, as operações de transformação não são particionáveis.

#### Correspondência Exata do Filtro

Determina se a transformação aplica o algoritmo de comparação em uma estratégia de correspondência a pares de registros idênticos nos dados de entrada.

Quando a transformação encontra um par de registros idênticos, o algoritmo não precisa analisar os níveis de similaridade entre os registros. A transformação pode transmitir os registros diretamente para

a fase de saída sem análise adicional. Para configurar a transformação para transmitir os registros idênticos diretamente para a fase de saída, selecione Filtrar Correspondência Exata. Quando os dados de entrada contêm muitas linhas idênticas, o algoritmo de comparação realiza menos cálculos e o mapeamento é executado mais rapidamente.

Selecione a opção quando os dados de entrada contiverem muitas linhas idênticas. Não selecione a opção se os dados de entrada não contiverem muitas linhas idênticas porque a transformação poderá ser executada mais lentamente.

**Nota:** A saída de transformação contém os mesmos dados de registro quando você seleciona ou desmarca a opção. A transformação pode atribuir diferentes pontuações de link e de driver para os registros de saída quando você seleciona e desmarca a opção.

#### Nível de Rastreamento

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.

## Estudo de Caso do Índice Persistente

Você tem um administrador de dados em um banco com várias filiais. Você gerencia um conjunto mestre dos registros de conta do cliente de todas as filiais. Você usa um conjunto de tabelas de banco de dados de índice para verificar se o banco de dados da conta do cliente contém registros redundantes ou duplicados.

Para criar e gerenciar o armazenamento de dados de índice, realize as seguintes operações:

- Crie o armazenamento de dados.
- Atualize-o com os dados mais recentes das filiais do banco.

Você pode adicionar dados da conta ao armazenamento de dados ou atualizar os dados atuais no armazenamento de dados.

- Remova registros obsoletos do armazenamento de dados.

Você compreende que cada operação pode criar registros duplicados no armazenamento de dados. Você decide desenvolver uma política para analisar os dados das filiais antes de adicioná-los aos dados do armazenamento de dados mestre. Use a análise de correspondência de identidade para analisar os dados das filiais e verificar se eles criam identidades duplicadas no armazenamento de dados. Configure as opções de índice persistente na transformação de Correspondência para analisar os dados das filiais e o armazenamento de dados.

### Desenvolver uma Política de Gerenciamento de Dados de Índice Persistente

Como um administrador de dados, você define uma regra comercial que afirma que o armazenamento de dados da conta do cliente não pode conter identidades duplicadas. Você projeta um mapeamento de correspondência de identidade para analisar os dados das filiais em um banco de dados preparação antes de adicioná-los ao armazenamento de dados.

As operações para adicionar os dados das filiais ao armazenamento de dados podem criar identidades duplicadas nos seguintes casos:

- Os dados das filiais contêm identidades duplicadas.
- Os dados das filiais contêm uma identidade que o índice também contém.
- Os dados das filiais contêm uma versão mais recente de uma identidade no armazenamento de dados, e a versão mais recente corresponde a outra identidade no índice.

Quando você comparar o banco de dados de preparação com o armazenamento de dados, selecione as opções de índice persistente que refletem o status de registro duplicado dos dados das filiais. Antes de atualizar o armazenamento de dados, talvez você decida comparar os dados das filiais com os dados de índice.

**Nota:** Você pode ativar e desativar a análise de correspondência em algumas das opções. Ative a análise de correspondência para analisar os dados de mapeamento ou para comparar o armazenamento de dados de índice com os dados de mapeamento. Desative a análise de correspondência quando você não precisar comparar os dados. Você também pode usar as propriedades de correspondência na guia Saída de Correspondência para incluir ou excluir os dados da análise de correspondência.

## Comparar uma Fonte de Dados de Mapeamento com o Armazenamento de Dados de Índice

Para comparar os dados de entrada do mapeamento com o armazenamento de dados de índice e para não fazer alterações no armazenamento de dados, selecione a seguinte opção:

- Não atualizar o banco de dados

O mapeamento compara os dados de entrada com o armazenamento de dados de índice. O mapeamento não adiciona, remove ou atualiza os dados no armazenamento de dados de índice.

Você não pode desativar a análise de correspondência de identidade ao selecionar a opção.

Como você não atualiza os dados de índice, não pode criar linhas duplicadas no armazenamento. Selecione a opção das propriedades de Correspondência na guia Saída de Correspondência que atenda às necessidades atuais do projeto de dados. Por exemplo, selecione a opção **Completa**. A opção **Completa** verifica se os dados de mapeamento contêm duplicatas e se eles adicionam duplicatas ao armazenamento de dados.

**Nota:** Use a opção para comparar os dados de mapeamento e o armazenamento de dados antes de atualizar o armazenamento de dados. Se a saída do mapeamento indicar que os dados de mapeamento não adicionam duplicatas ao armazenamento de dados, execute o mapeamento novamente. Selecione a opção para atualizar o banco de dados quando você executar o mapeamento novamente.

## Criar e Adicionar Linhas ao Armazenamento de Dados

Para criar um armazenamento de dados ou para adicionar linhas dos dados de mapeamento a um armazenamento de dados, selecione a seguinte opção:

- Atualizar o banco de dados com novos IDs

O mapeamento adicionará uma linha ao armazenamento de dados se a linha não compartilhar um identificador de sequência com uma linha no armazenamento de dados. O mapeamento não substitui linhas nas tabelas de índice. Quando você especifica tabelas de banco de dados vazias, o mapeamento grava todos os dados de índice de mapeamento nas tabelas.

Você pode ativar ou desativar a análise de correspondência de identidade ao selecionar a opção. A opção ativa a análise de correspondência por padrão.

Como você não atualiza as linhas de índice, selecione a opção **Exclusiva** ou **Parcial** das propriedades de Correspondência na guia Saída de Correspondência. Use a opção **Exclusiva** se você tiver verificado a exclusividade das linhas de dados de mapeamento em um processo anterior.

## Atualizar as Linhas no Armazenamento de Dados

Para atualizar uma linha atual no armazenamento de dados com os dados de mapeamento, selecione a seguinte opção:

- Atualize os IDs atuais no banco de dados.

O mapeamento atualizará um registro atual no armazenamento de dados se o registro compartilhar um identificador de sequência com um registro nos dados de mapeamento. O mapeamento não adiciona linhas às tabelas de índice.

Você pode ativar ou desativar a análise de correspondência de identidade ao selecionar a opção. A opção desativa a análise de correspondência por padrão.

Como você não adiciona linhas de índice às tabelas de índice, selecione a opção **Completa** das propriedades de Correspondência na guia saída de correspondência.

**Nota:** Ao atualizar as linhas no armazenamento de dados, você espera localizar duplicatas entre os dados de origem de mapeamento e o armazenamento de dados. Selecione a opção **Completa** para verificar se os dados de identidade que você adiciona ao armazenamento correspondem aos dados atuais no armazenamento.

## Remover Linhas do Armazenamento de Dados

Para remover linhas do armazenamento de dados, selecione a seguinte opção:

- Remover IDs do banco de dados

O mapeamento excluirá uma linha do armazenamento de dados se a linha compartilhar um identificador de sequência com um registro nos dados de mapeamento.

Você pode ativar ou desativar a análise de correspondência de identidade ao selecionar a opção. A opção desativa a análise de correspondência por padrão.

**Nota:** Quando você remover os dados de um armazenamento de dados, altere os relacionamentos entre as linhas no armazenamento. Se o armazenamento contiver identidades duplicadas, você poderá remover os dados de um registro de driver ou um registro vinculado em um cluster. Ou você pode remover os dados da melhor correspondência em um par correspondido. Quando você executar o mapeamento novamente, ele poderá gerar diferentes clusters ou pares duplicados. Se você remover linhas de um armazenamento de dados que não contém registros duplicados, não poderá alterar o status duplicado dos registros. Quando você executa o mapeamento depois de excluir as linhas, ele gera as mesmas pontuações de correspondência para as identidades que permanecem no conjunto de dados.

# Exemplo de Análise de Correspondência de Identidade

Você tem é um responsável pelos recursos humanos em uma organização de software com centros de desenvolvimento em diferentes cidades. A organização armazena os registros de pessoal para dos membros da equipe em um banco de dados no escritório matriz. Os centros de desenvolvimento contratam equipes em intervalos regulares, e os centros enviam para você os dados de pessoal dos membros da equipe que contratam.

Adicione os registros de pessoal a um arquivo de planilha e use os dados do arquivo para atualizar o banco de dados de funcionários. Você está preocupado com a possibilidade de que o arquivo contenha identidades duplicadas.

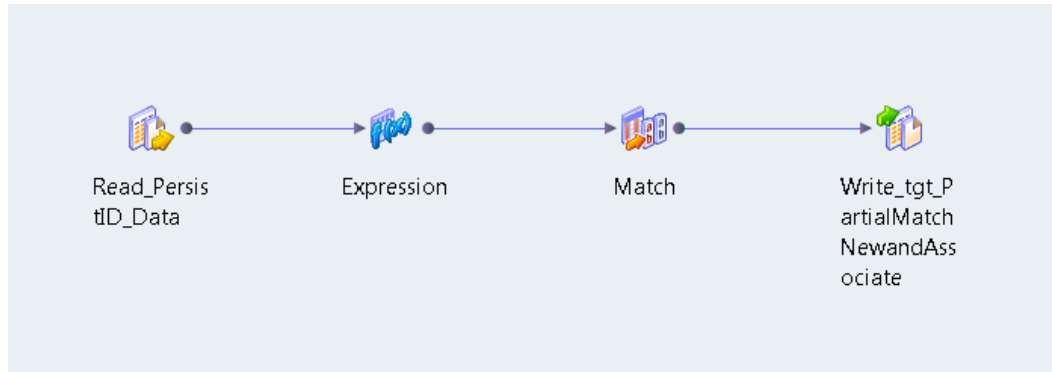
Por isso, decide criar um mapeamento para realizar uma análise de identidade nos registros de funcionários. Configure uma transformação de Correspondência para pesquisar por identidades duplicadas no arquivo de planilha. Você também deve verificar se os dados do arquivo duplicam algum dado dos funcionários no banco de dados mestre. Configure a transformação de Correspondência para comparar os dados do arquivo com os dados mestre que você armazena para os funcionários da organização.

Como o banco de dados é um conjunto de dados mestre, você armazena os dados de índice dos registros de equipe em um armazenamento de dados persistentes.

## Criar o Mapeamento

Crie um mapeamento que procure identidades duplicadas. O mapeamento lê uma fonte de dados, adiciona um identificador de sequência aos registros de origem, realiza a análise de identidade e grava os resultados em um destino de dados.

A seguinte imagem mostra o mapeamento na ferramenta Developer:



O mapeamento criado contém os seguintes objetos:

Nome do Objeto	Descrição
Read_PersistID_Data	Fonte de dados. Contém os nomes e detalhes dos funcionários.
Expressão	Transformação de Expressão. Adiciona valores de identificador de sequência aos dados de origem.
Correspondência	Transformação de Correspondência. Analisa os níveis de duplicação nas identidades de dados de origem.
Write_tgt_PartialMatchNewandAssociate	Destino de dados. Contém os resultados da análise de identidade.

**Nota:** O mapeamento não usa uma transformação de Gerador de Chaves. Na análise de correspondência de identidade, a transformação de Gerador de Chaves é opcional.

## Amostra de Dados de Entrada

O arquivo da equipe contém o nome do funcionário, a cidade na qual ele trabalha e sua função designada. Você cria uma fonte de dados a partir do arquivo da equipe no repositório do Modelo. Você adiciona a fonte de dados ao mapeamento.

O seguinte fragmento de dados apresenta uma amostra dos dados dos funcionários no arquivo da equipe:

Nome	Cidade	Designação
Chaithra	Bangalore	SE
Ramanan	Chennai	SSE

Nome	Cidade	Designação
Ramesh	Chennai	SSE
Ramesh	Chennai	Líder
Sunil	Bangalore	Principal
Venu	Hyderabad	Principal
Harish	Bangalore	SE
Sachin	Bangalore	SSE

## Configuração da Transformação de Expressão

Ao configurar a transformação de Expressão, conecte todas as portas de fonte de dados que você deseja incluir na saída de mapeamento. Conecte as portas como portas de passagem. Criar uma expressão para adicionar um valor de Identificação de sequência às portas.

A seguinte expressão cria a variável *Init3* e adiciona o valor inteiro 1267 a cada identificador de sequência:

*Init3+1267*

A seguinte tabela descreve as portas na transformação de Expressão que lê a fonte de dados de funcionários:

Nome	Tipo de Porta	Grupo de Porta
SEQID	bigint	Somente Saída
Nome	string	Passagem
Cidade	string	Passagem
Designação	string	Passagem
Init3	integer	Variável

## Configuração da Transformação de Correspondência

Adicione uma transformação de Correspondência não reutilizável ao mapeamento para realizar a análise de identidade.

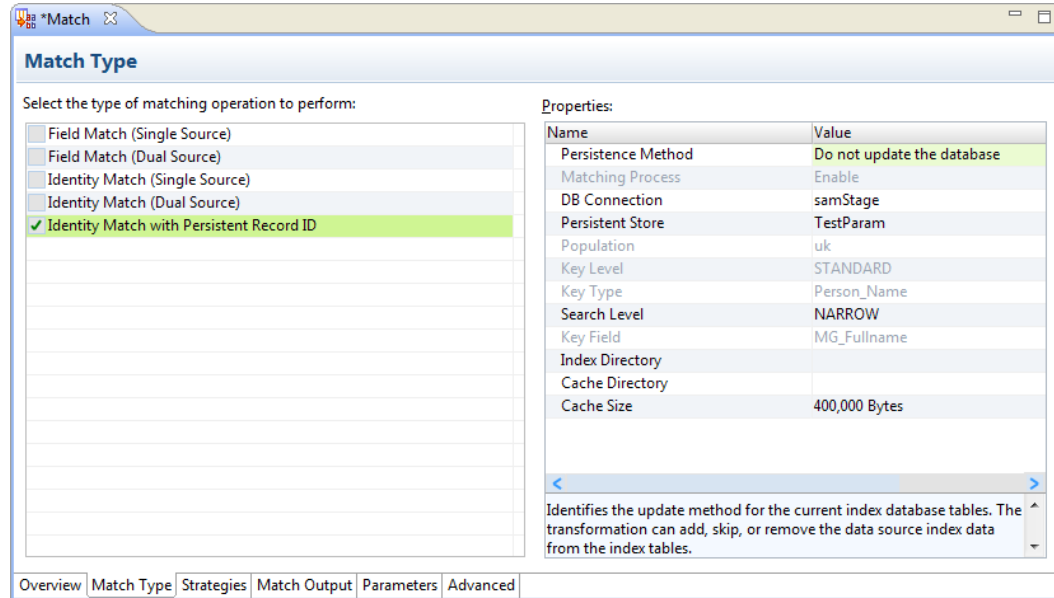
Realize as seguintes tarefas para configurar a transformação:

1. Selecione o tipo de análise de correspondência a ser realizada.
2. Conecte as portas de entrada à transformação.
3. Configure uma estratégia para comparar os dados de registros.
4. Selecione o tipo dos dados de saída de correspondência criados pela transformação.
5. Conecte as portas de saída a um destino de dados.

## Selecionar o Tipo de Operação de Correspondência

Use as opções na exibição Tipo de Correspondência para selecionar a operação de correspondência.

A seguinte imagem mostra a exibição Tipo de Correspondência:



Para comparar os dados de índice da fonte de dados com os dados de índice no conjunto de dados mestre, selecione **Correspondência de Identidade com ID de Registro Persistente**. Atualize o método de persistência para que a análise de correspondência não adicione dados às tabelas de índice. Você poderá determinar se as tabelas de índice serão atualizadas depois de revisar os resultados do mapeamento.

Use a opção **Conexão do DB** para identificar o banco de dados que contém os dados de índice. Use a opção **Armazenamento Persistente** para selecionar as tabelas de índice.

**Nota:** A transformação de Correspondência lê os valores das propriedades de preenchimento de identidade, nível de chave, tipo de chave e campo de chave dos metadados nas tabelas de banco de dados de índice. Os valores coincidem com as propriedades correspondentes na transformação que criou o armazenamento de dados de índice.

## Conectar as Portas de Entrada

Conecte as portas de dados de entrada à transformação. Verifique se os nomes de portas, a ordem das portas, os tipos de dados e as precisões devem corresponder às configurações de porta da transformação que criou o armazenamento de dados.

A transformação de Correspondência usa portas de entrada predefinidas para determinar a ordem na qual ela processa os registros. A transformação usa um identificador de sequência para rastrear os registros das portas de entrada até os pares correspondentes ou clusters que ela grava como saída. A transformação usa uma chave de grupo para classificar os registros processados.

Conecte as portas predefinidas às seguintes portas na transformação de Expressão:

- SequenceID. Conecte à porta SEQID da transformação de Expressão.
- GroupKey. Conectar à porta City da transformação de Expressão.

## Configurar uma Estratégia para Análise de Identidade

Use as opções na exibição Estratégias para configurar uma estratégia. A estratégia determina o tipo de análise que a transformação realiza nos dados de registros.

Selecione o algoritmo Person\_Name para os dados de registros. Selecione a porta de entrada Name para análise. Como a transformação cria cópias dos dados da porta, você seleciona as portas Name\_1 e Name\_2.

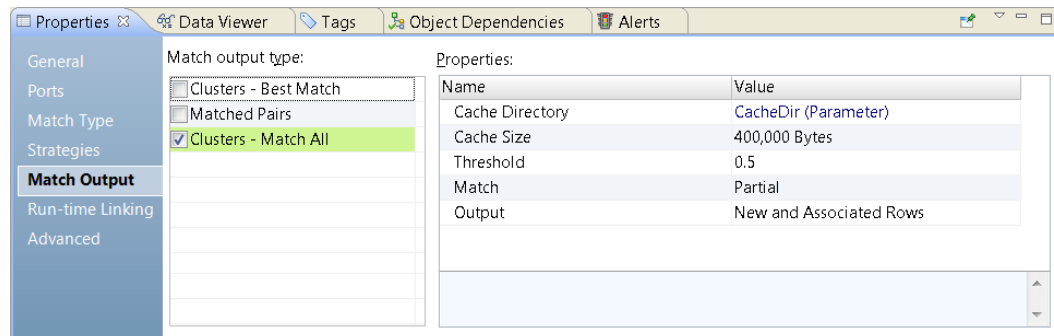
### TÓPICOS RELACIONADOS:

- [“Algoritmos de Correspondência de Identidade” na página 504](#)

## Selecionar o Tipo de Saída de Correspondência

Use as opções na exibição Saída de Correspondência para definir o formato de saída dos resultados da análise de correspondência

A seguinte imagem mostra a exibição Saída de Correspondência para a análise de identidade de origem única:



Você configura a transformação de forma a organizar os registros de saída em clusters. Cada cluster contém todos os registros que correspondem a pelo menos um outro registro, com base nas propriedades de correspondência que você especifica. As propriedades de correspondência determinam como a transformação compara os registros de fonte de dados com os registros de índice.

A seguinte tabela descreve as opções de propriedade de correspondência que você pode especificar para analisar os dados de registros da equipe:

Propriedade de Correspondência	Opção	Descrição da Opção
Correspondência	Parcial	A transformação compara os registros de fonte de dados com o armazenamento de dados de índice. A transformação também compara os registros dentro da fonte de dados entre si.
Saída	Linhas Novas e Associadas	A transformação grava cada cluster que contém pelo menos um registro proveniente da fonte de dados. Os clusters podem incluir registros do armazenamento de dados de índice.  Como um cluster pode conter um único registro, a saída contém todos os registros na fonte de dados.

### TÓPICOS RELACIONADOS:

- [“Configuração da Propriedade Correspondência” na página 508](#)
- [“Configuração da Propriedade Saída” na página 508](#)



## Conectar as Portas de Saída

Conecte as portas de saída da transformação de Correspondência ao destino de dados no mapeamento. Selecione as portas contendo os dados de registros que você deseja gravar no destino de dados.

A transformação inclui uma série de portas predefinidas para os dados clusterizados. Selecione as portas predefinidas que indicam o status duplicado dos registros e identifique a fonte de dados que armazena cada registro.

As seguintes portas contêm dados que você pode usar para localizar registros duplicados e determinar a origem ou os registros:

- A porta **ClusterSize** indica o número de registros em um cluster. Se um registro pertencer a um cluster cujo tamanho é maior que 1, a transformação considerará o registro como uma duplicata de outro registro.
- A porta **ClusterID** identifica o cluster ao qual um registro pertence. Use os dados em ClusterID para localizar os registros que são duplicatas do registro atual.
- A porta **PersistenceStatus** usa um valor de código para descrever o relacionamento entre os dados de índice da origem do mapeamento e os dados de índice no armazenamento de dados.
- A porta **PersistenceStatusDesc** retorna uma descrição de texto dos valores no código da porta PersistenceStatus.

É possível usar outras portas para analisar os relacionamentos entre os registros de cluster. Os valores de porta de vínculo e os valores de porta de driver indicam a extensão da semelhança entre os registros em cada cluster.

No exemplo atual, você conecta todas as portas ao destino de dados. Para exibir os dados de saída nessas portas, execute o Visualizador de Dados.

## TÓPICOS RELACIONADOS:

- [“Códigos de Status de Persistência e Descrições de Status de Persistência” na página 478](#)
- [“Valores do Código de Status e Valores de Descrição de Status” na página 479](#)

## Executar o Visualizador de Dados

Execute o Visualizador de Dados para analisar os resultados da análise de correspondência. Por padrão, o Visualizador de Dados mostra todas as portas de saída na transformação de Correspondência. Ao executar o mapeamento, você atualiza o destino de dados com os dados das portas de saída.

A seguinte imagem mostra os dados de saída no Visualizador de Dados:

	Name	ClusterSize	PersistenceStatusDesc	Designation	GroupKey	LinkId	ClusterId	LinkScore	PersistenceSt...
1	Chaith	2	Input, Exists, Ignored	SE	Bangalore	S - 1	000000001	1	IEI00000
2	Chaith	2	Store, -, No change	SE	Bangalore	1 - 1267	000000001	1	SON00000
3	Ramana	2	Input, Exists, Ignored	SSE	Chennai	S - 2	000000002	1	IEI00000
4	Ramana	2	Store, -, No change	SSE	Chennai	1 - 2534	000000002	1	SON00000
5	Ramesh	4	Input, Exists, Ignored	SSE	Chennai	S - 3	000000003	1	IEI00000
6	Ramesh	4	Input, Exists, Ignored	Lead	Chennai	S - 3	000000003	1	IEI00000
7	Ramesh	4	Store, -, No change	SSE	Chennai	1 - 3801	000000003	1	SON00000
8	Ramesh	4	Store, -, No change	Lead	Chennai	1 - 3801	000000003	1	SON00000
9	Sunil	2	Input, Exists, Ignored	Principal	Bangalore	S - 5	000000004	1	IEI00000
10	Sunil	2	Store, -, No change	Principal	Bangalore	1 - 6335	000000004	1	SON00000
11	Venu	2	Input, Exists, Ignored	Principal	Hydrabad	S - 6	000000005	1	IEI00000
12	Venu	2	Store, -, No change	Principal	Hydrabad	1 - 7602	000000005	1	SON00000
13	Harish	2	Input, Exists, Ignored	SE	Bangalore	S - 7	000000006	1	IEI00000
14	Harish	2	Store, -, No change	SE	Bangalore	1 - 8869	000000006	1	SON00000
15	Sachin	2	Input, Exists, Ignored	SSE	Bangalore	S - 8	000000007	1	IEI00000
16	Sachin	2	Store, -, No change	SSE	Bangalore	1 - 10136	000000007	1	SON00000

Row 1 to 16

O Visualizador de Dados verifica se o arquivo contém registros que duplicam os dados no conjunto de dados mestre.

Considere os seguintes valores de dados no Visualizador de Dados:

- Cada registro do conjunto de dados pertence a um cluster que contém dois ou mais registros. Portanto, cada registro é uma duplicata de pelo menos um outro registro. A transformação atribui um tamanho de cluster de 1 a registros exclusivos. A fonte de dados não contém registros que o conjunto de dados mestre não contém.
- Os dados de **PersistenceStatusDesc** identificam a origem de registro e indicam se a transformação de Correspondência adiciona o registro às tabelas de índice. A coluna indica que cada registro de entrada existe no conjunto de dados mestre. A transformação não adiciona dados ao índice de dados mestre.

## Conclusão

Os resultados da análise de correspondência indicam que o arquivo de registros da equipe não contém registros que o conjunto de dados mestre não contém. As descrições de status de persistência indicam que

o mapeamento não atualiza as tabelas de índice com os dados da fonte de dados. Você descarta o arquivo de registros da equipe.

Quando receber outra atualização dos escritórios regionais, você poderá criar outro arquivo e compará-lo ao conjunto de dados mestre. É possível reutilizar o mapeamento e as tabelas de índice. Como você armazena os dados de índice do conjunto de dados mestre nas tabelas de banco de dados, não precisa gerar novamente os dados de índice.

## CAPÍTULO 31

# Transformação Normalizador

Este capítulo inclui os seguintes tópicos:

- [Visão Geral de Transformação do Normalizador, 520](#)
- [Campos de Ocorrência Múltipla, 521](#)
- [Registros de Ocorrência Múltipla, 522](#)
- [Definição de Hierarquia de Entrada, 523](#)
- [Grupos e Portas de Saída de transformação de Normalizador, 531](#)
- [Geração de Chave para Grupos de Saída, 535](#)
- [Propriedades Avançadas da Transformação de Normalizador, 535](#)
- [Criando uma Transformação de Normalizador, 536](#)
- [Criando uma Transformação de Normalizador de uma Origem Upstream, 537](#)
- [Exemplo de Mapeamento de Normalizador, 537](#)
- [Transformação de Normalizador em um ambiente não nativo, 541](#)

## Visão Geral de Transformação do Normalizador

A transformação de Normalizador é uma transformação ativa que transforma uma linha de origem em várias linhas de destino. Quando o Normalizador recebe uma linha que contém dados de ocorrência múltipla, ele retorna uma linha para cada instância dos dados de ocorrência múltipla.

A transformação de Normalizador analisa os dados de ocorrência múltipla em linhas de saída separadas. Por exemplo, uma linha de origem relacional pode conter quatro trimestres de vendas. A transformação de Normalizador gera uma linha de saída separada para cada ocorrência de vendas.

Ao definir a transformação de Normalizador, você configura uma hierarquia de linha de entrada que descreve a estrutura dos dados de origem. Opcionalmente, você pode definir registros na hierarquia de entrada da transformação. Um registro é um contêiner para um grupo de campos. Defina um registro quando um grupo de campos ocorrer várias vezes nos dados de origem. A hierarquia de entrada determina como você pode configurar os grupos de saída de transformação.

A transformação de Normalizador transforma dados de tabelas relacionais ou de origens de arquivo simples.

## Campos de Ocorrência Múltipla

Quando um campo for repetido várias vezes na fonte de dados, você poderá defini-lo como um campo de ocorrência múltipla na hierarquia de linha de entrada. A transformação do Normalizador pode retornar uma linha separada para cada ocorrência de um campo de ocorrência múltipla ou um grupo de campos em uma origem.

Uma linha de origem pode conter quatro trimestres de vendas por loja:

Loja	Vendas (1)	Vendas (2)	Vendas (3)	Vendas (4)
Loja1	100	300	500	700
Loja2	250	450	650	850

Quando você define a hierarquia de entrada do Normalizador, pode combinar os quatro campos Vendas em um campo de ocorrência múltipla. Defina um nome de campo como Qtr\_Sales e configura-o para ocorrer quatro vezes na origem.

Quando o grupo de saída contém os dados de loja e os dados de vendas, a transformação do Normalizador retorna uma linha para cada combinação de Store e Qtr\_Sales. A linha de saída contém um índice que identifica qual instância de Qtr\_Sales está na linha de saída.

A transformação retorna as seguintes linhas:

Loja	Qtr_Sales	Trim (GCID)
Loja1	100	1
Loja1	300	2
Loja1	500	3
Loja1	700	4
Loja2	250	1
Loja2	450	2
Loja2	650	3
Loja2	850	4

Quando um grupo de saída contém colunas de ocorrência única e uma coluna de ocorrência múltipla, o Normalizador retorna dados duplicados para as colunas de ocorrência única em cada linha de saída. Por exemplo, Store1 e Store2 se repetem para cada instância de Qtr\_Sales.

Uma linha de origem pode conter mais de um nível de dados de ocorrência múltipla. Você pode configurar a transformação de Normalizador para retornar linhas separadas em cada nível com base no modo como você define a hierarquia de entrada.

## ID da Coluna Gerada

A transformação de Normalizador retorna uma porta de saída da ID da coluna gerada (GCID) para cada instância de um campo de ocorrência múltipla.

A porta de ID de coluna gerada tem um índice para a instância dos dados de ocorrência múltipla. Por exemplo, se um campo ocorrer quatro vezes em um registro de origem, a ferramenta Developer retornará o valor 1, 2, 3 ou 4 na porta de ID de coluna gerada com base em qual instância dos dados de ocorrência múltipla ocorre na linha.

# Registros de Ocorrência Múltipla

Você pode definir registros de ocorrência múltipla nos dados de origem da transformação de Normalizador. Os registros são grupos de campos. Defina os registros na transformação de Normalizador quando você precisar definir os grupos de campos de origem que são de ocorrência múltipla.

## Exemplo de Registros de Ocorrência Múltipla

A seguinte linha do Cliente contém informações do cliente com informações de endereço residencial e comercial:

```
CustomerID
FirstName
LastName
Home_Street
Home_City
Home_State
Home_Country
Business_Street
Business_City
Business_State
Business_Country
```

Ao configurar a transformação de Normalizador, você pode definir uma estrutura de entrada que contenha os campos do cliente e um registro de endereço de ocorrência múltipla. O registro de endereço ocorre duas vezes. Ao configurar os grupos de saída de transformação de Normalizador, você pode retornar o registro de Endereço para um destino diferente dos campos CustomerID, FirstName e LastName.

O seguinte exemplo mostra uma estrutura de entrada com um registro de endereço de ocorrência múltipla:

```
CustomerID
FirstName
LastName
Address (occurs twice)
    Street
    City
    State
    Country
```

Subregistros são registros em registros. Ao definir os registros e os subregistros, você define uma hierarquia de entrada de campos na linha de origem. Cada registro é um nó em uma hierarquia que você pode referenciar ao definir a saída de transformação.

Por exemplo, a linha de origem pode conter vários números de telefone para cada tipo de endereço:

```
CustomerID
FirstName
LastName
Home_Street
Home_City
Home_State
Home_Country
Telephone_No
Cell_Phone_No
Alternate_Phone_No
Business_Street
Business_City
Business_State
Business_Country
Business_Telephone_No
Business_Cell_Phone_No
Business-Alternate_Phone1
```

Você define uma hierarquia de entrada onde o Endereço é o pai do Telefone. Ao definir a saída de transformação de Normalizador, você pode retornar os endereços e os números de telefone para destinos separados das informações do cliente.

Defina uma hierarquia de entrada semelhante ao seguinte exemplo:

```
CustomerID
FirstName
LastName
Address (occurs twice)
  Street
  City
  State
  Country
  Phone
  Telephone_No (occurs three times)
```

## Definição de Hierarquia de Entrada

Ao criar uma transformação de Normalizador, você define uma hierarquia de entrada que descreve os registros e os campos na origem. Defina a hierarquia de entrada na exibição **Normalizador** da transformação.

A Developer tool cria portas de entrada de transformação com base nos campos definidos na hierarquia de entrada. Defina a estrutura do grupo de entrada antes de definir os grupos de saída de transformação.

Ao definir uma hierarquia de entrada, você deve definir uma estrutura de entrada que corresponda à estrutura dos dados de origem. Os dados de origem podem conter mais de um grupo de campos de ocorrência múltipla. Para definir a estrutura, você pode configurar um registro que ocorra no mesmo nível de outro registro na origem. Ou você pode definir registros que ocorram em outros registros.

### Exemplo de Hierarquia de Entrada

A seguinte linha de origem contém campos do cliente e um registro de endereço que ocorre duas vezes:

```
CustomerID
FirstName
LastName
Address
  Street
  City
  State
  Country
Address1
  Street1
  City1
  State1
  Country1
```

Ao definir a estrutura de entrada na exibição **Normalizador**, você pode adicionar CustomerID, FirstName e LastName como campos. Defina um registro Endereço e inclua os campos Rua, Cidade e Estado no endereço. Altere o valor de Ocorrência do Endereço para 2.

A seguinte imagem mostra a hierarquia de entrada na exibição **Normalizador**:

Normalizer					
Name	Level	Occurs	Type	Precision	Scale
CustomerID	1	1	string	10	0
FirstName	1	1	string	10	0
LastName	1	1	string	10	0
Address	1	2			
Street	2	1	string	10	0
City	2	1	string	10	0
State	2	1	string	10	0
Country	2	1	string	10	0

A coluna **Ocorrência** na exibição **Normalizador** identifica o número de instâncias de um campo ou de um registro em uma linha de origem. Altere o valor na coluna **Ocorrência** para registros ou campos de ocorrência múltipla. Neste exemplo, os campos de cliente ocorrem uma vez e o registro de Endereço ocorre duas vezes.

A coluna **Nível** na exibição **Normalizador** indica onde um campo ou um registro é exibido na hierarquia de entrada. Os campos de cliente estão no nível 1 na hierarquia. O registro de Endereço também está no nível 1.

## Transformação Normalizador Portas de Entrada

A ferramenta Developer cria as portas de entrada da transformação de Normalizador quando você define a hierarquia de entrada na exibição **Normalizador**. Quando você altera os campos na hierarquia de entrada, a ferramenta Developer altera as portas de entrada.

Exiba as portas de entrada da transformação de Normalizador na exibição **Visão Geral**. Você pode reordenar as portas de entrada na exibição **Visão Geral**. Para alterar as portas de entrada, atualize a hierarquia de entrada na exibição **Normalizador**.

Quando você define um campo como ocorrência múltipla na hierarquia de entrada, a ferramenta Developer cria uma porta de entrada para cada instância do campo de ocorrência múltipla. Quando um registro é de ocorrência múltipla, a ferramenta Developer cria uma porta de entrada para cada instância dos campos no registro.

### Exemplo de Portas de Entrada

A seguinte imagem mostra as portas de entrada que a ferramenta Developer cria para os dados de cliente e os dados de endereço de ocorrência múltipla:



Ports					
	Name	Type	Precision	Scale	Location
	Input				
1	CustomerID	string	10	0	CustomerID
2	FirstName	string	10	0	FirstName
3	LastName	string	10	0	LastName
4	Street	string	10	0	Address.Street
5	Street1	string	10	0	Address.Street
6	City	string	10	0	Address.City
7	City1	string	10	0	Address.City
8	State	string	10	0	Address.State
9	State1	string	10	0	Address.State
10	Country	string	10	0	Address.Country
11	Country1	string	10	0	Address.Country

## Mesclar Campos

Você pode mesclar os campos de dados semelhantes em um único campo de ocorrência múltipla na exibição **Normalizador**. Talvez você precise mesclar os campos quando arrastar portas de outro objeto para criar uma transformação de Normalizador em um mapeamento.

Uma linha de origem pode conter vários campos com diferentes tipos de dados de salário, como Base\_Salary, Bonus\_Pay e Sales\_Commissions. Você pode mesclar os campos para criar um campo de Salário que ocorra três vezes.

A seguinte imagem mostra uma linha de funcionário com três tipos de salários selecionados na exibição Normalizador:

Normalizer					
Name	Level	Occurs	Type	Preci...	Scale
EmployeeID	1	1	string	10	0
Base_Salary	1	1	decimal	10	0
Bonus_Pay	1	1	decimal	10	0
Sales Commissions	1	1	decimal	10	0

Você pode mesclar os três tipos de dados de salário em um campo de Salário que ocorra três vezes.

A seguinte imagem mostra o campo de Salário:

Normalizer						
Name	Level	Occurs	Type	Preci...	Scale	
EmployeeID	1	1	string	10	0	
Salary	1	3	decimal	10	0	

## Mesclando Campos

Mesclando os campos do mesmo tipo em um campo de ocorrência múltipla na exibição Normalizador.

1. Clique na exibição **Normalizador**.
2. Selecione os campos que você deseja mesclar.
3. Clique no botão **Mesclar**.  
A caixa de diálogo **Mesclar Campos** é exibida.
4. Insira um nome para o campo de mesclagem, o tipo, a precisão, a escala e a ocorrência do campo de ocorrência múltipla.
5. Clique em **OK**.

## Nivelar campos

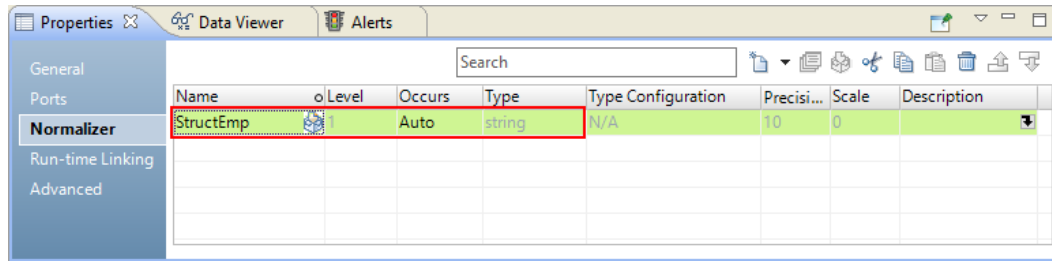
Você pode nivelar campos de um tipo de dados complex em mapeamentos executados no Mecanismo Spark. Você nivela campos na exibição **Normalizador** para modificar dados hierárquicos que passam por uma porta complexa.

A saída da ação de nivelamento depende do tipo de dados complex. Quando você nivela um tipo de dados array ou struct, a transformação de Normalizador cria uma linha para cada elemento no tipo de dados complex. Quando você nivela um tipo de dados map, a transformação de Normalizador cria duas colunas para os elementos chave de mapa e valor de mapa.

A ação de nivelamento em um tipo de dados aninhado extrai elementos no primeiro nível. Para nivelar um tipo de dados aninhado em todos os níveis, use o assistente de conversão hierárquica **Porta Complexa de Nivelamento** na Developer tool. A opção **Nivelar Tudo** extrai elementos em cada nível e retorna dados relacionais do tipo de dados primitivo. Para obter mais informações sobre os assistentes de conversão hierárquica, consulte o *Guia do Usuário do Data Engineering Integration*.

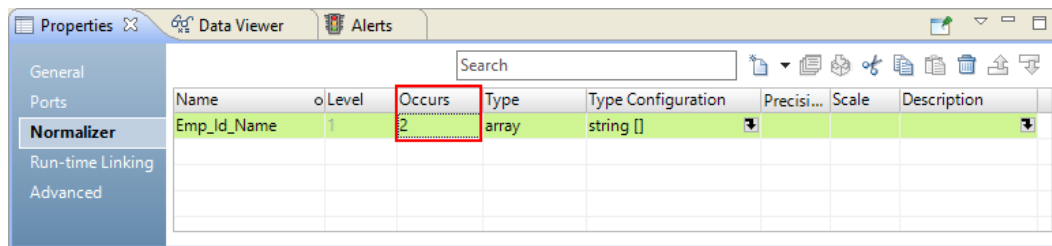
A ação de nivelamento altera o valor da coluna Ocorre na exibição Normalizador para Automático e adiciona um ícone de nivelamento ao lado do campo nivelado. O valor Automático indica que a transformação nivela todos os elementos do tipo de dados complex.

A imagem a seguir mostra uma estrutura que é nivelada em um campo de cadeia com um ícone de nivelamento ao lado e o valor Ocorre como Automático:



Você não pode nivelar um campo de múltiplas ocorrências. Por exemplo, você não pode nivelar um campo de matriz com valor Ocorre como 2.

A imagem a seguir mostra um campo de múltiplas ocorrências de um tipo de dados array que você não pode nivelar:



## Nivelar matriz

A transformação de Normalizador nivela uma matriz unidimensional para um tipo de dados primitivo e uma matriz n-dimensional para uma matriz (n-1)-dimensional. O número de linhas que a transformação cria é o mesmo que o tamanho da matriz.

Por exemplo, se você nivelar uma porta de matriz com 10 elementos de cadeia, a saída retornará 10 portas de cadeia. Se você nivelar uma matriz tridimensional, a saída retornará uma matriz bidimensional.

Uma tabela contém um Nome de porta de cadeia e Telefones de porta de matriz. Você deseja nivelar a porta de matriz. A tabela contém os seguintes valores:

Name	Phones
Adams	[205-128-6478, 722-515-2889, 650-213-4020]
Jane	[650-321-4506]

Quando você nivela a porta de matriz, a saída é a seguinte:

Name	Phones	GCID_Phones
Adams	205-128-6478	1
Adams	722-515-2889	2

Name	Phones	GCID_Phones
Adams	650-213-4020	3
Jane	650-321-4506	1

Você pode editar o valor Ocorre de um campo nivelado para extrair um número específico de elementos na matriz. O valor deve ser um inteiro positivo maior que 1. O valor determina o número de elementos a serem extraídos. Por exemplo, você pode alterar o valor Ocorre para 2 para extrair os dois primeiros elementos da matriz. A saída é a seguinte:

Name	Phones	GCID_Phones
Adams	205-128-6478	1
Adams	722-515-2889	2
Jane	650-321-4506	1

## Nivelar estrutura

A transformação nivela uma estrutura em um campo do tipo de dados dos elementos na estrutura. Para nivelar um tipo de dados struct, todos os elementos de estrutura devem ser do mesmo tipo de dados. A transformação cria uma linha para cada elemento no tipo de dados struct.

Por exemplo, você deseja nivelar o seguinte campo de estrutura:

```
customer_address{
  city : string
  state : string
  zip : string
}
```

A tabela contém os seguintes valores:

Name	customer_address
Clara	{       New York       NY       10032     }

Quando você nivela a porta de estrutura, a saída é a seguinte:

Name	customer_address	GCID_customer_address
Clara	New York	1
Clara	NY	2
Clara	10032	3

Quando os elementos de estrutura são de tipos de dados diferentes e pelo menos os dois primeiros elementos são do mesmo tipo de dados, você pode nivelar os dados da estrutura para elementos consecutivos do mesmo tipo de dados. Para extrair elementos de estrutura consecutivos do mesmo tipo de

dados, edite o valor Ocorre. O valor deve ser um inteiro positivo maior que 1. Por exemplo, um struct emp\_address contém os seguintes elementos:

```
emp_address{
  city : string
  state : string
  zip : int
  country : string
}
```

Você pode definir o valor de Ocorre para 2 para extrair elementos de estrutura da cidade e do estado. Se você definir o valor como 3 ou 4, a validação do mapeamento falhará.

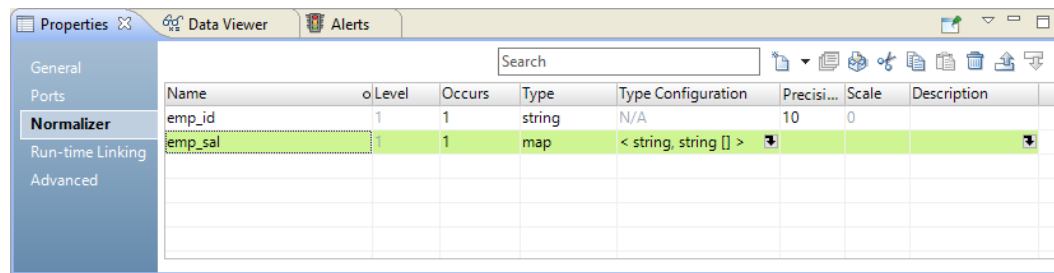
## Nivelar mapa

A transformação nivela um mapa em dois campos para os elementos chave e valor no mapa. Para um campo de mapa que você nivelou, não é possível alterar o valor de Ocorre de Automático para um valor inteiro.

Por exemplo, você deseja nivelar o seguinte campo de mapa emp\_sal com uma chave de cadeia e uma matriz de valores inteiros:

```
<emp_name -> [base_sal, bonus, commision]>
```

A imagem a seguir mostra o campo do mapa que você deseja nivelar na exibição Normalizador:



Name	Level	Occurs	Type	Type Configuration	Precisi...	Scale	Description
emp_id	1	1	string	N/A	10	0	
emp_sal	1	1	map	< string, string [] >			

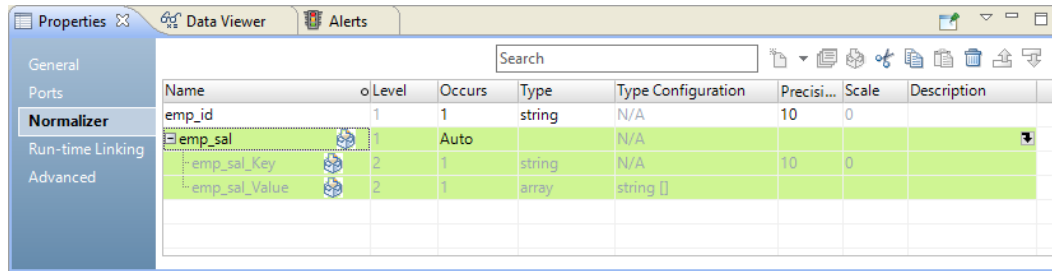
A tabela contém os seguintes valores:

emp_id	emp_sal
12200	<Greg -> [4000, 1000, 500]>
12201	<Patricia -> [3800, 1500, 1000]>

Quando você nivela a porta do mapa, a saída retorna um campo de cadeia para a chave de mapa e um campo da matriz para o valor de mapa da seguinte forma:

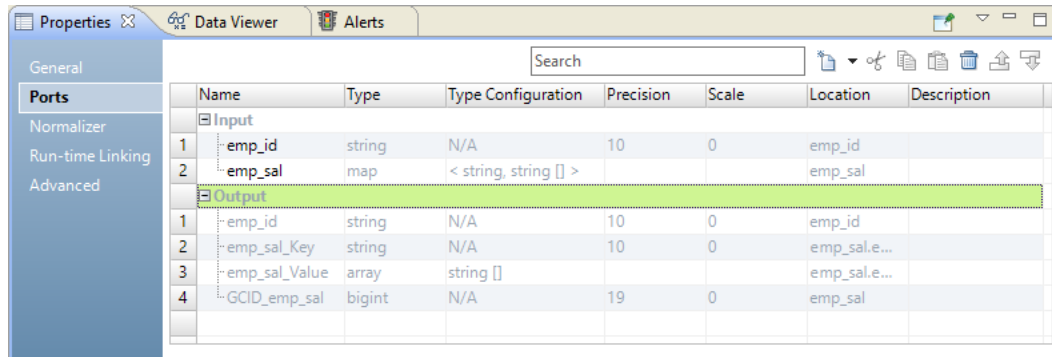
emp_id	emp_sal_Key	emp_sal_Value	GCID_emp_salary
12200	Greg	[4000, 1000, 500]	1
12201	Patricia	[3800, 1500, 1000]	1

A imagem a seguir mostra o campo de mapa que é nivelado em um campo de chave de cadeia e um campo de valor de matriz na exibição Normalizador:



Name	o Level	Occurs	Type	Type Configuration	Precisi...	Scale	Description
emp_id	1	1	string	N/A	10	0	
emp_sal	1	Auto	N/A	N/A			
emp_sal_Key	2	1	string	N/A	10	0	
emp_sal_Value	2	1	array	string []			

A imagem a seguir mostra o grupo Saída na exibição Portas:



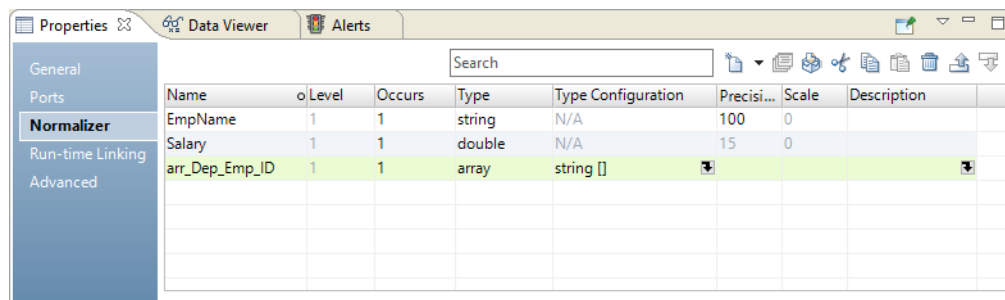
Name	Type	Type Configuration	Precision	Scale	Location	Description
<b>Input</b>						
1 emp_id	string	N/A	10	0	emp_id	
2 emp_sal	map	< string, string [] >			emp_sal	
<b>Output</b>						
1 emp_id	string	N/A	10	0	emp_id	
2 emp_sal_Key	string	N/A	10	0	emp_sal.e...	
3 emp_sal_Value	array	string []			emp_sal.e...	
4 GCID_emp_sal	bigint	N/A	19	0	emp_sal	

## Nivelando campos

Nivele campos de um tipo de dados complex para modificar dados hierárquicos ou converter em dados relacionais.

1. Clique na exibição **Normalizador**.
2. Selecione o campo de um tipo de dados complex.

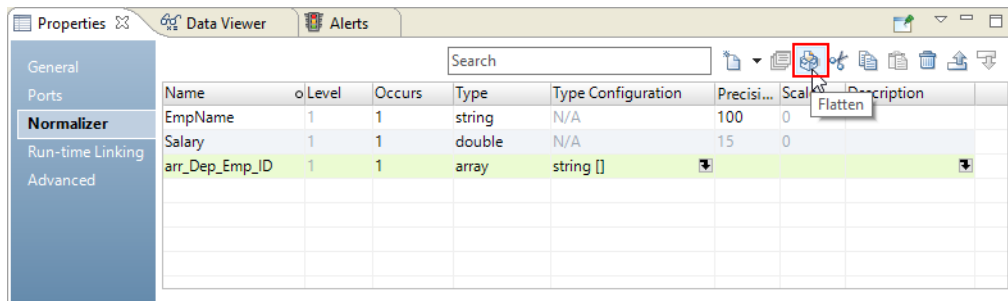
A imagem a seguir mostra um campo de matriz de tipo com elementos de cadeia:



Name	o Level	Occurs	Type	Type Configuration	Precisi...	Scale	Description
EmpName	1	1	string	N/A	100	0	
Salary	1	1	double	N/A	15	0	
arr_Dep_Emp_ID	1	1	array	string []			

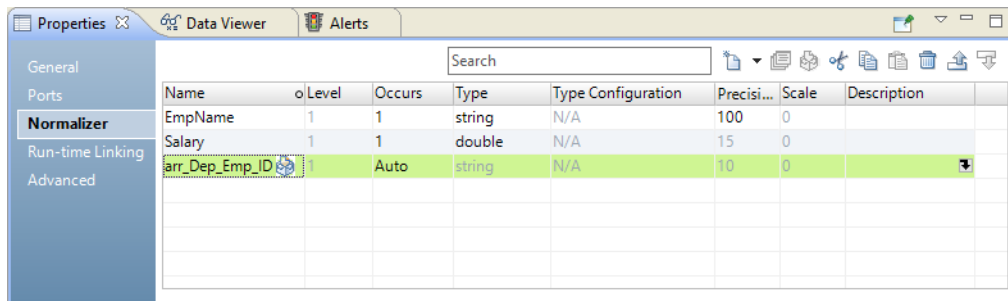
3. Clique no botão **Nivelar**.

A seguinte imagem mostra o botão Nivelar:

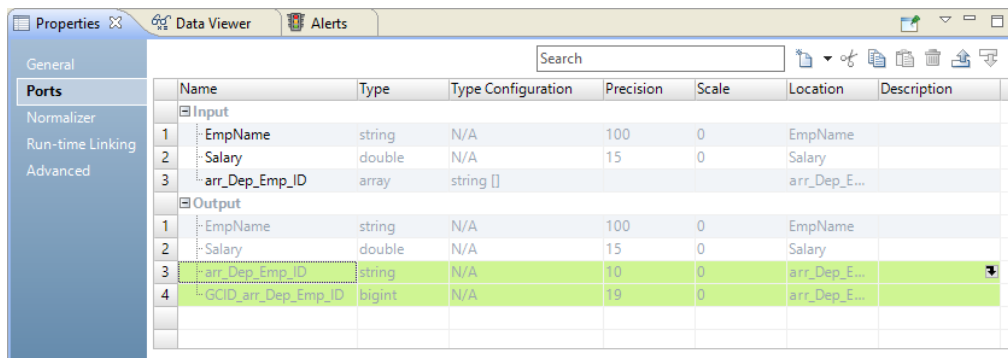


A ação de nivelamento substitui o campo de um tipo de dados complex por um campo nivelado e altera o valor de **Ocorre** para Automático. O tipo de dados do campo nivelado depende do tipo de dados complex que você nivela.

A imagem a seguir mostra o campo nivelado do tipo cadeia:



A imagem a seguir mostra a porta de saída da cadeia nivelada e a porta de saída do GCID na exibição **Portas**:



## Grupos e Portas de Saída de transformação de Normalizador

Defina os grupos e as portas de saída na exibição Visão Geral da transformação de Normalizador. Você pode definir os grupos de saída depois de definir a hierarquia de entrada da transformação.

A ferramenta Developer gera pelo menos um grupo de saída por padrão. O grupo de saída contém todos os campos de nível 1 das portas de entrada e o primeiro campo de ocorrência múltipla. Quando você define

mais de um campo de ocorrência múltipla na exibição do Normalizador, a Developer tool cria um grupo de saída para cada campo de ocorrência múltipla.

A seguinte linha de origem contém os dados de clientes, o campo de ocorrência múltipla de vendas e o campo de ocorrência múltipla de telefone:

```
CustomerID
LastName
FirstName
Sales (occurs 4 times)
Phone (occurs 3 times)
```

A ferramenta Developer cria dois grupos de saída da estrutura de entrada:

```
Output
CustomerID
LastName
FirstName
Sales

Output1
Phone
GCID_Phone
```

A ferramenta Developer cria um grupo Output1 porque os dados de origem contém mais de um campo de ocorrência múltipla. A ferramenta Developer não cria grupos para os campos que você definir nos registros. Quando você definir registros, defina os grupos de saída que contém os campos nos registros.

A seguinte linha de origem contém campos do cliente e um registro de endereço que ocorre duas vezes:

```
CustomerID
FirstName
LastName
Address
  Street
  City
  State
  Country
Address1
  Street1
  City1
  State1
  Country1
```

A ferramenta Developer cria um grupo de saída que contém os campos a seguir.

```
CustomerID
FirstName
LastName
```

A ferramenta Developer cria um grupo de saída padrão para os campos do cliente de nível 1. O grupo de saída padrão não inclui o registro de Endereço. Você deve configurar como os dados de Endereço são retornados na saída.

Crie os grupos de saída com base em como você precisa estruturar as linhas de saída. Quando a linha de origem contém dados de clientes e de endereço, você pode criar um grupo de saída para os campos de cliente. Você pode criar outro grupo de saída para os campos de endereço. Ou pode atualizar o grupo de saída padrão e adicionar os campos de endereço a ele. Os seguintes exemplos mostram diferentes resultados de saída com base na configuração do grupo de saída.

## Criar um Grupo de Saída

Crie um grupo de saída na exibição **Visão Geral** da transformação de Normalizador.

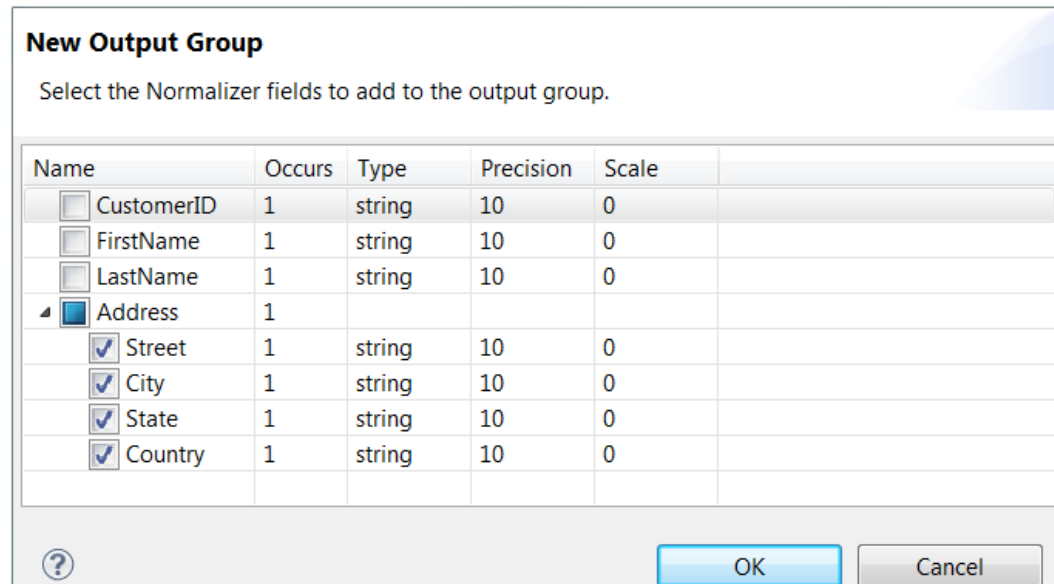
Quando você abre a exibição **Visão Geral** da transformação de Normalizador, a ferramenta Developer mostra os grupos padrão que ele cria da hierarquia de entrada.



Quando você cria um novo grupo de saída, uma caixa de diálogo mostra uma lista de campos e registros na hierarquia de entrada. Selecione os campos ou os registros a serem incluídos no grupo.

### Exemplo de Grupo de Saída

A seguinte imagem mostra a caixa de diálogo **Novo Grupo de Saída**:



Quando você seleciona o registro de Endereço, a ferramenta Developer cria um grupo de portas de saída que correspondem aos campos no registro de Endereço. O grupo Saída1 contém as portas Rua, Cidade, Estado e País. Você pode alterar as portas no grupo de saída.

A seguinte imagem mostra o grupo Saída e o grupo Saída1 na exibição **Visão Geral**:

	<b>Output</b>				
1	CustomerID	string	10	0	CustomerID
2	FirstName	string	10	0	FirstName
3	LastName	string	10	0	LastName
	<b>Output1</b>				
1	Street	string	10	0	Address.Street
2	State	string	10	0	Address.State
3	Country	string	10	0	Address.Country
4	GCID_Address	bigint	19	0	Address

Você pode configurar a transformação de Normalizador para retornar as linhas do grupo Saída para uma tabela Cliente.

A tabela Cliente recebe dados semelhantes às seguintes linhas:

```
100, Robert, Bold
200, James, Cowan
```

Você pode retornar as linhas do grupo Saída1 para uma tabela Endereço. A tabela Endereço recebe a Rua, a Cidade, o Estado, o País e o GCID.

A tabela Endereço recebe dados semelhantes às seguintes linhas:

```
100 Summit Dr, Redwood City, CA, United States,1
41 Industrial Way, San Carlos, CA, United States,2
```

```
85 McNulty Way, Los Angeles, CA, United States,1
55 Factory Street, Los Vegas, NV, United States,2
```

O GCID identifica qual instância do endereço do cliente está na linha de saída. Neste exemplo, a transformação de Normalizador retorna duas instâncias do registro de Endereço. Cada linha de saída contém o valor GCID de 1 ou 2.

## Atualizar um Grupo de Saída

Você pode atualizar um grupo de saída de transformação de Normalizador. Você pode adicionar ou remover os campos no grupo.

Por padrão, a ferramenta Developer cria grupos de saída de nível 1 quando você define uma hierarquia de entrada. A ferramenta Developer não inclui registros nos grupos. Você pode atualizar os grupos de saída padrão e adicionar registros a eles.

Para atualizar um grupo de saída, realce o nome do grupo e clique em **Novo > Atualizar Grupo**. A caixa de diálogo **Editar Grupo de Saída** mostra os campos na hierarquia de entrada. Escolha os campos a serem incluídos no grupo.

### Exemplo de Atualizar Grupo de Saída

No exemplo anterior, a ferramenta Developer criou um grupo de saída padrão com os campos CustomerID, FirstName e LastName.

A seguinte imagem mostra o grupo de saída padrão:

	Output				
1	CustomerID	string	10	0	CustomerID
2	FirstName	string	10	0	FirstName
3	LastName	string	10	0	LastName

Você pode atualizar o grupo de saída padrão e adicionar o registro de Endereço a ele.

A seguinte imagem mostra a caixa de diálogo **Editar Grupo de Saída**:

**Edit Output Group**  
Select the Normalizer fields to include in the output group.

Name	Occurs	Type	Precision	Scale
<input checked="" type="checkbox"/> CustomerID	1	string	10	0
<input checked="" type="checkbox"/> FirstName	1	string	10	0
<input checked="" type="checkbox"/> LastName	1	string	10	0
<input checked="" type="checkbox"/> Address	2			
<input checked="" type="checkbox"/> Street	1	string	10	0
<input checked="" type="checkbox"/> State	1	string	10	0
<input checked="" type="checkbox"/> Country	1	string	10	0

Nesse exemplo, o CustomerID, o FirstName e o LastName são nós de nível 1. O registro de Endereço também é um nó de nível 1. A transformação de Normalizador pode retornar o Endereço na mesma linha que os dados dos clientes. Como o Endereço é de ocorrência múltipla, a ferramenta Developer adiciona o índice GCID\_Address ao grupo de saída.

A seguinte imagem mostra as portas no grupo de Saída:

	Output				
1	CustomerID	string	10	0	CustomerID
2	FirstName	string	10	0	FirstName
3	LastName	string	10	0	LastName
4	Street	string	10	0	Address.Street
5	State	string	10	0	Address.State
6	Country	string	10	0	Address.Country
7	GCID_Address	bigint	19	0	Address

Quando os campos de cliente e os campos de endereço de ocorrência múltipla estão no grupo de saída, a transformação de Normalizador retorna os mesmos campos de cliente para cada instância dos dados de endereço.

O seguinte exemplo mostra as linhas que a transformação de Normalizador gera no grupo de saída:

```
100, Robert, Bold, 100 Summit Dr, Redwood City, CA, United States,1
100, Robert, Bold, 41 Industrial Way, San Carlos, CA, United States,2
200, James, Cowan, 85 McNulty Way, Los Angeles, CA, United States,1
200, James, Cowan, 55 Factory Street, Los Vegas, NV, United States,2
```

A porta GCID contém o número da instância do Endereço. O valor GCID é 1 ou 2.

## Geração de Chave para Grupos de Saída

Você pode configurar uma transformação de Gerador de Sequência para gerar chaves que vinculam cada linha de saída que a transformação de Normalizador retorna da mesma linha de origem.

Você pode adicionar uma transformação de Gerador de Sequência antes da transformação de Normalizador em um mapeamento. A transformação de Gerador de Sequência adiciona um número de sequência a cada linha de origem. Quando a transformação de Normalizador retorna vários grupos de saída ou linhas da mesma linha de origem, cada linha de saída recebe o mesmo número de sequência. Você pode usar o número como uma chave em um relacionamento de chave primária-chave externa entre as tabelas de destino.

Por exemplo, a transformação de Normalizador retorna informações de cliente em um grupo de saída e informações de pedido em outro grupo de saída. Você pode usar o número de sequência para vincular as informações de cliente em uma tabela às informações de pedido em outra tabela.

## Propriedades Avançadas da Transformação de Normalizador

Configure as propriedades da transformação de Normalizador na guia **Avançado**.

Configure as seguintes propriedades na guia **Avançado**:

### Nível de Rastreamento

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.

## Gerar Grupos de Saída de Primeiro Nível

A ferramenta Developer gera, por padrão, o grupo de saída de primeiro nível da transformação do Normalizador. A ferramenta Developer gera outros grupos de saída quando você define mais de um campo de ocorrência múltipla na exibição do Normalizador.

A ferramenta Developer cria um grupo de saída que contém todos os campos de ocorrência única de nível 1 e o primeiro campo de ocorrência múltipla de nível 1. A ferramenta Developer cria um grupo de saída adicional quando a transformação contém mais de um campo de ocorrência múltipla.

A ferramenta Developer não cria um grupo de saída para registros de nível 1. Quando você definir um registro de nível 1, configure onde deseja colocá-lo na saída.

Para impedir a ferramenta Developer de criar grupos de saída para campos de ocorrência múltipla, desative a propriedade avançada **Gerar Automaticamente Grupos de Saída de Primeiro Nível**.

## Criando uma Transformação de Normalizador

Você pode criar uma transformação de Normalizador reutilizável ou não reutilizável. Transformações reutilizáveis podem existir em vários mapeamentos. As transformações não reutilizáveis existem dentro de um único mapeamento.

1. Para criar uma transformação, use um dos métodos a seguir:

Opção	Descrição
Reutilizável	Selecione um projeto ou pasta na exibição <b>Object Explorer</b> . Clique em <b>Arquivo &gt; Novo &gt; Transformação</b> . Selecione a transformação de Normalizador e clique em <b>Avançar</b> .
Não reutilizável	Em um mapeamento ou em um mapplet, arraste uma transformação de Normalizador da paleta <b>Transformação</b> até o editor. Você pode arrastar portas dos objetos de dados de origem ou das transformações no mapeamento para definir a transformação.

O assistente de **Nova Transformação de Normalizador** é exibido.

2. Insira um nome para a transformação.
3. Clique em **Avançar**.  
A página **Definição do Normalizador** é exibida.
4. Para adicionar um registro, clique no botão **Novo** e selecione **Registro**.
5. Para adicionar um campo, clique no botão **Novo** e selecione **Campo**.  
Para adicionar um campo a um registro, você deve selecionar o registro antes de adicionar o campo.
6. Opcionalmente, clique duas vezes o valor da coluna **Ocorre** para alterar a ocorrência de um campo ou registro.
7. Clique em **Avançar**.  
A página **Portas do Normalizador** é exibida.
8. Para adicionar um grupo de saída, clique no botão **Novo** e selecione **Novo Grupo de Saída**.
9. Para editar um grupo de saída, selecione o grupo de saída que você deseja editar. Clique no botão **Novo** e selecione **Editar Grupo de Saída**.

10. Clique em **Concluir**.

A transformação aparece no editor.

## Criando uma Transformação de Normalizador de uma Origem Upstream

Você pode criar uma transformação de Normalizador vazia e arrastar as portas de um objeto de dados de origem ou de uma transformação até a transformação de Normalizador para criar entradas e portas.

1. Crie um mapeamento que inclua a origem ou a transformação para transmitir os dados de origem para a transformação de Normalizador.
2. Para criar a transformação de Normalizador, selecione-a na paleta Transformação e arraste-a até o editor.

A caixa de diálogo Normalizador é exibida.

3. Clique em **Concluir** para criar uma transformação vazia.
4. Selecione portas de uma origem ou uma transformação no mapeamento e arraste-as até a transformação de Normalizador.

As portas de entrada e de saída são exibidas na transformação de Normalizador. A ferramenta Developer cria um grupo de entrada e um grupo de saída.

5. Abra a exibição **Normalizador** para atualizar o grupo padrão e organizar os campos em registros conforme necessário.
6. Para mesclar vários campos em um único campo de ocorrência múltipla, selecione os campos na exibição **Normalizador** e clique na opção **Mesclar**.

Escolha um nome para o campo de ocorrência múltipla.

## Exemplo de Mapeamento de Normalizador

Uma organização de varejo recebe os totais de vendas das lojas na organização. A organização recebe uma linha de dados que contém as informações de loja e quatro valores de vendas. Cada valor de vendas representa o total de vendas de um trimestre do ano.

O seguinte exemplo mostra como definir uma transformação de Normalizador para retornar os dados de vendas para um destino de Loja e de Vendas. O destino de Loja recebe uma linha para cada loja. O destino de Vendas recebe quatro linhas de cada loja. Cada linha contém um trimestre de dados de vendas.

Uma transformação de Gerador de Sequência gera uma ID exclusiva para cada loja. A transformação de Normalizador retorna a StoreID com cada linha de saída.

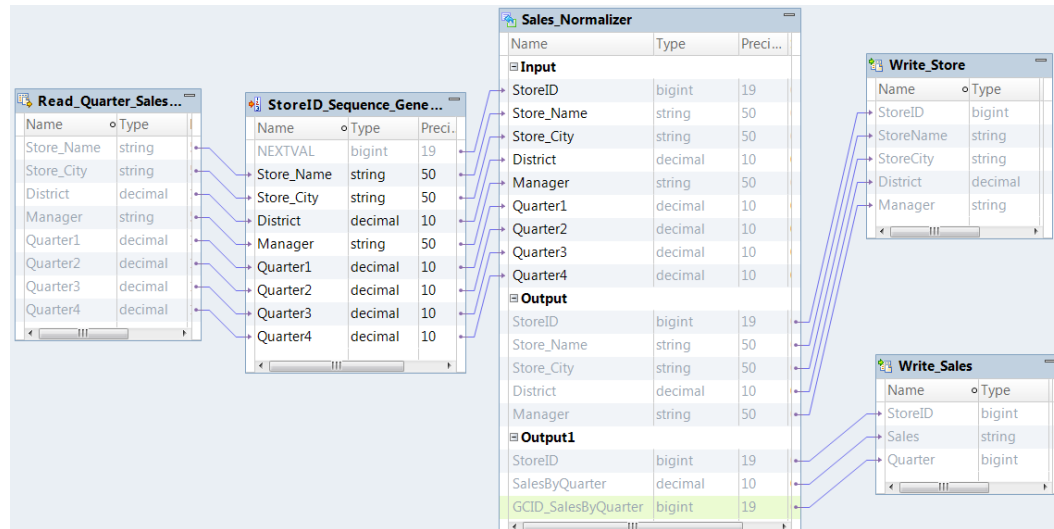
Crie um mapeamento com uma transformação de Leitura, uma transformação de Gerador de Sequência, a transformação do Normalizador e duas transformações de Gravação.

## Exemplo de Mapeamento do Normalizador

Crie um mapeamento que contém a transformação de Normalizador para normalizar várias ocorrências de dados de vendas trimestrais de uma origem de arquivo simples.

A transformação de Normalizador gera uma linha de saída separada para cada trimestre de vendas e grava os valores de vendas normalizados em um destino de Vendas. A transformação de Normalizador grava as informações de Loja em um destino de Loja.

A figura a seguir mostra o mapeamento da transformação de Normalizador:



O mapeamento contém os seguintes objetos:

### Read\_STORE

Uma origem de dados que contém campos recorrentes.

### Transformação de Gerador de Sequência StoreID

Uma transformação de Gerador de Sequência que gera uma chave storeID para vincular a tabela Loja à tabela Vendas.

### Sales\_Normalizer

Uma transformação de Normalizador que normaliza os dados de vendas de ocorrência múltipla.

### Write\_Store

Um destino que recebe as informações de loja da transformação de Normalizador.

### Write\_Sales

Um destino que recebe os números de vendas da transformação de Normalizador.

## Definição de Exemplo de Normalizador

A origem é um arquivo simples que contém informações de armazenamento e os dados de vendas trimestrais. Defina a estrutura dos dados de origem na exibição **Normalizador**.

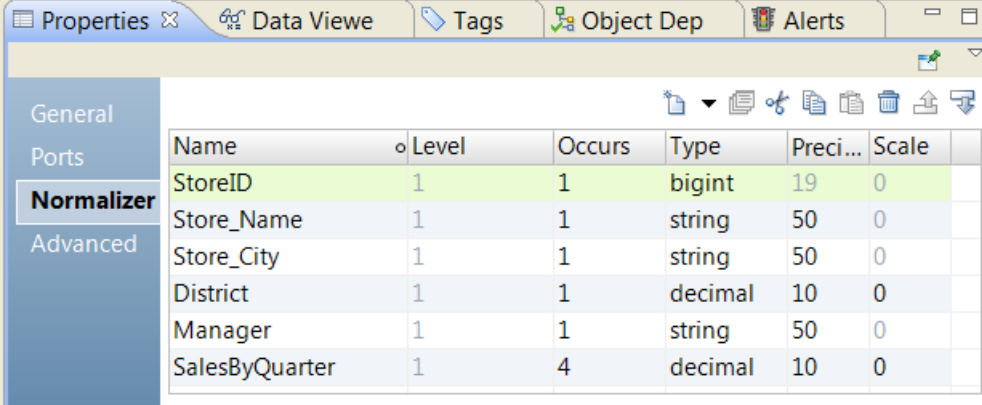
O arquivo simples STORE contém os seguintes dados de origem:

StoreID	Store_Name	Store_City	Distrito	Gerente	Trimestre1	Trimestre2	Trimestre3	Trimestre4
1	BigStore	Nova York	Leste	Robert	100	300	500	700
2	SmallStore	Phoenix	Oeste	Radhika	250	450	650	850

Adicione o arquivo simples a um mapeamento como uma transformação de Leitura e depois crie uma transformação do Normalizador vazia. Arraste as portas do objeto de leitura Read\_STORE até a transformação de Normalizador para criar a definição de Normalizador.

A exibição **Normalizador** contém uma instância dos campos Store\_Name, Store\_City, Distrito e Gerente. A exibição **Normalizador** contém quatro instâncias de um campo chamado QUARTER. Mescle os campos QUARTER para criar um campo SalesByQuarter que ocorre quatro vezes.

A seguinte figura mostra a definição do Normalizador com os campos de Trimestre mesclados:



Name	Level	Occurs	Type	Preci...	Scale
StoreID	1	1	bigint	19	0
Store_Name	1	1	string	50	0
Store_City	1	1	string	50	0
District	1	1	decimal	10	0
Manager	1	1	string	50	0
SalesByQuarter	1	4	decimal	10	0

## Exemplo de Grupos de Entrada e Saída do Normalizador

Depois que você modifica a hierarquia de entrada, a transformação de Normalizador tem um grupo de entrada e um grupo de saída padrão. Você deve reorganizar as portas de saída em dois grupos. Você precisa de um grupo que contenha as informações de Loja e um grupo que contenha as informações de Vendas.

O grupo de entrada contém uma porta para cada campo da origem. O grupo de saída contém portas para os campos de loja e uma porta para o campo de ocorrência múltipla SalesByQuarter. O grupo de saída também contém uma ID de coluna gerada, GCID\_SalesByQuarter, que corresponde ao campo de ocorrência múltipla SalesByQuarter.

Para retornar as vendas trimestrais para outro destino, crie um novo grupo na exibição **Visão Geral**. No grupo de Saída1, adicione os seguintes campos:

```
StoreID
SalesByQuarter
GCID_SalesByQuarter
```

Atualize o grupo de saída padrão. Remova os seguintes campos:

```
SalesByQuarter
GCID_SalesByQuarter
```

A seguinte imagem mostra o grupo de entrada e os grupos de saída na exibição **Visão Geral**:

	Name	Type	Precision	Scale	Location
<b>Input</b>					
1	StoreID	bigint	19	0	StoreID
2	Store_Name	string	50	0	Store_N...
3	Store_City	string	50	0	Store_City
4	District	decimal	10	0	District
5	Manager	string	50	0	Manager
6	Quarter1	decimal	10	0	SalesBy...
7	Quarter2	decimal	10	0	SalesBy...
8	Quarter3	decimal	10	0	SalesBy...
9	Quarter4	decimal	10	0	SalesBy...
<b>Output</b>					
1	StoreID	bigint	19	0	StoreID
2	Store_Name	string	50	0	Store_N...
3	Store_City	string	50	0	Store_City
4	District	decimal	10	0	District
5	Manager	string	50	0	Manager
<b>Output1</b>					
1	StoreID	bigint	19	0	StoreID
2	SalesByQuarter	decimal	10	0	SalesBy...
3	GCID_SalesByQuarter	bigint	19	0	SalesBy...

A StoreID é a chave gerada que vincula as informações de Loja às informações de Vendas. Verifique se ambos os grupos de saída retornam a StoreID.

## Saída de Mapeamento de Exemplo de Normalizador

Adicione a transformação de Gravação ao mapeamento e conecte as portas de saída da transformação do Normalizador aos objetos de dados.

Quando você executa o mapeamento, a transformação de Normalizador grava as seguintes linhas no destino de Armazenamento:

StoreID	Store_Name	Store_City	Distrito	Gerente
1	BigStore	Nova York	Leste	Robert
2	SmallStore	Phoenix	Oeste	Radhika

A transformação de Normalizador grava as seguintes linhas no destino de Vendas:

StoreID	SalesByQuarter	GCID_SalesByQuarter
1	100	1
1	300	2
1	500	3
1	700	4
2	250	1
2	450	2
2	650	3
2	850	4



# Transformação de Normalizador em um ambiente não nativo

O processamento da transformação de Normalizador em um ambiente não nativo depende do mecanismo que executa a transformação.

Considere o suporte para os seguintes mecanismos de tempo de execução não nativos:

- Mecanismo Blaze. Suportado sem restrições.
- Mecanismo Spark. Com suporte sem restrições em mapeamentos de fluxo e em lote.
- Mecanismo Databricks Spark. Suportado sem restrições.

## CAPÍTULO 32

# Transformação de Mesclagem

Este capítulo inclui os seguintes tópicos:

- [Transformação de Mesclagem - Visão Geral, 542](#)
- [Configurando uma Estratégia de Mesclagem, 542](#)
- [Propriedades Avançadas da Transformação de Mesclagem, 543](#)
- [Transformação de Mesclagem em um ambiente não nativo, 543](#)

## Transformação de Mesclagem - Visão Geral

A transformação de Mesclagem é uma transformação passiva que lê os valores de dados de várias colunas de entrada e cria uma única coluna de saída.

Use a transformação de Mesclagem para criar dados em um formato de sua preferência. Por exemplo, você pode combinar os campos Customer\_Firstname e Customer\_Surname para criar um campo chamado Customer\_FullName.

Dentro de uma transformação de Mesclagem, é possível criar várias estratégias de mesclagem. A transformação de Mesclagem fornece um assistente usado para criar estratégias.

## Configurando uma Estratégia de Mesclagem

Para configurar uma estratégia de mesclagem, edite as configurações na exibição **Estratégias** de uma transformação de Mesclagem.

1. Selecione a exibição **Estratégias**.
2. Clique em **Novo**.  
O assistente de **Nova Estratégia** é aberto.
3. Clique no campo **Entradas** para selecionar portas de entrada para a estratégia.
4. Para definir o caractere de mesclagem a ser colocado entre os itens mesclados, clique em **Escolher**. Se você não selecionar um caractere de mesclagem, a transformação de Mesclagem usará um caractere de espaço por padrão.
5. Opcionalmente, selecione **Incluir strings vazias na saída mesclada** para incluir strings de entrada vazias na saída.
6. Clique em **Concluir**.

# Propriedades Avançadas da Transformação de Mesclagem

Configure propriedades que ajudam a determinar como o Serviço de Integração de Dados processa dados para a transformação de Mesclagem.

Você pode configurar níveis de rastreamento para logs.

Configure a seguinte propriedade na guia **Avançado**:

## Nível de Rastreamento

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.

# Transformação de Mesclagem em um ambiente não nativo

O processamento da transformação de Mesclagem em um ambiente não nativo depende do mecanismo que executa a transformação.

Considere o suporte para os seguintes mecanismos de tempo de execução não nativos:

- Mecanismo Blaze. Suportado sem restrições.
- Mecanismo Spark. Com suporte sem restrições em mapeamentos em lote. Sem suporte em mapeamentos de fluxo.
- Mecanismo Databricks Spark. Sem suporte.

## CAPÍTULO 33

# Transformação de Analisador

Este capítulo inclui os seguintes tópicos:

- [Transformação de Analisador - Visão Geral, 544](#)
- [Modos de Transformação de Analisador, 545](#)
- [Quando Usar uma Transformação de Analisador, 545](#)
- [Dados de Referência Usados na Transformação de Analisador, 546](#)
- [Operações de Análise de Token, 548](#)
- [Portas de Análise de Token, 549](#)
- [Propriedades de Análise de Token, 550](#)
- [Modo de Análise Baseada em Padrões, 552](#)
- [Configurando uma Estratégia de Análise de Token, 553](#)
- [Configurando uma Estratégia de Análise de Padrão, 554](#)
- [Propriedades Avançadas da Transformação de Analisador, 555](#)
- [Transformação de Analisador em um ambiente não nativo, 555](#)

## Transformação de Analisador - Visão Geral

A transformação de Analisador é uma transformação passiva que analisa valores de dados de entrada para novas portas. A transformação grava os valores em novas portas de acordo com os tipos de informação contidos nos valores e a posição dos valores na string de entrada.

Use uma transformação de Analisador quando desejar alterar a estrutura de um conjunto de dados. A transformação de Analisador pode adicionar colunas a um conjunto de dados e gravar valores de dados em novas colunas. Use uma transformação de Analisador quando uma coluna de dados contiver vários valores em uma única coluna e você desejar gravar os valores de dados em colunas individuais com base no tipo de informação que elas contiverem.

A transformação de Analisador analisa os valores de dados para as portas de saída que você define. Se a transformação puder identificar um valor de dados de entrada, mas não houver uma porta de saída definida disponível, a transformação gravará o valor em uma porta de overflow. Se a transformação não puder identificar um valor de dados de entrada, ela gravará o valor em uma porta de dados não analisados.

# Modos de Transformação de Analisador

Ao criar uma transformação de Analisador, selecione o modo de análise de token ou análise baseada em padrões.

Selecione um dos seguintes modos:

- Modo de análise de token. Use este modo para analisar valores de entrada que correspondem aos valores em objetos de dados de referência, como conjuntos de tokens, expressões regulares, modelos probabilísticos e tabelas de referência. Você pode usar várias estratégias de análise de token em uma transformação.
- Modo de análise baseada em padrões. Use este modo para analisar valores de entrada que correspondem aos valores em conjuntos de padrões.

## Quando Usar uma Transformação de Analisador

Use a transformação de Analisador quando os campos de dados em uma coluna contiverem mais de um tipo de informação e você desejar mover os valores de campo para novas colunas. A transformação de Analisador permite que você crie uma nova coluna para cada tipo de informação em um conjunto de dados.

Os exemplos a seguir descrevem alguns tipos de alterações estruturais que você pode executar com uma transformação de Analisador.

### **Criar novas colunas para dados de contato**

Você pode criar uma estrutura de dados que analisa dados de nome de uma única coluna em várias colunas. Por exemplo, você pode criar colunas para saudações, nomes, nomes do meio e sobrenomes.

Você configura a transformação com um modelo probabilístico que representa as estruturas dos nomes de pessoas na porta de entrada. Use uma amostra dos dados da porta de entrada para definir o modelo.

Você cria uma estratégia de análise de token que aplica o modelo probabilístico à porta de entrada e grava os valores de nome nas novas colunas. A transformação grava os valores de nome nas novas portas com base na posição de cada valor na string de entrada e o tipo de nome que o valor representa.

**Nota:** Você também pode usar uma estratégia de análise baseada em padrões para analisar dados de contato. Ao configurar uma estratégia de análise baseada em padrões, você define os padrões que representam as estruturas dos nomes na porta de entrada.

### **Criar colunas de endereço**

Você pode criar uma estrutura de dados que analisa uma única coluna de dados de endereço em várias colunas que descrevem um endereço para entrega.

Configure a transformação com tabelas de referência que contêm elementos de endereço reconhecíveis, como códigos postais, nomes de estado e nomes de cidade. Crie uma estratégia de análise de token que grava cada elemento de endereço em uma nova porta.

Não é possível usar uma tabela de referência para analisar dados de rua de uma string de entrada, porque as informações de nome da rua e número são muito gerais para serem capturadas em uma tabela de referência. No entanto, você pode usar a porta de Overflow para capturar esses dados. Quando você tiver analisado todos os dados de cidade, estado e código postal de um endereço, os dados restantes conterão informações de rua.

Por exemplo, use uma estratégia de análise de token para dividir o seguinte endereço em elementos de endereço:

```
123 MAIN ST NW STE 12 ANYTOWN NY 12345
```

A estratégia de análise pode gravar os elementos de endereço nas seguintes colunas:

Nome da Coluna	Dados
Overflow	123 MAIN ST NW STE 12
Cidade	ANYTOWN
Estado	NY
CEP	12345

#### Criar colunas de dados de produto

Você pode criar uma estrutura de dados que analisa uma única coluna de dados de produto e a converte em várias colunas que descrevem os detalhes de inventário do produto.

Configure a transformação com conjuntos de tokens que contêm elementos de inventário, como dimensão, cor e peso. Criar uma estratégia de análise de token que grava cada elemento de inventário em uma nova porta

Por exemplo, use uma estratégia de análise de token para dividir a seguinte descrição de tintas em elementos de inventário separados:

```
500ML Red Matt Exterior
```

A estratégia de análise pode gravar os elementos de endereço nas seguintes colunas:

Nome da Coluna	Dados
Tamanho	500ML
Cor	Vermelho
Estilo	Fosco
Exterior	Y

## Dados de Referência Usados na Transformação de Analisador

O Informatica Developer é instalado com vários objetos de dados de referência que você pode usar com a transformação de Analisador. Você também pode criar objetos de dados de referência na ferramenta Developer.

Quando você adicionar um objeto de dados de referência a uma transformação de Analisador, a transformação gravará as strings correspondentes a um valor no objeto nas novas colunas que você especificar.

A seguinte tabela descreve os tipos de dados de referência que você pode usar:

Tipo de Dados de Referência	Descrição
Conjuntos de padrões	Identifica os valores de dados com base na posição relativa de cada valor na cadeia.
Modelos probabilísticos	Adiciona recursos de correspondência difusa a operações de análise de token. A transformação pode usar um modelo probabilístico para inferir o tipo de informação em uma cadeia. Para ativar os recursos de correspondência difusa, compile o modelo probabilístico na ferramenta Developer.
Tabelas de referência	Localiza cadeias que correspondem às entradas em uma tabela de banco de dados.
Expressões regulares	Identifica cadeias que correspondem às condições definidas por você. Você pode usar uma expressão regular para localizar uma cadeia em uma cadeia maior.
Conjuntos de tokens	Identifica as cadeias com base nos tipos de informações que elas contêm. O Informatica instala, com os conjuntos de tokens, diferentes tipos de definições de token, como definições de palavra, número de telefone, CEP e código do produto.

## Conjuntos de Padrões

Um conjunto de padrões contém expressões que identificam padrões de dados na saída de uma operação de rotulagem de token. Você pode usar conjuntos de padrões para analisar a porta de saída de Dados com Token e gravar strings correspondentes em uma ou mais portas de saída. Use conjuntos de padrões nas transformações de Analisador que usam o modo de análise de padrão.

Por exemplo, você pode configurar uma transformação de Analisador para usar conjuntos de padrões que identifiquem nomes e iniciais. Essa transformação usa os conjuntos de padrões para analisar a saída de uma transformação de Rotulador no modo de rotulagem de token. Você pode configurar a transformação de Analisador para gravar nomes e iniciais na saída em portas separadas.

## Modelos Probabilísticos

Um modelo probabilístico identifica tokens pelos tipos de informações que eles contêm e pelas posições que ocupam em uma cadeia de entrada.

Um modelo probabilístico contém valores de dados de referência e de rótulos. Os valores de dados de referência representam os dados em uma porta de entrada que você conectar à transformação. Os valores de rótulos descrevem os tipos de informações contidas nos valores de dados de referência. Atribua um rótulo para cada valor de dados de referência no modelo.

Para vincular os valores de dados de referência aos rótulos em um modelo probabilístico, compile o modelo. O processo de compilação gera uma série de associações lógicas entre os valores de dados e os rótulos. Quando você executa um mapeamento que lê o modelo, o Serviço de Integração de Dados aplica a lógica de modelo aos dados de entrada da transformação. O Serviço de Integração de Dados retorna o rótulo descreve os valores de dados de entrada com mais precisão.

Você pode criar um modelo probabilístico na Developer tool. O repositório do Modelo armazena o objeto do modelo probabilístico. A Developer tool grava os valores de dados, os rótulos e os dados de compilação em um arquivo na estrutura de diretórios da Informatica.

**Nota:** Se você adicionar um modelo probabilístico a uma operação de análise de token e depois editar a configuração do rótulo nesse modelo, invalidará a operação. Ao atualizar a configuração do rótulo em um modelo probabilístico, recrie qualquer operação de análise que use o modelo.

## Tabelas de Referência

Uma tabela de referência é uma tabela de banco de dados que contém pelo menos duas colunas. Uma coluna contém o padrão ou versão necessária de um valor de dados, e outras colunas contêm versões alternativas do valor. Quando você adiciona uma tabela de referência a uma transformação, a transformação pesquisa os dados da porta de entrada em busca de valores que também apareçam na tabela. Você pode criar tabelas com quaisquer dados que sejam úteis para o projeto de dados no qual você está trabalhando.

## Expressões Regulares

No contexto de operações de análise, uma expressão regular é uma expressão que você pode usar para identificar uma ou mais strings em dados de entrada. A transformação de Analisador grava strings identificadas em uma ou mais portas de saída. Você pode usar expressões regulares nas transformações de Analisador que usam o modo de análise de token.

Transformações de Analisador usam expressões regulares para corresponder padrões em dados de entrada e analisar todas as strings correspondentes em uma ou mais saídas. Por exemplo, você pode usar uma expressão regular para identificar todos os endereços de e-mail em dados de entrada e analisar cada componente de endereço de e-mail em uma saída diferente.

## Conjuntos de Tokens

Um conjunto de tokens contém expressões que identificam tokens específicos. Você pode usar conjuntos de tokens em transformações de Analisador que utilizam o modo de análise de token.

Use conjuntos de tokens para identificar tokens específicos como parte de operações de análise. Por exemplo, você pode usar um conjunto de tokens para analisar todos os endereços de e-mail que usam um formato "NomeConta@NomeDomínio".

## Operações de Análise de Token

No modo de análise de token, a transformação de Analisador analisa strings que correspondem a dados em conjuntos de tokens, expressões regulares, modelos probabilísticos ou entradas de tabela de referência.

Para executar a análise de token, adicione estratégias na exibição **Estratégias** da transformação. Você pode adicionar um ou mais operações a cada estratégia. A transformação fornece um assistente que você usa para criar estratégias.

Você pode adicionar os seguintes tipos de operações a uma estratégia de análise de token:

### **Analisar Usando o Conjunto de Tokens**

Use conjuntos de tokens predefinidos ou definidos pelo usuário para analisar dados de entrada. Operações de conjuntos de tokens podem usar expressões regulares personalizadas que gravam dados em uma ou mais saídas.

Você também pode usar modelos probabilísticos para identificar e analisar valores de dados de entrada.

### **Analisar Usando Tabela de Referência**

Use tabelas de referência para analisar dados de entrada.

A transformação executa as operações na ordem em que elas aparecem na estratégia.



# Portas de Análise de Token

Configure as portas de análise de token com configurações apropriadas para seus dados.

Uma transformação de Analisador no modo de análise de token tem os seguintes tipos de porta:

## Entrada

Contém dados que você passa para a transformação de Analisador. A transformação mescla todas as portas de entrada em uma string de dados combinados usando o **Caractere de Associação de Entrada** especificado na guia **Estratégias**. Se você não especificar um caractere de associação de entrada, a transformação usará um caractere de espaço por padrão.

## Portas de Saída Analisadas

Portas de saída definidas pelo usuário que contêm strings analisadas com êxito. Em casos onde várias estratégias de análise usam a mesma saída, a transformação mescla a saída em uma string de dados combinados usando o **Caractere de Associação de Saída** especificado na guia **Estratégias**. Se você não especificar um caractere de associação de saída, a transformação usará um caractere de espaço por padrão.

## Overflow

Contém strings analisadas com êxito que não se encaixam no número de saídas definido na transformação. Por exemplo, se a transformação tem apenas duas saídas "WORD", a string "John James Smith" resulta em uma saída de overflow de "Smith". A transformação de Analisador criará uma porta de overflow para cada estratégia que você adicionar.

Quando você seleciona a opção Overflow Detalhado, a transformação cria uma porta de overflow para cada rótulo do modelo.

## Não Analisado

Contém strings que a transformação não consegue analisar com êxito. A transformação de Analisador criará uma porta não analisada para cada estratégia que você adicionar.

## Portas de Saída em Correspondência Probabilística

Quando você configura uma estratégia de análise para usar técnicas de correspondência probabilística, a transformação de Analisador adiciona uma porta para armazenar as pontuações de correspondência para cada porta de saída.

A seguinte tabela descreve os tipos de porta:

Tipo de Porta	Portas Criada em Correspondência Probabilística
Porta de saída analisadas	Saída de [nome do rótulo] Saída de pontuação de [nome do rótulo]
Porta de dados de overflow	Saída de [dados de overflow] Saída de pontuação de [dados de overflow]
Porta de dados não analisados	Saída de [dados não analisados] Saída de pontuação de [dados não analisados]

# Propriedades de Análise de Token

Configure propriedades para operações de análise de token na exibição **Estratégias** na transformação de Analisador.

## Propriedades Gerais

Propriedades gerais aplicam-se a todas as operações de análise de token que você define na estratégia. Use as propriedades gerais para nomear a estratégia, especificar as portas de entrada e saída e especificar se a estratégia permite técnicas de correspondência probabilística.

A tabela a seguir descreve as propriedades gerais.

Propriedade	Descrição
Nome	Fornecer um nome para a estratégia.
Entradas	Identifica as portas de entrada que as operações de estratégia podem ler.
Saídas	Identifica as portas de saída nas quais as operações de estratégia podem gravar.
Descrição	Descreve a estratégia. A propriedade é opcional.
Usar técnicas de correspondência probabilística	Especifica que a estratégia pode usar um modelo probabilístico para identificar tokens.
Caractere de Associação de Entrada	Especifica o caractere usado para associar portas de dados de entrada. A transformação mescla todas as portas de entrada em uma string de dados combinados e analisa a string completa.
Caractere de Associação de Saída	Especifica o caractere usado para associar valores de dados de saída quando várias operações de análise usam a mesma saída.
Inversão Ativada	Configura a estratégia para analisar dados da direita para a esquerda. Esta propriedade é desativada para correspondência probabilística.
Reversão de Overflow Ativada	Configura a estratégia para analisar os dados de overflow da direita para a esquerda. Esta propriedade é desativada para correspondência probabilística.
Overflow Detalhado Ativado	Cria um campo de overflow exclusivo para cada operação de análise.
Delimitadores	Especifica o delimitador que separa os dados de entrada em tokens distintos. O padrão é espaço.

## Propriedades do Modelo Probabilístico

Você pode selecionar um modelo probabilístico no lugar de um conjunto de tokens ao configurar uma estratégia de análise de token. Selecione a operação **Analisar Usando o Conjunto de Tokens** e selecione a opção para usar técnicas de correspondência probabilística.

A seguinte tabela descreve as propriedades do modelo probabilístico:

Propriedade	Descrição
Nome	Fornecer um nome para a operação.
Filtrar Texto	Usa caracteres ou caracteres curinga inseridos por você para filtrar a lista de conjuntos de tokens, modelos probabilísticos ou expressões regulares.
Modelo Probabilístico	Identifica o modelo probabilístico selecionado por você.

## Propriedades da Tabela de Referência

As propriedades da tabela de referência aplicam-se quando você configura uma operação de rotulagem para usar uma tabela de referência.

A seguinte tabela descreve as propriedades da tabela de referência:

Propriedade	Descrição
Nome	Fornecer um nome para a operação.
Tabela de Referência	Especifica a tabela de referência usada pela operação para analisar valores de entrada.
Distinção Entre Maiúsculas e Minúsculas	Determina se as strings de entrada devem corresponder à distinção entre maiúsculas e minúsculas das entradas da tabela de referência.
Substituir Correspondências por Valores Válidos	Substitui dados analisados pelos dados da coluna Válido na tabela de referência.
Saídas	Especifica as portas de saída para os dados analisados.

## Propriedades do Conjunto de Tokens

As propriedades do conjunto de tokens se aplicam quando você configura uma operação de análise para usar conjuntos de tokens.

Selecione a operação **Analisar Usando o Conjunto de Tokens** para analisar a entrada com conjuntos de tokens. Desmarque a opção para usar técnicas de correspondência probabilística.

As seguintes tabelas descrevem as propriedades do conjunto de tokens:

Propriedade	Descrição
Nome	Fornecer um nome para a operação.
Conjuntos de Tokens (Somente Saída Única)	Especifica o conjunto de tokens usado pela operação para analisar dados. A operação grava os dados em uma única porta.
Expressão Regular (Saídas Únicas ou Múltiplas)	Especifica a expressão regular usada pela operação para analisar dados. A operação gravará os dados em várias portas se encontrar várias cadeias no campo de entrada.
Saídas	Identifica as portas de saída nas quais as operações gravam.

Você pode adicionar, editar, importar ou remover um conjunto de tokens ou uma expressão regular. Você também pode filtrar a lista dos conjuntos de tokens.

A seguinte tabela descreve as propriedades que você usa para realizar as tarefas:

Propriedade	Descrição
Filtrar texto	Filtra a lista de conjuntos de tokens ou de expressões regulares. Use caracteres de texto e caracteres curinga como um filtro.
Adicionar	Use para definir um conjunto de tokens personalizado ou uma expressão regular.
Editar	Edita o conteúdo de um conjunto de tokens personalizado.
Importar	Importa uma cópia não reutilizável de um conjunto de tokens ou de uma expressão regular de uma pasta no repositório do Modelo. Se você atualizar o objeto de origem do conjunto de tokens ou da expressão regular, o Serviço de Integração de Dados não atualizará a cópia não reutilizável.
Remover	Exclui um conjunto de tokens personalizado ou uma expressão regular.

## Modo de Análise Baseada em Padrões

No modo de análise baseada em padrões, a transformação de Analisador analisa os padrões compostos de várias strings.

Você pode usar os seguintes métodos para definir padrões no modo de análise baseada em padrões:

- Analisar dados de entrada usando padrões definidos em tabelas de referência. Você pode criar uma tabela de referência de padrões a partir da saída com perfil criado de uma transformação de Rotulador que usa o modo de rotulagem de token.
- Analisar dados de entrada usando padrões definidos por você.
- Analisar dados de entrada usando padrões importados de um conjunto de padrões reutilizáveis do repositório do Modelo. Alterações no conjunto de padrões reutilizáveis não atualizam os dados adicionados na transformação de Analisador.

Você pode usar os caracteres curinga "+" e "\*" para definir um padrão. Use caracteres "\*" caracteres para corresponder a qualquer string e "+" para corresponder a uma ou mais instâncias da string precedente. Por

exemplo, use "WORD+" para encontrar várias instâncias consecutivas de um token de palavra e use "WORD \*" para localizar um token de palavra seguido de um ou mais tokens de qualquer tipo.

Você pode usar várias instâncias desses métodos na transformação de Analisador. A transformação usa as instâncias na ordem em que estão listadas na exibição **Configuração**.

**Nota:** No modo de análise baseada em padrões, a transformação de Analisador requer a saída de uma transformação de Rotulador que usa o modo de rotulagem de token. Crie e configure a transformação de Rotulador antes de criar uma transformação de Analisador que usa o modo de análise baseada em padrões.

## Portas de Análise Baseada em Padrões

Configure as portas de análise baseada com configurações apropriadas para seus dados.

Uma transformação de Analisador que usa o modo de análise baseada em padrões tem os seguintes tipos de porta:

### Label\_Data

Conecte esta porta à porta `Labeled_Output` de uma transformação de Rotulado que usa o modo de rotulagem de token.

### Tokenized\_Data

Conecte esta porta à porta de saída `Tokenized_Data` de uma transformação de Rotulador que usa o modo de rotulagem de token.

### Parse\_Status

Se uma correspondência é encontrada para o padrão de entrada, esta porta gera o valor `Com Correspondência`. Se nenhuma correspondência é encontrada, ela gera `Sem Correspondência`.

### Overflow

Strings analisadas com êxito que não se encaixam no número de saídas definido na transformação. Por exemplo, se apenas duas saídas "WORD" são definidas, a string "John James Smith" resulta em uma saída de overflow de "Smith" por padrão.

### Analisadas

Strings analisadas com êxito strings em portas definidas pelo usuário.

## Configurando uma Estratégia de Análise de Token

Para configurar uma estratégia de análise de token, abra uma transformação de Analisador no modo de análise de token e selecione a exibição **Estratégias**.

1. Selecione a exibição **Estratégias**.
2. Clique em **Novo**.  
O assistente de **Nova Estratégia** é aberto.
3. Clique no campo **Entradas** para selecionar portas para a estratégia.
4. Configure as propriedades da estratégia e clique em **Avançar**.
5. Escolha uma operação e clique em **Avançar**.
6. Configure as propriedades da operação e selecione portas de saída para os dados analisados com êxito.

7. Opcionalmente, clique em **Avançar** para adicionar mais operações à estratégia.
8. Após adicionar todas as operações à estratégia, clique em **Concluir**.
9. Opcionalmente, adicione mais estratégias à transformação.
10. Opcionalmente, altere a ordem na qual a transformação processa estratégias e operações. Selecione uma estratégia ou operação e clique em **Mover para Cima** ou **Mover para Baixo**.

## Configurando uma Estratégia de Análise de Padrão

Para configurar uma estratégia de análise de padrão, abra uma transformação de Analisador no modo de análise de padrão e selecione a exibição **Padrões**.

Antes de configurar a transformação para analisar padrões, verifique se a exibição **Padrões** mostra os nomes de porta de saída esperados. A transformação de Analisador analisa tokens para as portas de saída selecionadas. Crie portas de saída adicionais, se necessário.

1. Selecione a exibição **Padrões**.
2. Adicione um ou mais padrões à estratégia. Você pode adicionar padrões das seguintes maneiras:
  - Insira valores de dados para criar um padrão. Clique em **Novo** e selecione **Novo Padrão**.  
Se você selecionar **Novo Padrão**, clique em **Inserir Padrões Aqui** e insira um ou mais tipos de token. Os tokens inseridos devem corresponder à estrutura de token de um campo de dados de entrada. Adicione os padrões necessários para descrever as estruturas de token na porta de entrada.
  - Importe valores de dados de uma tabela de referência. Clique em **Novo** e selecione **Nova Tabela de Referência**.  
Se você selecionar **Nova Tabela de Referência**, navegue no repositório do Modelo e selecione uma tabela de referência que contenha uma lista de estruturas de token. A tabela de referência deve conter duas colunas. A segunda coluna na tabela de referência deve conter valores numéricos.
  - Importe valores de dados de um conjunto de padrões. Clique em **Importar** e selecione um conjunto de padrões reutilizáveis do repositório do Modelo.  
Se você selecionar **Importar**, procure os conjuntos de conteúdo no repositório do Modelo e selecione um conjunto de padrões reutilizáveis.

**Nota:** Você pode usar o campo **Filtrar texto** para filtrar as listas de tabelas de referência e conjuntos de padrões.

É possível misturar conjuntos de padrões e tabelas de referência na coluna Padrões.
3. Atribua cada token na coluna Padrões para uma porta de saída.
  - Para atribuir um token a uma porta de saída, clique duas vezes na coluna da porta e selecione o nome do token no menu.
  - Para analisar vários tokens para uma única saída, clique duas vezes na porta de coluna e selecione . Atribua tokens à porta e selecione o delimitador a ser usado.

Atribua os tokens em cada linha do padrão a uma ou mais portas de saída.
4. Salve a transformação.

# Propriedades Avançadas da Transformação de Analisador

Configure propriedades que ajudam a determinar como o Serviço de Integração de Dados processa dados para a transformação de Analisador.

Você pode configurar níveis de rastreamento para logs.

Configure a seguinte propriedade na guia **Avançado**:

## Nível de Rastreamento

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.

# Transformação de Analisador em um ambiente não nativo

O processamento da transformação de Analisador em um ambiente não nativo depende do mecanismo que executa a transformação.

Considere o suporte para os seguintes mecanismos de tempo de execução não nativos:

- Mecanismo Blaze. Suportado sem restrições.
- Mecanismo Spark. Com suporte sem restrições em mapeamentos de fluxo e em lote.
- Mecanismo Databricks Spark. Sem suporte.

# Transformação Python

A transformação de Python é uma transformação passiva que fornece uma interface para definir a funcionalidade de transformação usando o código Python. Você faz referência ao código Python e aos arquivos de recursos que você usa no código Python na transformação de Python.

Você pode usar uma transformação de Python para implementar um modelo de máquina nos dados transmitidos para a transformação. Por exemplo, você pode usar a transformação de Python para gravar o código Python que carrega um modelo pré-treinado. Você pode usar o modelo pré-treinado para classificar dados de entrada ou criar previsões.

Antes de poder usar a transformação de Python no mecanismo Spark, você deve instalar o Python na máquina do Serviço de Integração de Dados e configurar as propriedades avançadas do Spark correspondentes na conexão do Hadoop.

Para obter mais informações sobre a instalação de Python, consulte o *Guia do Data Engineering Integration*.

Você só pode executar a transformação Python no mecanismo Spark ou Databricks Spark. Você não pode executar a transformação Python no ambiente nativo.

Na versão 10.4.0, a transformação Python é suportada para visualização técnica no mecanismo Databricks Spark.

A funcionalidade de visualização técnica tem suporte para fins de avaliação, mas o suporte não é garantido, e o produto não está pronto para produção. A Informatica recomenda que você use apenas em ambientes de não produção. A Informatica pretende incluir a funcionalidade de visualização em uma próxima versão para uso em produção, mas pode optar por não fazer isso dependendo de eventuais mudanças nas circunstâncias técnicas ou de mercado. Para obter mais informações, entre em contato com o Suporte Global a Clientes da Informatica.

Para obter mais informações sobre a transformação Python, consulte o *Guia do Usuário do Data Engineering Integration*.



## CAPÍTULO 35

# Transformação de Classificação

Este capítulo inclui os seguintes tópicos:

- [Transformação de Classificação - Visão Geral, 557](#)
- [Transformações de Classificação em Mapeamentos Dinâmicos, 558](#)
- [Transformação de Classificação - Portas, 559](#)
- [Porta de Classificação, 560](#)
- [Definir Portas de Agrupamento, 560](#)
- [Caches de Classificação, 562](#)
- [Transformação de Classificação - Propriedades Avançadas, 562](#)
- [Transformação de Classificação em um ambiente não nativo, 563](#)

## Transformação de Classificação - Visão Geral

A transformação de Classificação é uma transformação ativa que limita os registros a um intervalo superior ou inferior. Use uma transformação de Classificação para retornar o maior ou menor valor numérico em uma porta ou um grupo. É possível usar uma transformação de Classificação também para retornar strings na parte superior ou inferior de uma ordem de classificação de mapeamento.

Durante uma execução de mapeamento, o Serviço de Integração de Dados armazena dados de entrada no cache até que ele possa realizar os cálculos de classificação.

A transformação de Classificação difere das funções de transformação MAX e MIN. A transformação de Classificação retorna um grupo de valores superiores ou inferiores, não apenas um valor. Por exemplo, use a transformação de Classificação para selecionar os dez melhores vendedores em um determinado território. Ou, para gerar um relatório financeiro, você poderia usar uma transformação de Classificação para identificar os três departamentos com os gastos mais baixos com salários e despesas gerais. Enquanto a linguagem SQL fornece muitas funções para controlar grupos de dados, não é possível identificar segmentos superiores ou inferiores de um conjunto de linhas usando funções SQL padrão.

Você conecta à transformação todas as portas que representam o mesmo conjunto de linhas. As linhas dentro dessa classificação, com base em alguma medida definida durante a configuração da transformação, passam pela transformação de Classificação.

Como transformação ativa, a transformação Classificação pode alterar o número de linhas que passam por ela. Você pode passar 100 linhas para a transformação de Classificação, mas só é possível selecionar as primeiras 10 linhas para classificar. As 10 primeiras linhas passam da transformação de Classificação para outra transformação.

Você pode conectar portas de uma transformação para a transformação de Classificação. Você também pode criar variáveis locais e gravar expressões não agregadas.

## Classificando Valores de String

Você pode configurar a transformação de Classificação para retornar os primeiros ou últimos valores de uma porta de string. O Data Integration Service classifica strings com base na ordem de classificação selecionada para o mapeamento implantado.

Ao configurar o aplicativo que contém o mapeamento, você seleciona a ordem de classificação usada pelo Data Integration Service para executar o mapeamento. Você pode selecionar binário ou um idioma específico, como francês ou alemão. Se você selecionar binário, o Data Integration Service calculará o valor binário de cada string e classificará as strings usando os valores binários. Se você selecionar um idioma, o Data Integration Service classificará as strings em ordem alfabética usando a ordem de classificação para o idioma.

## Propriedades da Transformação de Classificação

Quando cria uma transformação Classificação, você pode configurar as propriedades a seguir:

- Insira um diretório de cache.
- Selecione a ordem superior ou inferior.
- Selecione a porta de entrada/saída que contém valores usados para determinar a classificação. Você pode selecionar somente uma porta para definir uma classificação.
- Selecione o número de linhas que você deseja classificar.
- Defina grupos para ordens, como os 10 produtos mais baratos de cada fabricante.

# Transformações de Classificação em Mapeamentos Dinâmicos

É possível usar uma transformação de Classificação em um mapeamento dinâmico. Você pode configurar portas dinâmicas na transformação e fazer referência às portas geradas.

Se você fizer referência a uma porta gerada na transformação de Classificação e essa porta não existir em tempo de execução, o mapeamento falhará.

Se você especificar uma porta dinâmica como porta de Classificação, essa porta dinâmica não poderá ter mais de uma porta gerada.

Se você especificar uma porta dinâmica como a porta Agrupar por, o Serviço de Integração de Dados considerará todas as portas geradas como portas Agrupar por. O mapeamento não será válido se você especificar uma porta gerada como uma porta Agrupar por e especificar a porta dinâmica pai como uma porta de Classificação ou Agrupar por.

É possível parametrizar a porta de Classificação e as portas Agrupar por. Use um parâmetro de tipo de porta para a porta de Classificação. Use um parâmetro de tipo de lista de portas para as portas Agrupar por.

# Transformação de Classificação - Portas

A transformação de Classificação inclui portas de entrada, de entrada/saída ou de saída que estejam conectadas a outra transformação no mapeamento. A transformação também inclui portas de passagem e variáveis.

Uma transformação de Classificação tem os seguintes tipos de porta:

## Entrada

Recebe dados de transformações upstream. Você pode designar portas de entrada como portas de entrada/saída. A transformação deve ter pelo menos uma porta de entrada.

## Porta Dinâmica

Uma porta que pode receber várias colunas para criar um número dinâmico de portas geradas. Uma porta gerada é uma porta dentro de uma porta dinâmica que representa uma única coluna. É possível criar portas dinâmicas de entrada, saída e variáveis.

## Saída

Passa dados para transformações downstream. Você pode designar portas de saída como portas de entrada/saída. A transformação deve ter pelo menos uma porta de saída.

## Passagem

Passa os dados inalterados.

## Variável

Usado para variáveis locais. Você pode usar uma porta variável para armazenar valores ou cálculos a serem usados em uma expressão. Portas variáveis não podem ser portas de entrada ou de saída. Elas passam dados dentro da transformação.

## Índice de Classificação

A ferramenta Developer cria uma porta RANKINDEX para cada transformação de Classificação. O Data Integration Service utiliza a porta Índice de Classificação para armazenar a posição de classificação de cada linha em um grupo.

Por exemplo, você pode criar uma transformação Classificação para identificar os 50 funcionários com os salários mais altos na empresa. Você identifica a coluna SALARY como a porta de entrada/saída usada para medir as classificações e configura a transformação para filtrar todas as linhas, exceto as 50 primeiras.

Depois que a transformação Classificação identifica todas as linhas que pertencem à primeira ou à última classificação, ela atribui os valores de índice. No caso dos 50 primeiros funcionários, medidos por salário, o funcionário com o salário mais alto recebe um índice de classificação de 1. O próximo funcionário de salário mais alto recebe um índice de 2 e assim por diante. Ao medir uma classificação inferior, como os 10 produtos de preço mais baixo no inventário, a transformação Classificação atribui um índice de classificação desde o mais baixo até o mais alto. Portanto, o item menos caro receberia um índice de classificação de 1.

Se dois valores de classificação corresponderem, eles recebem o mesmo valor no índice de classificação e a transformação ignora o próximo valor. Por exemplo, se você deseja ver as cinco principais lojas de varejo no país e duas lojas têm as mesmas vendas, os dados retornados podem ser semelhantes aos mostrados abaixo:

RANKINDEX	SALES	STORE
1	10000	Orange
1	10000	Brea

RANKINDEX	SALES	STORE
3	90000	Los Angeles
4	80000	Ventura

RANKINDEX é uma porta somente de saída. Você pode passar o índice de classificação para outra transformação no mapeamento ou diretamente para um destino.

## Porta de Classificação

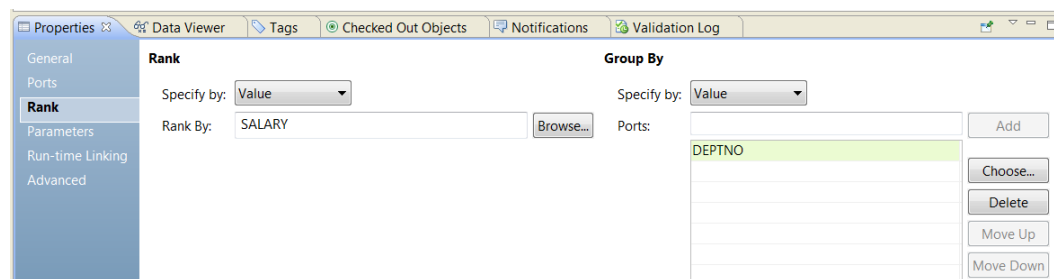
A porta de classificação determina a coluna pela qual classificar valores.

Você deve designar uma porta de entrada/saída ou portas de saída como a porta de classificação. Por exemplo, você cria uma transformação de Classificação para classificar os principais funcionários em cada departamento com base no salário. A porta Salário contém o salário para cada funcionário. Designe a porta de entrada/saída Salário como a porta de Classificação.

Selecione a porta de classificação na guia **Classificação** da exibição **Propriedades**. Você pode usar um parâmetro para a porta de classificação. Para usar um parâmetro, selecione **Especificar por Parâmetro**. Procure ou crie um parâmetro de porta. O valor padrão do parâmetro é o nome de uma porta ou de uma porta gerada.

É necessário vincular a porta de classificação a outra transformação.

A seguinte imagem mostra a guia **Classificação**:



**Nota:** A porta de rank não suporta tipos de dados binários.

## Definir Portas de Agrupamento

Você pode configurar a transformação de Classificação para criar grupos para as linhas classificadas.

Por exemplo, se você deseja selecionar os 10 itens mais caros por fabricante, deve definir primeiro um grupo para cada fabricante. No painel **Agrupar por** da guia **Classificação**, é possível definir uma das portas de entrada, de entrada/saída ou de saída como uma porta de agrupamento.

Para cada valor exclusivo na porta de grupo, a transformação cria um grupo de linhas que se encaixam na definição da classificação (primeira ou última e um número específico em cada classificação).

A transformação de Classificação altera o número de linhas de duas maneiras diferentes. Filtrando todas, menos as linhas que se encaixam na primeira ou na última classificação, você reduz o número de linhas transferidas pela transformação. Definindo grupos, você cria um conjunto de linhas classificadas para cada grupo.

Por exemplo, se você criar uma transformação de Classificação que classifique os cinco melhores vendedores agrupados por trimestre, o índice de classificação numerará os vendedores de 1 a 5 para cada trimestre:

RANKINDEX	SALES_PERSON	SALES	QUARTER
1	Sam	10,000	1
2	Mary	9,000	1
3	Alice	8,000	1
4	Ron	7,000	1
5	Alex	6,000	1

Defina o número de linhas a serem incluídas em uma classificação na guia **Avançado** da exibição **Propriedades**.

## Parâmetros de Agrupamento

É possível configurar um parâmetro de lista de portas que contém uma ou mais portas para inclusão no grupo. Crie um parâmetro de lista de portas selecionando portas em uma lista das portas na transformação.

A seguinte imagem mostra a guia **Agrupar por** quando você usa um parâmetro para identificar as portas do grupo:

É possível procurar um parâmetro de lista de portas ou clicar em **Novo** para criar esse parâmetro. Se você optar por criar um parâmetro de lista de portas, poderá selecionar as portas em uma lista das portas da transformação.

# Caches de Classificação

Quando você executa um mapeamento que usa uma transformação de Classificação, o Serviço de Integração de Dados cria um cache de índice e um cache de dados na memória para processar a transformação. Se o Serviço de Integração de Dados exigir mais espaço que o disponível no cache de memória, ele armazenará os dados excedentes nos arquivos de cache.

Quando você executa um mapeamento que usa uma transformação de Classificação, o Serviço de Integração de Dados compara uma linha de entrada com as linhas no cache de dados. Se a linha de entrada tiver classificação superior à de uma linha armazenada no cache, o Serviço de Integração de Dados substituirá a linha no cache pela linha de entrada. Se você configurar a transformação de Classificação para linhas de grupo, o Serviço de Integração de Dados classificará as linhas dentro de cada grupo.

O Serviço de Integração de Dados cria os seguintes caches para a transformação de Classificação:

- Cache de índice que armazena valores de grupo como configurado no grupo por portas.
- Cache de dados que armazena informações com base no grupo por portas.

## Transformação de Classificação - Propriedades Avançadas

Configure propriedades para ajudar a determinar como o Serviço de Integração de Dados processa dados para a transformação de Classificação.

Configure as seguintes propriedades na guia **Avançado**:

### Top/Bottom

Especifica se você deseja a primeira ou a última classificação para uma coluna.

### Número de Classificações

Número de linhas a serem incluídas na primeira ou última classificação.

### Comparação de Strings com Diferenciação de Maiúsculas/Minúsculas

Especifica se o Serviço de Integração de Dados utiliza comparações de strings com diferenciação de maiúsculas e minúsculas ao classificar strings. Desmarque essa opção para que o Serviço de Integração de Dados ignore o uso de maiúsculas e minúsculas nas strings. Por padrão, essa opção é selecionada.

### Diretório de Cache

O diretório onde o Serviço de Integração de Dados cria os arquivos de cache de índice e de dados. Verifique se o diretório existe e se contém espaço em disco suficiente para os arquivos de cache.

Insira vários diretórios separados por ponto-e-vírgula para melhorar o desempenho durante o particionamento do cache. O particionamento do cache cria um cache separado para cada partição que processa a transformação.

O padrão é o parâmetro do sistema CacheDir. Você pode configurar outro parâmetro do sistema ou parâmetro definido pelo usuário para essa propriedade.

#### **Tamanho do Cache de Dados de Classificação**

Quantidade de memória que o Serviço de Integração de Dados aloca para o cache de dados para a transformação no início da execução do mapeamento. Selecione Automático para que o Serviço de Integração de Dados calcule automaticamente os requisitos de memória no tempo de execução. Insira um valor específico em bytes quando você ajustar o tamanho do cache. O padrão é Automático.

#### **Tamanho do Cache de Índice de Classificação**

Quantidade de memória que o Serviço de Integração de Dados aloca para o cache de índice para a transformação no início da execução do mapeamento. Selecione Automático para que o Serviço de Integração de Dados calcule automaticamente os requisitos de memória no tempo de execução. Insira um valor específico em bytes quando você ajustar o tamanho do cache. O padrão é Automático.

#### **Nível de Rastreamento**

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.

#### **TÓPICOS RELACIONADOS:**

- [“Tamanho do Cache” na página 72](#)

## **Transformação de Classificação em um ambiente não nativo**

O processamento da transformação de Classificação em um ambiente não nativo depende do mecanismo que executa a transformação.

Considere o suporte para os seguintes mecanismos de tempo de execução não nativos:

- Mecanismo Blaze. Suportado com restrições.
- Mecanismo Spark. Com suporte com restrições em mapeamentos de fluxo e em lote.
- Mecanismo Databricks Spark. Suportado com restrições.

## **Transformação de Classificação no mecanismo Blaze**

Algumas regras de processamento para o mecanismo Blaze são diferentes das regras de processamento do Serviço de Integração de Dados.

O cache de dados da transformação de Classificação é otimizado para usar o comprimento variável para armazenar tipos de dados binary e string que passam pela transformação de Classificação. A otimização está ativada para tamanhos de registros de até 8 MB. Se o tamanho do registro for maior que 8 MB, a otimização de tamanho variável será desativada.

Quando o comprimento de variável é usado para armazenar dados que passam pela transformação de Classificação no cache de dados, a transformação de Classificação é otimizada para usar entrada classificada e uma transformação de Classificador de passagem é inserida antes da transformação de Classificação no mapeamento de tempo de execução. Para visualizar a transformação de Classificador, visualize o mapeamento otimizado ou visualize o plano de execução no ambiente de validação do Blaze.

Durante a otimização do cache de dados, o cache de dados e o cache de índice da transformação de Classificação são definidos como Automático. O cache de classificação da transformação de Classificador é configurado para o mesmo tamanho que o cache de dados para a transformação de Classificação. Para

configurar o cache de classificação, você deve configurar o tamanho do cache de dados para a transformação de Classificação.

## Transformação de Classificação no mecanismo Spark

Algumas regras de processamento para o mecanismo Spark são diferentes das regras de processamento do Serviço de Integração de Dados.

### Validação de Mapeamento

A validação do mapeamento falha nas seguintes situações:

- A diferenciação de maiúsculas e minúsculas está desativada.

### Otimização do cache de dados

Você não pode otimizar o cache de dados para a transformação para armazenar dados usando o tamanho da variável.

## Transformação de Classificação em um mapeamento de fluxo

Os mapeamentos de fluxo têm regras de processamento adicionais que não se aplicam a mapeamentos em lote.

### Validação de Mapeamento

A validação do mapeamento falha nas seguintes situações:

- Uma transformação de Classificação está no mesmo pipeline de fluxo que uma transformação de Pesquisa passiva configurada com uma condição de pesquisa de desigualdade.
- Uma transformação de Classificação está upstream de uma transformação de Associador.
- Um pipeline de fluxo contém mais de uma transformação de Classificação.
- Um pipeline de fluxo contém uma transformação de Agregador e uma transformação de Classificação.

## Transformação de Classificação no mecanismo Databricks Spark

Algumas regras de processamento para o mecanismo Databricks Spark são diferentes das regras de processamento do Serviço de Integração de Dados.

### Validação de Mapeamento

A validação do mapeamento falha nas seguintes situações:

- A diferenciação de maiúsculas e minúsculas está desativada.

### Otimização do cache de dados

Você não pode otimizar o cache de dados para a transformação para armazenar dados usando o tamanho da variável.



## CAPÍTULO 36

# Transformação de Leitura

Este capítulo inclui os seguintes tópicos:

- [Visão Geral da Transformação de Leitura, 565](#)
- [Propriedades da Transformação de Leitura , 566](#)
- [Sincronizar Objetos de Dados Relacionais, 569](#)
- [Alterar o Objeto de Dados de Origem, 570](#)
- [Parâmetros da Transformação de Leitura, 572](#)
- [Restrições , 572](#)
- [Criar uma Transformação de Leitura, 573](#)

## Visão Geral da Transformação de Leitura

A transformação de Leitura é uma transformação passiva que lê dados de uma origem. A transformação de Leitura não é reutilizável.

É possível criar uma transformação de Leitura a partir de um objeto de dados físicos ou de um objeto de dados lógicos. Se quiser criar uma transformação de Leitura a partir de um objeto de dados físicos importado de uma origem de adaptador do PowerExchange®, o editor de mapeamento poderá solicitar que você especifique uma operação de leitura antes de criar uma transformação de Leitura a partir desse objeto de dados.

Você pode configurar diferentes propriedades para uma transformação de Leitura com base no tipo de objeto de dados que usou para criar a transformação. Por exemplo, se você criar uma transformação de Leitura a partir de um objeto de dados relacional, poderá configurar substituições de SQL e definir restrições. As propriedades que você pode configurar também dependem de você ter configurado parâmetros para a transformação.

Transformações de Leitura podem conter origens dinâmicas. É possível configurar uma transformação de Leitura para atualizar dinamicamente suas portas, metadados e outras propriedades. Para obter informações sobre como configurar fontes dinâmicas, leia o capítulo "Mapeamentos Dinâmicos" no *Guia de Mapeamento do Informatica Developer*.

# Propriedades da Transformação de Leitura

Depois de criar uma transformação de Leitura, você pode configurar propriedades para a transformação.

Configurar propriedades de transformação de Leitura em guias de propriedades. As guias que você pode usar dependem do tipo de origem que a transformação de Leitura representa.

A seguinte tabela descreve cada guia de propriedade e identifica o tipo de origem para o qual você usa a guia:

Guia Propriedade	Descrição	Tipo de Origem
Geral	Especifique as propriedades e o comportamento da transformação. Para fontes de objetos de dados relacionais e personalizados, sincronize portas de entrada de transformação com a origem.	Tudo
Objeto de Dados	Especifique a fonte de dados da transformação.	<ul style="list-style-type: none"><li>- Relacional</li><li>- Arquivo simples</li><li>- Objeto de dados personalizados</li></ul>
Portas	Defina a configuração de porta pelo objeto de dados associado.	<ul style="list-style-type: none"><li>- Relacional</li><li>- Objeto de dados personalizados</li><li>- Objeto de dados lógicos</li></ul>
Formato	Configurações de entrada para uma fonte de dados de arquivo simples	Arquivo simples
Consulta	Especifique uma consulta para a origem.	<ul style="list-style-type: none"><li>- Relacional</li><li>- Objeto de dados personalizados</li><li>- Objeto de dados lógicos</li></ul>
Tempo de execução	Defina o comportamento de tempo de execução.	<ul style="list-style-type: none"><li>- Relacional</li><li>- Arquivo simples</li><li>- Objeto de dados personalizados</li></ul>
Origens	Selecione tabelas de origem e configure detalhes da origem.	<ul style="list-style-type: none"><li>- Relacional</li><li>- Objeto de dados personalizados</li></ul>
Parâmetros de Objeto de Dados	Defina propriedades de parâmetro.	<ul style="list-style-type: none"><li>- Arquivo simples</li><li>- Objeto de dados personalizados</li><li>- Objeto de dados lógicos</li></ul>
Vinculação em Tempo de Execução	Configure um link de grupo para grupo entre transformações que usa um parâmetro, uma política de link ou ambos para determinar quais portas devem ser vinculadas em tempo de execução.	Tudo
Avançado	Definir o nível de rastreamento e a ordem das linhas. Para uma origem relacional, defina a opção para criar ou substituir a tabela de destino em tempo de execução.	Tudo

## Propriedades Gerais

É possível configurar o nome e a descrição da transformação de Leitura. Você também pode configurar as seguintes propriedades:

### Quando os Metadados da Coluna Forem Alterados

Disponível para fontes relacionais. Selecione uma das seguintes opções:

- Sincronizar portas de saída. A Developer tool atualiza portas de saída de transformações de Leitura com alterações de metadados que o repositório do Modelo armazena para o objeto de dados.
- Não sincronizar. A transformação de Leitura não mostra alterações de metadados no objeto de dados.

### Objeto de Dados Físicos

Disponível para origens de arquivo simples e personalizadas. O objeto usado para criar a transformação.

Você pode selecionar o nome do objeto de dados e configurar suas propriedades.

## Propriedades do Objeto de Dados

Na guia Objeto de Dados, é possível especificar ou alterar a origem da transformação de Leitura e tornar dinâmicas as fontes de objetos de dados relacionais, de arquivo simples e personalizadas.

As seguintes propriedades podem ser configuradas:

### Especificar por

Para especificar colunas de origem e metadados para a transformação de Leitura, selecione uma das seguintes opções:

- Valor. A transformação de Leitura usa o objeto de dados associado para especificar metadados e colunas de origem.
- Parâmetro. A transformação de Leitura usa um parâmetro para especificar metadados e colunas de origem.

### Objeto de Dados

Se você tiver criado a transformação de Leitura a partir de um objeto de dados existente, o campo exibirá o nome do objeto. Clique em **Procurar** para alterar o objeto de dados a ser associado à transformação de Leitura.

### Em tempo de execução, obter as colunas de objeto de dados da fonte de dados

Ao ativar essa opção, o Serviço de Integração de Dados busca alterações de metadados e definições de dados de tabelas de origem na transformação de Leitura.

## Propriedades de Consulta

Configure uma consulta SQL para um objeto de dados de recurso relacional ou objeto de dados personalizados.

Ao configurar propriedades na guia **Consulta**, opte por configurar propriedades simples ou avançadas.

Na exibição de propriedades **Simples**, você configura a instrução SQL padrão como uma instrução Definir Distintos e para editar as dicas e as condições de associação, filtro e classificação para a instrução.

Na exibição de propriedades **Avançadas**, é possível definir uma consulta SQL personalizada. Você pode selecionar de colunas no objeto de dados associado, ou de parâmetros, ou criar um novo parâmetro para representar um objeto de dados.

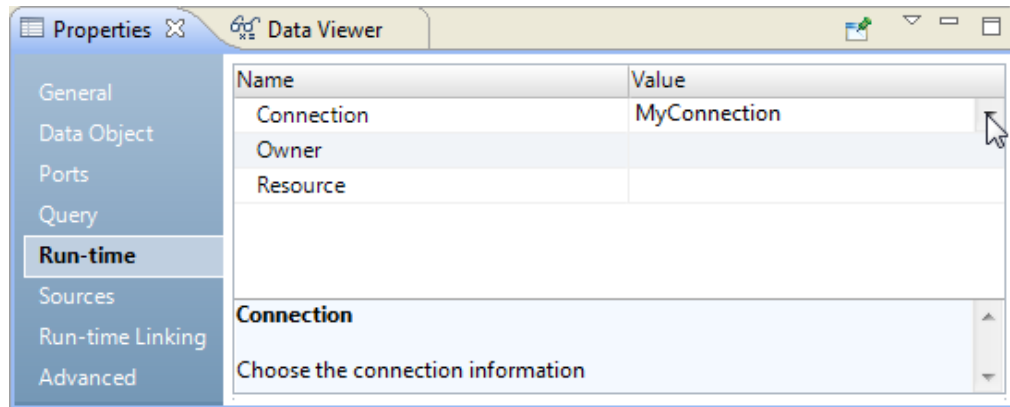
## Propriedades de Tempo de Execução

É possível configurar as seguintes propriedades da transformação de Leitura na guia **Tempo de execução**:

### Conexão

Disponível para fontes relacionais. Conexão usada pela transformação. Clique no lado direito do campo para alterar a conexão.

A seguinte imagem mostra a localização do botão suspenso a ser clicado:



## Propriedades de Origens

Configure os detalhes das origens para recursos relacionais e objetos de dados personalizados. Você pode alterar a definição do objeto de dados relacionais depois de importá-lo para o repositório. Você pode adicionar e remover portas, definir chaves primárias e configurar relacionamentos entre vários objetos de dados relacionais no repositório.

A guia **Origens** permite que você defina as seguintes configurações:

### Todas as origens

Use os botões Adicionar e Remover para adicionar e remover origens adicionais para a transformação.

### Guia Geral

Altere o nome e a descrição para a origem selecionada. Clique no nome da origem para alterar outros detalhes.

### Guia Chaves

Designa colunas de recursos como chaves.

### Guia Relacionamentos

Adicione e remova relacionamentos entre vários recursos relacionais.

## Propriedades Avançadas

Configure propriedades avançadas para determinar como o Serviço de Integração de Dados processa dados para a transformação de Leitura.

Configure as seguintes propriedades na guia Avançado:

### Nível de rastreamento

Controla a quantidade de detalhes no arquivo de log de mapeamento.

### PreSQL

Comando SQL que o Serviço de Integração de Dados executa no banco de dados de origem antes de ler a origem.

A Developer tool não valida o SQL.

### PostSQL

Comando SQL que o Serviço de Integração de Dados executa no banco de dados de origem depois de gravar no destino.

A Developer tool não valida o SQL.

### Restrições

Instruções SQL para restrições de integridade referencial em nível de tabela. Aplicável somente em origens relacionais.

## Sincronizar Objetos de Dados Relacionais

É possível sincronizar objetos de dados físicos quando suas origens são alteradas. Quando você sincroniza um objeto de dados físicos, a Developer tool reimporta os metadados do objeto a partir da origem selecionada.

É possível sincronizar todos os objetos de dados físicos. Ao sincronizar objetos de dados relacionais ou personalizados, você pode manter ou substituir os relacionamentos de chave definidos na Developer tool.

Escolha entre vários métodos para sincronizar objetos de mapeamento:

#### Sincronize um recurso relacional.

Para sincronizar qualquer objeto de dados físicos, clique com o botão direito do objeto na exibição **Object Explorer** e selecione **Sincronizar**.

#### Sincronize portos de transformação com o objeto de dados físicos.

Na guia Objeto de Dados de uma transformação, selecione a opção **No tempo de execução, obter colunas de objetos de dados da fonte de dados**.

No tempo de execução, o Serviço de Integração de Dados busca as alterações de metadados e a definição de dados da fonte de dados e atualiza a definição do objeto de dados no repositório do Modelo.

Para visualizar como o Serviço de Integração de Dados busca as alterações de metadados e a definição de dados, visualize o mapeamento com parâmetros resolvidos.

#### Sincronizar portas quando metadados forem alterados

Na guia Geral de uma transformação, selecione a opção para sincronizar portas. O rótulo exato dessa opção depende do tipo de transformação que você configurar. Por exemplo, para uma transformação de Leitura, a opção é **Quando metadados forem alterados, sincronizar portas de saída**.

Quando o mapeamento é executado, o Serviço de Integração de Dados sincroniza metadados de coluna na transformação com os metadados na fonte de dados.

# Alterar o Objeto de Dados de Origem

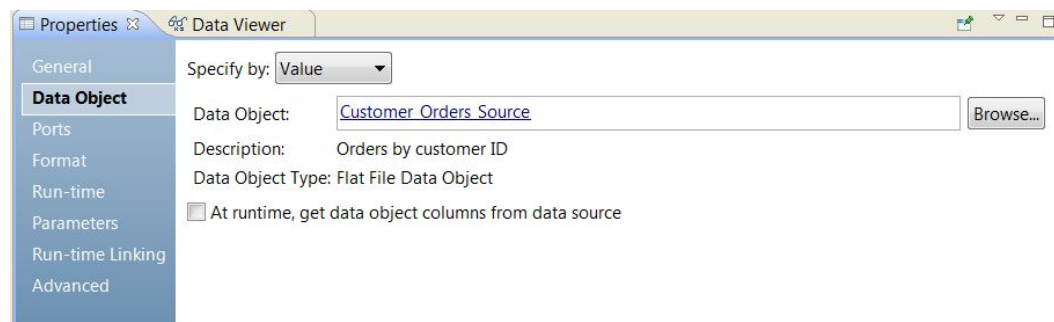
Uma transformação de Leitura se baseia em objeto de dados físicos ou lógicos no repositório do Modelo. Você pode alterar esse objeto de dados ao configurar uma transformação de Leitura. É possível parametrizar o objeto de dados para alterá-lo em tempo de execução. Por exemplo, você pode testar um mapeamento com um arquivo de origem diferente daquele utilizado para uma execução de mapeamento de produção.

Quando você cria uma transformação de um objeto de dados físicos, informações sobre esse objeto de dados são exibidas na guia **Objeto de Dados** das propriedades da transformação. Você pode clicar no nome desse objeto de dados para exibir a definição do objeto de dados físicos no repositório do Modelo.

É possível alterar o objeto de dados da transformação procurando um objeto de dados físicos diferente no repositório do Modelo. Quando você altera o objeto de dados, a transformação usa as propriedades de tempo de execução e as propriedades avançadas do objeto de dados selecionado.

É possível atualizar a estrutura do objeto de dados em tempo de execução com base em alterações na fonte de dados. A fonte de dados é o arquivo físico ou a tabela de banco de dados que o objeto de dados representa. Quando você permite que o Serviço de Integração de Dados obtenha colunas de dados da fonte de dados, o Serviço de Integração de Dados examina a estrutura da fonte de dados. O Serviço de Integração de Dados atualiza as portas de objetos de dados na instância de transformação com base na fonte de dados. O Serviço de Integração de Dados não altera a definição do objeto de dados físicos no repositório do Modelo.

A seguinte imagem mostra a guia **Objeto de Dados**:



A guia **Objeto de Dados** possui os seguintes campos:

## Especificar por

Escolha **Valor** para inserir um nome de objeto de dados específico. Escolha **Parâmetro** para parametrizar o objeto de dados.

## Objeto de dados

O nome do objeto de dados no repositório do Modelo. Você pode clicar no link **Objeto de Dados** para abrir a definição de objeto de dados do repositório. Você também pode procurar um objeto de dados diferente no repositório do Modelo.

## Descrição

A descrição do objeto de dados no repositório. Somente leitura.

## Tipo de objeto de dados

Descreve o tipo de objeto de dados, como um objeto de dados de arquivo simples, um objeto de tabela relacional ou um objeto de dados personalizado.

### Em tempo de execução, obter colunas de objetos da fonte de dados

O Serviço de Integração de Dados busca alterações de metadados e definição de dados do arquivo de dados ou da tabela a que o objeto de dados se refere e atualiza a estrutura do objeto de dados para a instância de transformação no tempo de execução.

Para visualizar como o Serviço de Integração de Dados busca as alterações de metadados e definição de dados no tempo de execução, visualize o mapeamento com parâmetros resolvidos.

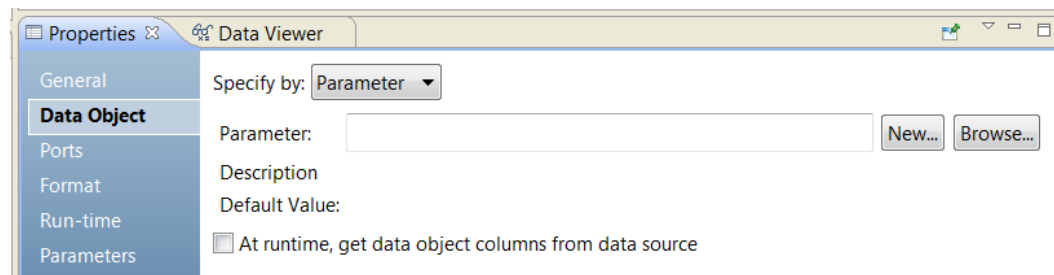
## Parametrizar a transformação de Leitura

Você pode parametrizar a transformação de Leitura e alterar o objeto de dados em tempo de execução.

Para parametrizar um objeto de dados, escolha **Especificar por Parâmetro** na guia **Objeto de Dados**. As propriedades na guia **Objeto de Dados** são modificadas.

Para parametrizar o objeto de dados, crie um parâmetro de tipo de recurso ou procure um parâmetro de recurso que você já tenha criado. O valor padrão do parâmetro é o nome do objeto de dados físicos no repositório do Modelo. Ao criar um valor de parâmetro padrão, você seleciona um nome de objeto de dados físicos em uma lista de objetos de dados no repositório.

A seguinte imagem mostra a guia **Objeto de Dados** quando você especifica o objeto de dados por um parâmetro:



A guia **Objeto de Dados** tem as seguintes opções por parâmetro:

#### Parâmetro

O nome de um parâmetro de recurso que você configurou como o objeto de dados. Somente leitura.

#### Descrição

A descrição do parâmetro. Somente leitura.

#### Novo

Crie um parâmetro de recurso. Procure e selecione um objeto de dados no repositório do Modelo para o valor padrão do parâmetro.

#### Procurar

Procure um parâmetro de recurso e selecione esse parâmetro.

#### Valor padrão

O valor padrão do parâmetro de recurso que você configurou para o objeto de dados. O valor padrão é um nome de objeto de dados físicos e o caminho para o objeto no repositório do Modelo. Somente leitura.

# Parâmetros da Transformação de Leitura

Você pode parametrizar algumas das propriedades de uma transformação de Leitura e algumas das propriedades do objeto de dados físicos reutilizáveis a partir do qual você cria a transformação de Leitura.

Ao criar um objeto de dados físicos, você configura as propriedades de leitura e gravação. Os parâmetros que você configura para propriedades de leitura em um objeto de dados físicos aparecem na guia **Parâmetros de Objeto de Dados** ao adicionar o objeto de dados a um mapeamento.

É possível configurar parâmetros para as seguintes propriedades de Leitura do objeto de dados físicos:

- Diretório do arquivo de controle
- Nome do arquivo de controle
- Escala padrão
- Delimitador
- Delimitador de arquivo simples
- Diretório do arquivo de mesclagem
- Nome do arquivo de origem
- Diretório do arquivo de origem

Depois de adicionar o objeto de dados físicos a um mapeamento, é possível visualizar os parâmetros na guia **Parâmetros do Objeto de Dados** da transformação de Leitura. Você pode expor esses parâmetros como parâmetros de mapeamento para substituir os valores dos parâmetros em tempo de execução.

**Nota:** Você não pode aninhar parâmetros definidos pelo usuário dentro de uma origem parametrizada. Se o objeto de dados de origem for parametrizado, você não poderá expor um parâmetro definido pelo usuário como um parâmetro de mapeamento para substituir os valores de parâmetro no tempo de execução. Em vez disso, o mapeamento usa o valor padrão.

Os seguintes parâmetros de mapeamento podem ser configurados para a transformação de Leitura:

- Conexão (relacional)
- Objeto de dados
- Ordem de resolução de links
- Nome do recurso (relacional)
- Nome do proprietário da tabela (relacional)

Você pode ver esses parâmetros na guia **Parâmetros do Objeto de Dados** do mapeamento.

## Restrições

Uma restrição é uma expressão condicional que os valores em uma linha de dados devem satisfazer.

Ao definir uma restrição, você insere uma expressão avaliada como TRUE para cada linha de dados.

O Serviço de Integração de Dados pode ler restrições de origens relacionais, objetos de dados lógicos, objetos de dados físicos ou tabelas virtuais. Para definir uma restrição em um objeto de dados físicos reutilizável, crie um objeto de dados personalizados.

Quando o Serviço de Integração de Dados lê restrições, ele pode descartar as linhas que não são avaliadas como TRUE para as linhas de dados com base no método de otimização aplicado.



Antes de definir uma restrição, verifique se os dados de origem satisfazem a condição definida por ela. Por exemplo, um banco de dados de origem tem uma coluna AGE que parece ter linhas com AGE < 70. Você pode definir uma restrição com AGE < 70 no banco de dados de origem. A Integração de Dados lê os registros do banco de dados de origem com a restrição AGE < 70. Se o Serviço de Integração de Dados ler os registros com AGE >= 70, ele poderá descartar as linhas com AGE >= 70.

No banco de dados, você pode usar comandos SQL para definir restrições no ambiente de banco de dados quando você se conectar ao banco de dados. O Serviço de Integração de Dados executa o SQL de ambiente de conexão sempre que se conectar ao banco de dados.

## Criar uma Transformação de Leitura

Ao criar uma transformação de Leitura, você escolhe um dos seguintes métodos com base no recurso a partir do qual a transformação é criada:

### Crie a transformação de um objeto de dados no repositório do Modelo.

Realize as seguintes etapas para criar uma transformação de Leitura a partir de um objeto de dados no repositório do Modelo:

1. Abra o mapeamento no editor.
2. Arraste um objeto de dados do **Object Explorer** até a exibição do editor.
3. Selecione **Leitura** e clique em **OK**.

A transformação de Leitura no mapeamento contém as portas e propriedades do objeto de dados.

### Crie a transformação usando o editor de mapeamento.

Use esse método se quiser definir configurações detalhadas da transformação de Leitura. Também é possível usar esse método se você desejar uma transformação de Leitura em um parâmetro.

Para criar uma transformação de Leitura no editor de mapeamento, consulte [“Criando uma Transformação de Leitura no Editor de Mapeamento” na página 573](#).

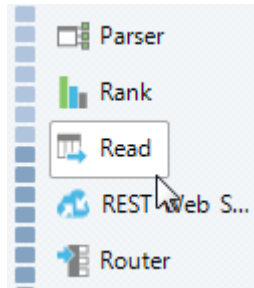
## Criando uma Transformação de Leitura no Editor de Mapeamento

É possível uma transformação de Leitura para representar a fonte de dados e metadados de colunas, bem como as propriedades em um mapeamento.

Realize as seguintes etapas:

1. Escolha um dos seguintes métodos para criar uma transformação de Leitura:
  - Clique com o botão direito do mouse no editor de mapeamento e selecione **Adicionar Transformação**.  
A caixa de diálogo **Adicionar Transformação** é aberta.  
Selecione a transformação de Leitura e clique em **Avançar**.
  - Role para baixo na paleta de mapeamento para localizar e clique duas vezes no ícone da transformação de Leitura.

A seguinte imagem mostra o ícone da transformação de Leitura:



A caixa de diálogo **Nova Transformação de Leitura** é aberta.

2. Para usar um objeto de dados de arquivo simples, recurso relacional ou personalizado como origem, realize as etapas a seguir:

- a. Selecione **Objeto de Dados Físicos** como o tipo de objeto de dados.
- b. Clique em **Procurar** para selecionar um objeto de dados de arquivo simples, recurso relacional ou personalizado.

A janela **Selecionar Objeto de Dados** é aberta.

- c. Selecione um objeto de dados e clique em **OK**.
- d. Opcionalmente, configure a transformação para buscar colunas de objetos de dados a partir da origem em tempo de execução. Selecione **Em tempo de execução, obter as colunas de objeto de dados da fonte de dados**.

O Serviço de Integração de Dados atualiza metadados de colunas para a transformação de Leitura quando o mapeamento é executado.

3. Para usar um parâmetro como origem, realize as seguintes etapas:

- a. Selecione **Criar usando um parâmetro**.
- b. Clique em **Novo** para criar um novo parâmetro ou clique em **Procurar** para selecionar um parâmetro existente.
- c. Selecione um parâmetro e clique em **OK**.
- d. Opcionalmente, configure a transformação para buscar colunas de objetos de dados a partir da origem em tempo de execução. Selecione **Em tempo de execução, obter as colunas de objeto de dados da fonte de dados**.

O Serviço de Integração de Dados atualiza metadados de colunas para a transformação de Leitura quando o mapeamento é executado.

4. Para usar um objeto de dados lógicos como origem, realize as seguintes etapas:

- a. Selecione **Objeto de Dados Lógicos** como o tipo de objeto de dados.
- b. Clique em **Procurar** para selecionar um objeto de dados e depois clique em **OK**.

5. Opcionalmente, digite um nome para a transformação de Leitura.

6. Clique em **Concluir**.

## CAPÍTULO 37

# Transformação de Relacional para Hierárquica

Este capítulo inclui os seguintes tópicos:

- [Visão geral da transformação de Relacional para Hierarquia, 575](#)
- [Exemplo - transformação de Relacional para Hierárquica, 576](#)
- [Portas relacionais de entrada e a exibição Visão Geral, 578](#)
- [Portas da transformação de Relacional para Hierarquia, 579](#)
- [Referências de esquema, 579](#)
- [Desenvolvimento de uma transformação de Relacional para Hierárquica, 579](#)

## Visão geral da transformação de Relacional para Hierarquia

A transformação de Relacional para Hierárquica processa a entrada relacional e a transforma em uma saída hierárquica. Uma transformação de Relacional para Hierárquica lê a entrada relacional de portas de entrada e transforma os dados para a saída hierárquica na porta de saída de transformação. Para transformar a entrada relacional em saída hierárquica, use um objeto de esquema para definir a estrutura hierárquica.

É possível usar o assistente de transformação de Relacional para Hierárquica para criar uma estrutura hierárquica que reflete as portas de entrada relacionais. O mapeamento para as portas de saída hierárquicas pode ser visualizado na exibição **Visão Geral** da transformação.

Depois de criar a transformação, você pode passar os dados da porta de saída hierárquica para outra transformação em um mapeamento.

**Nota:** A transformação Relacional em Hierárquica pode processar até 10.000 elementos de esquema em um arquivo .xsd. Para processar mais de 10.000 elementos, divida os dados em vários arquivos.

No modelo relacional, cada esquema de tabela identifica uma coluna, chamada de chave primária, para identificar exclusivamente cada linha. Você identifica o relacionamento entre cada linha da tabela e uma linha em outra tabela com uma chave externa. O assistente gera chaves ao criar a transformação. Você pode alterar uma transformação automaticamente gerada e adicionar, editar ou excluir portas.

Você pode vincular uma porta relacional de entrada a um nó na hierarquia. Vincule uma chave primária do elemento ou atributo relevante na hierarquia a um grupo relacional na entrada. A chave primária identifica cada linha nas tabelas relacionais.

Vincule uma chave externa do elemento ou atributo relevante na hierarquia a um grupo relacional na entrada. Uma chave externa na entrada relacional é uma coluna em uma tabela que aponta para a chave primária de outra tabela.

A porta relacional de entrada e o nó hierárquico devem ter tipos de dados compatíveis.

## Exemplo - transformação de Relacional para Hierárquica

O departamento de Finanças da empresa Electronics Superstore precisa processar cheques de pagamento para os funcionários da empresa. Eles precisam para transformar dados de funcionários armazenados em um banco de dados relacional em um formato hierárquico que seu sistema de pagamentos possa processar.

O mapeamento precisa usar uma transformação de Relacional para Hierárquica que processa a entrada dos detalhes de funcionários, como o nome, a ID, o endereço e os dados de conta bancária do funcionário, e gera a saída dos detalhes em um formato hierárquico utilizável.

Na entrada relacional, o elemento Bank\_ID é uma chave primária na tabela Employee e uma Chave externa na tabela Bank:

Employee_ID	Last_Name	First_Name	Address	Bank_ID	Bank_Account
9173327437	Sandrine	Jacques	74 Mobile Avenue	74845	8723487234
9174562342	Race	Tom	266 Crouse St.	9234734	45324734
8484526471	Jones	Charles	3815 LaValle Boulevard	389236	234638437
7023847265	Smith	Delilah	193 Short Drive	74845	8723463432
9174596725	Frederick	George	17 Serenity Road	9234734	6342636699

Bank_ID	Bank_Name	SWIFT_Code
74845	National Bank	9173327
9234734	International Bank	9174562
389236	Star National Bank	8484526

Na saída Payment em formato hierárquico, os elementos são combinados a partir das tabelas:

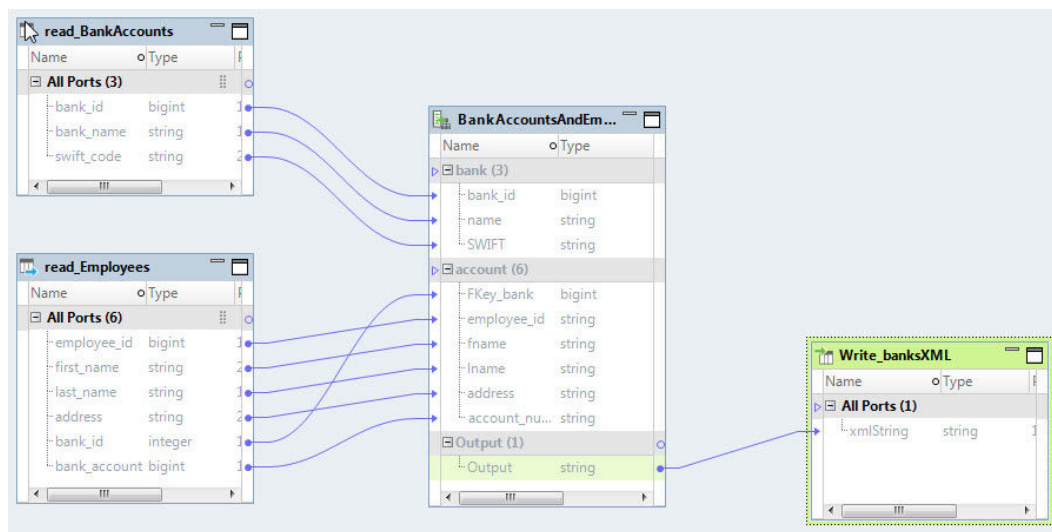
```
<banks>
  <bank name="National Bank" SWIFT="9173327">
    <account id="8723487234">
      <employee_id>9173327437</employee_id>
      <fname>Sandrine</fname>
      <lname>Jacques</lname>
      <address>74 Mobile Avenue</address>
    </account>
    <account id="8723463432">
      <employee_id>9082745558</employee_id>
      <fname>Delilah</fname>
      <lname>Smith</lname>
      <address>193 Short Drive</address>
    </account>
  </bank>
  <bank name="International Bank" SWIFT="9174562">
    <accounts>
      <account id="45324734">
        <employee_id>5534398889</employee_id>
        <fname>Race</fname>
        <lname>Tom</lname>
        <address>266 Crouse St.</address>
      </account>
      <account id="6342636699">
        <employee_id>9174596725</employee_id>
        <fname>Frederick</fname>
```

```

        <lname>George</lname>
        <address>17 Serenity Road</address>
    </account>
</accounts>
</bank>
<bank name="Star National Bank" SWIFT="8484526">
    <accounts>
        <account id="234638437">
            <employee_id>8484526471</employee_id>
            <fname>Jones</fname>
            <lname>Charles</lname>
            <address>3815 LaValle Boulevard</address>
        </account>
    </accounts>
</bank>
</banks>

```

A seguinte imagem mostra o mapeamento nesse exemplo:



O mapeamento contém os seguintes objetos:

#### **Read\_BankAccounts**

A origem que contém os dados bancários.

#### **Read\_Employees**

A origem que contém os dados dos funcionários.

#### **BankAccountsAndEmployees\_To\_PaymentsSystemXML**

Uma transformação de Relacional para Hierárquica que transforma a entrada relacional que contém informações de funcionários e contas bancárias em um formato XML que é consumido pelo sistema de pagamento.

#### **Write\_BanksXML**

Um caminho de destino para o arquivo que armazena os dados transformados sempre que você executa o mapeamento.

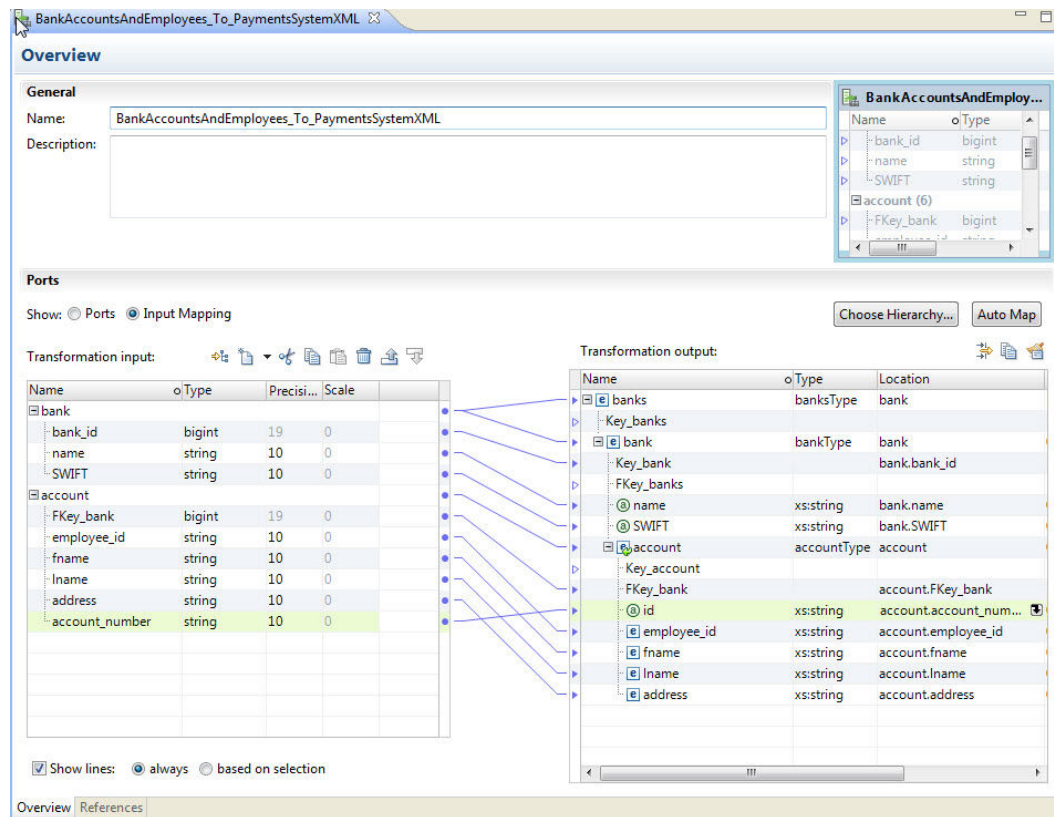
O mapeamento usa os arquivos Read\_BankAccount e Read\_Employees para fornecer a entrada relacional. O mapeamento processa e transforma os dados com a transformação BankAccountsAndEmployees\_To\_PaymentsSystemXML. Em seguida, ele armazena a saída no caminho de destino listado no arquivo simples Write\_BanksXML.

# Portas relacionais de entrada e a exibição Visão Geral

Para transformar dados relacionais em dados hierárquicos, o assistente cria uma estrutura hierárquica que reflete as portas de entrada relacionais. Você pode usar a exibição **Visão Geral** para vincular portas relacionais a portas hierárquicas.

Para exibir os links entre a entrada relacional e a saída hierárquica, use a exibição **Visão Geral**. Selecione **Mapeamento de Entrada**. O painel **Portas** aparece na exibição **Visão Geral**.

A seguinte imagem mostra o painel **Portas**:



À esquerda do painel **Portas**, está a área **Entrada da transformação**, que contém os elementos e grupos relacionais. À direita está a área **Saída de transformação**, que contém os nós de esquema hierárquicos.

Você pode criar portas na área **Entrada da transformação** e vincular elementos relacionais aos nós de esquema. Também é possível arrastar o ponteiro de um nó no esquema até um campo vazio na área **Entrada da transformação** para criar uma porta. Quando você conecta uma porta relacional a um nó do esquema, a Developer tool mostra um link entre eles.

# Portas da transformação de Relacional para Hierarquia

As portas da transformação de Relacional para Hierárquica são definidas na exibição **Visão Geral** da transformação.

Uma transformação de Relacional para Hierárquica pode ler a saída de dados relacionais a partir de um buffer. As portas de saída retornam dados hierárquicos para um buffer.

## Referências de esquema

Uma transformação de Relacional para Hierárquica requer um esquema hierárquico para definir a hierarquia de saída na transformação. Para usar o esquema na transformação, defina uma referência de esquema.

É possível definir referências de esquema de transformação na exibição **Referências** da transformação.

A transformação de Relacional para Hierárquica faz referência a objetos de esquema no repositório do Modelo. Os objetos de esquema podem existir no repositório antes de você criar a transformação.

Um esquema pode fazer referência a esquemas adicionais. A exibição **Referências** mostra o espaço de nome e o prefixo para cada esquema referenciado pela transformação de Relacional para Hierárquica.

## Desenvolvimento de uma transformação de Relacional para Hierárquica

Use o assistente de Nova Transformação para gerar automaticamente uma transformação de Relacional para Hierárquica. Escolha um arquivo de amostra de esquema ou hierárquico para definir a hierarquia de saída.

### Criando a transformação de Relacional para Hierárquica

1. Na Developer tool, clique em **Arquivo > Novo > Transformação**.
2. Selecione a transformação de Relacional para Hierárquica e clique em **Avançar**.
3. Insira um nome para a transformação, procure uma localização do repositório do Modelo para inserir a transformação e clique em **Avançar**.
4. Para selecionar um esquema, selecione um dos seguintes métodos:
  - Para usar um esquema do repositório do Modelo para definir a hierarquia de saída, perto do campo **Objeto de Esquema**, navegue para selecionar o arquivo de esquema do repositório.
  - Para importar um arquivo de esquema, clique em **Criar um novo objeto de esquema**. Na janela **Novo objeto de esquema**, você pode procurar e selecionar um arquivo de esquema ou pode optar por criar um esquema a partir de arquivo hierárquico de amostra.
5. Escolha a raiz da hierarquia de saída. Na caixa de diálogo **Raiz da hierarquia**, selecione o elemento no esquema que é o elemento raiz do arquivo hierárquico de saída. Para ajudar a selecionar o objeto raiz,

você pode adicionar um arquivo hierárquico de amostra. Para adicionar um arquivo de amostra, ao lado do campo **Arquivo de amostra**, procure e selecione o arquivo no sistema de arquivos.

6. Clique em **Concluir**.

O assistente cria a transformação no repositório.

## Criando as portas

Configure as portas na exibição **Visão Geral**.

1. Para visualizar o mapeamento, na área **Portas** da exibição **Visão Geral**, selecione **Mapeamento de Entrada**.
2. Selecione o tipo de porta de entrada, a precisão e a escala.
3. Expanda as árvores na grade **Portas**. À esquerda, o painel **Entrada de transformação** mostra a entrada relacional e, à direita, o painel **Saída de transformação** mostra a saída hierárquica esperada.
4. Para definir um nó como uma raiz, clique em **Escolher Hierarquia**.  
A Developer tool exibe apenas os nós do nível raiz e abaixo dele na área **Entrada de transformação**.
5. Para exibir linhas que conectam as portas aos nós hierárquicos, clique em **Mostrar Linhas**. Selecione para exibir todas as linhas de conexão ou apenas as linhas das portas selecionadas.
6. Para adicionar uma porta ou um grupo de entrada à área **Entrada da transformação**, use um dos seguintes métodos:
  - Arraste um elemento simples ou complexo na área **Saída de transformação** até uma coluna vazia na área **Entrada de transformação**. Se o nó for um nó de grupo, a Developer tool adicionará um grupo relacional sem portas.
  - Para adicionar um grupo relacional, selecione uma linha e clique com o botão direito do mouse para selecionar **Novo > Grupo**.
  - Para adicionar uma porta relacional, clique com o botão direito do mouse para selecionar **Novo > Campo**.
7. Para limpar as configurações de nós hierárquicos para localizações de portas, use um dos seguintes métodos:
  - Selecione um ou mais nós na área **Saída de transformação**, clique com o botão direito do mouse e selecione **Limpar**.
  - Selecione uma ou mais linhas que conectam as portas de entrada relacionais aos nós hierárquicos, clique com o botão direito do mouse e selecione **Excluir**.
8. Para exibir as portas de saída em uma hierarquia, clique em **Mostrar como Hierarquia**. Cada grupo filho é exibido embaixo do grupo pai.



## CAPÍTULO 38

# Transformação de Consumidor de Serviço da Web REST

Este capítulo inclui os seguintes tópicos:

- [Visão Geral da Transformação de Consumidor de Serviço da Web REST, 581](#)
- [Configuração da Transformação de Consumidor de Serviço da Web REST, 583](#)
- [Métodos HTTP, 585](#)
- [Portas da Transformação de Consumidor de Serviço da Web REST, 588](#)
- [Mapeamento de Entradas da Transformação de Consumidor de Serviço da Web REST, 591](#)
- [Mapeamento da Saída da Transformação de Consumidor de Serviço da Web REST, 593](#)
- [Propriedades Avançadas da Transformação de Consumidor de Serviço da Web REST, 595](#)
- [Criação da Transformação de Consumidor de Serviço da Web REST, 596](#)
- [Analisando uma mensagem de resposta JSON que contém matrizes, 597](#)

## Visão Geral da Transformação de Consumidor de Serviço da Web REST

A Transformação de Consumidor de Serviço da Web REST é uma transformação ativa que se conecta a um serviço da Web REST como um cliente de serviços da Web para acessar ou transformar dados. Use uma transformação de Consumidor de Serviço da Web REST para se conectar a um serviço da Web REST. A transformação de Consumidor de Serviço da Web REST pode enviar uma solicitação para um serviço da Web REST e receber uma resposta desse serviço.

A transformação de Consumidor de Serviço da Web REST se conecta a um serviço da Web através de uma URL que você define na transformação ou em uma conexão HTTP. Também é possível usar uma conexão HTTPS. As transformações de Consumidor de Serviço da Web REST podem usar o TLS 1.2, TLS 1.1 ou TLS 1.0.

Um serviço da Web REST contém um método HTTP para cada ação à qual o serviço da Web oferece suporte. Quando o Serviço de Integração de Dados se conecta a um serviço da Web REST, ele pode enviar uma solicitação para obter, publicar, inserir ou excluir dados. A solicitação pode atuar em recursos individuais ou em conjuntos de recursos. Depois que o Serviço de Integração de Dados envia uma mensagem de solicitação, ele recebe uma mensagem de resposta do serviço da Web.

A solicitação e as mensagens de resposta contêm dados XML ou JSON com elementos que podem formar uma hierarquia. Quando uma solicitação ou uma mensagem de resposta contém elementos de ocorrência múltipla, grupos de elementos formam níveis na hierarquia XML ou JSON. Os grupos são relacionados quando um nível está aninhado dentro de outro.

Na transformação de Consumidor de Serviço da Web REST, a entrada e a saída do método definem a estrutura da solicitação e das mensagens de resposta. A entrada e a saída do método incluem mapeamentos que definem como mapear os elementos da mensagem para as portas de entrada e de saída.

A transformação de Consumidor de Serviço da Web REST dá suporte a um servidor proxy. Você também pode se conectar a um aplicativo do Microsoft SharePoint com a transformação de Consumidor de Serviço da Web REST.

## Exemplo

Uma loja online define recursos para um banco de dados de produtos. Esse banco de dados identifica cada produto por número de peça.

Os clientes de serviços da Web acessam os detalhes dos produtos por meio de um serviço da Web REST. O serviço da Web usa a seguinte URL:

```
http://www.HypoStores.com/products/ProductDetails
```

Você precisa recuperar detalhes sobre um determinado produto, como a descrição e o preço unitário, e transmitir esses detalhes para uma transformação downstream em um mapeamento. Crie uma transformação de Consumidor de Serviço da Web REST para recuperar detalhes sobre um produto e transmiti-los para outra transformação.

A seguinte tabela mostra os detalhes da transformação que você pode configurar:

Detalhe da Transformação	Valor
Método HTTP	Get
URL Base	http://www.HypoStores.com/products/ProductDetails
Porta de argumentos de entrada	Part_No
Portas de saída	Descrição, Unit_Price
Saída do método	<A estrutura da mensagem de resposta.>

A saída do método inclui um mapeamento de saída que define como os elementos na mensagem de resposta são mapeados para as portas de saída.

Quando o Serviço de Integração de Dados envia a solicitação ao serviço da Web, ele acrescenta o valor na porta de argumentos à URL base. Por exemplo, para recuperar detalhes sobre a peça 0716, o Serviço de Integração de Dados usa a seguinte URL:

```
http://www.HypoStores.com/products/ProductDetails?Part_No=0716
```

Quando o Serviço de Integração de Dados recebe uma resposta, ele converte a descrição do produto e o preço unitário da mensagem de resposta em dados para as portas de saída.

Você também pode transmitir Part\_No como parâmetro e substituir o valor midstream quando executar o mapeamento.

## Processo da Transformação de Consumidor de Serviço da Web REST

A transformação de Consumidor de Serviço da Web REST cria uma mensagem de solicitação de acordo com os dados nas portas de entrada e na entrada do método. Ela converte os elementos da mensagem de resposta em dados para as portas de saída com base na saída do método.

As portas de entrada da transformação de Consumidor de Serviço da Web REST contêm dados relacionais de transformações upstream em um mapeamento. O Serviço de Integração de Dados usa a entrada do método para converter dados das portas de entrada em elementos na mensagem de solicitação.

Para se conectar ao serviço da Web, o Serviço de Integração de Dados lê a URL base que você configura nas propriedades da transformação ou na conexão HTTP. Ele identifica o recurso que você deseja obter, publicar, inserir ou excluir acrescentando valores das portas de URL ou das portas de argumentos à URL base.

Quando o Serviço de Integração de Dados recebe uma resposta, ele transmite os dados na mensagem de resposta para as portas de saída da transformação. O Serviço de Integração de Dados transmite dados com base em como você configura a saída do método. As portas de saída contêm dados relacionais. O Serviço de Integração de Dados envia os dados nas portas de saída para transformações downstream no mapeamento ou para o destino.

## Configuração da Transformação de Consumidor de Serviço da Web REST

Ao criar uma transformação de Consumidor de Serviço da Web REST, você seleciona o método HTTP e define a entrada e a saída do método. Se você selecionar o método Get, não defina a entrada do método.

Os elementos de entrada no mapa de mensagens de solicitação HTTP são mapeados para portas de entrada. Os elementos de saída no mapa de mensagens de resposta HTTP são mapeados para portas de saída. A ferramenta Developer cria portas para os elementos de primeiro nível.

Ao configurar a transformação, você conclui as seguintes tarefas:

1. Seleciona o método HTTP.
2. Configura portas para representar elementos no cabeçalho e no corpo das mensagens de solicitação e de resposta.
3. Configura o mapeamento de entrada.
4. Configura o mapeamento de saída.
5. Configura propriedades avançadas, como a conexão e a URL base do serviço da Web.

Se o serviço da Web REST exigir autenticação, crie um objeto de conexão HTTP.

## Configuração de Mensagens

O Serviço de Integração de Dados gera mensagens de solicitação e interpreta mensagens de resposta com base na entrada e saída do método e nas portas que você configura na transformação de Consumidor de Serviço da Web REST.

Portas de entrada representam diferentes partes da mensagem de solicitação. É possível adicionar portas de entrada que identificam o recurso que você deseja recuperar ou alterar. Também é possível adicionar portas de entrada que representam cabeçalhos HTTP, informações de cookies e elementos na mensagem de solicitação.

Portas de saída representam elementos na mensagem de resposta que você deseja enviar para transformações downstream ou para o destino em um mapeamento. Você pode adicionar portas de saída que representam cabeçalhos HTTP, informações de cookies, o código de resposta e elementos na mensagem de resposta.

## Identificação do Recurso

Para identificar o recurso em uma solicitação HTTP, o Serviço de Integração de Dados acrescenta os valores em portas de entrada específica à URL base. Você define a URL base na conexão HTTP ou nas propriedades da transformação. Use portas de URL ou de argumentos para identificar um determinado recurso.

Use portas de URL quando o serviço da Web identifica um recurso por meio de uma string de caracteres exclusiva.

Por exemplo, o serviço da Web REST da HypoStores identifica peças por número de peça através da seguinte URL:

```
http://www.HypoStores.com/products/ProductDetails/<Part_No>
```

Para identificar uma peça, defina os seguintes detalhes da transformação:

1. Defina a URL base como a seguinte URL:

```
http://www.HypoStores.com/products/ProductDetails
```

2. Defina uma porta de URL e transmita o número da peça à transformação por meio dessa porta de URL.

Se o mapeamento transmitir o número da peça 500 à porta de URL, o Serviço de Integração de Dados usará a seguinte URL na mensagem de solicitação:

```
http://www.HypoStores.com/products/ProductDetails/500
```

Use portas de argumentos quando o serviço da Web identifica a localização de um recurso por meio de argumentos.

Por exemplo, você deseja transmitir um número de peça ao serviço da Web REST da HypoStores através do argumento "Part\_No".

Para identificar uma peça, defina os seguintes detalhes da transformação:

1. Defina a URL base como a seguinte URL:

```
http://www.HypoStores.com/products/ProductDetails
```

2. Crie uma porta de argumentos com o nome do argumento "Part\_No" e transmita o número da peça à transformação através da porta de argumentos.

Se o mapeamento transmitir o número da peça 600 à porta de argumentos, o Serviço de Integração de Dados usará a seguinte URL na mensagem de solicitação:

```
http://www.HypoStores.com/products/ProductDetails?Part_No=600
```

Crie várias portas de argumentos para definir vários argumentos. O Serviço de Integração de Dados separa cada argumento com um caractere de E comercial (&).

Por exemplo, você deseja recuperar detalhes de funcionários de um serviço da Web REST e transmitir o nome e o sobrenome dos funcionários através dos argumentos "First\_Name" e "Last\_Name". Crie portas de argumentos com os nomes dos argumentos "First\_Name" e "Last\_Name". Se o mapeamento transmitir o nome "John Smith" para a transformação, o Serviço de Integração de Dados usará uma URL como a seguinte na mensagem de solicitação:

```
http://www.HypoStores.com/employees/EmpDetails?First_Name=John&Last_Name=Smith
```

Se você não especificar uma URL ou uma porta de argumento, o Serviço de Integração de Dados usará a URL base das propriedades da transformação ou da conexão HTTP para identificar o recurso. A URL base na conexão HTTP substitui a URL base na transformação.

## Métodos HTTP

Ao criar uma transformação de Consumidor de Serviço da Web REST, selecione o método HTTP que usado pelo Serviço de Integração de Dados na mensagem de solicitação. Depois de criar a transformação, você não poderá alterar o método HTTP.

Você configura a transformação para usar um dos seguintes métodos HTTP:

### Obter

Recupera um recurso ou uma coleção de recursos do serviço da Web. Por exemplo, você pode recuperar uma tabela de produtos ou recuperar informações sobre um produto.

### Post

Envia dados para um serviço da Web. Use o método Post para criar um recurso ou uma coleção de recursos. Por exemplo, você pode adicionar os detalhes de uma nova transação da loja.

### Put

Substitui um recurso ou uma coleção de recursos. Se os dados não existirem, o método Put os publicará. Por exemplo, você pode atualizar o endereço de remessa de um cliente.

### Delete

Exclui um recurso ou uma coleção de recursos. Por exemplo, você pode excluir o registro de um funcionário que não trabalha mais em uma organização.

## Método Get HTTP

O Serviço de Integração de Dados usa o método Get HTTP para recuperar dados de um serviço da Web REST. Use o método Get para recuperar um recurso ou uma coleção de recursos.

Ao configurar a transformação de Consumidor de Serviço da Web REST para usar o método Get, você configura as portas de entrada, a saída do método e as portas de saída. Você não configura a entrada do método.

### Exemplo

Você deseja recuperar a descrição e o preço do número de peça 500 no banco de dados de produtos da HypoStores. O serviço da Web usa a seguinte URL para identificar uma peça:

`http://www.HypoStores.com/products/ProductDetails?Part_No=<Part_No>`

Insira a seguinte URL base:

`http://www.HypoStores.com/products/ProductDetails`

A seguinte tabela mostra a porta de entrada que você pode definir:

Tipo de Porta	Nome do Argumento	Valor de entrada
Argumento	Part_No	500

A tabela a seguir mostra as portas de saída que você pode definir:

Tipo de Porta	Nome da Porta	Valor de Retorno
Saída	Part_Desc	...<desc>ACME ball point pens, 12-pk, black, 0.7 mm</desc>...
Saída	Price_USD	...<price>9.89</price>...

## Método Post HTTP

O Serviço de Integração de Dados usa o método Post HTTP para enviar dados a um serviço da Web REST. O serviço da Web determina a função real realizada pelo método Post. É possível usar o método Post para criar um recurso ou uma coleção de recursos.

Ao configurar a transformação de Consumidor de Serviço da Web REST para usar o método Post, você configura as portas de entrada, a entrada do método, a saída do método e as portas de saída.

### Exemplo

Você deseja publicar a nova peça 501 no banco de dados de produtos da HypoStores. O serviço da Web usa a seguinte URL para a peça 501:

`http://www.HypoStores.com/products/ProductDetails/501`

Insira a seguinte URL base:

`http://www.HypoStores.com/products/ProductDetails`

A seguinte tabela mostra as portas de entrada que você pode definir:

Tipo de Porta	Nome da Porta	Valor de entrada
URL	URL_Part_No	501
Entrada	Part_Desc	ACME ball point pens, 12-pk, black, 0.5 mm
Entrada	Price_USD	9.89

A tabela a seguir mostra as portas de saída que você pode definir:

Tipo de Porta	Nome da Porta	Valor de Retorno
Saída	Resposta	<Resposta retornada pelo serviço da Web>

## Método Put HTTP

O Serviço de Integração de Dados usa o método Put HTTP para atualizar dados através de um serviço da Web REST. Use o método Post para atualizar um recurso ou uma coleção de recursos. Se os dados não existirem, o Serviço de Integração de Dados criará o recurso ou a coleção de recursos.

Ao configurar a transformação de Consumidor de Serviço da Web REST para usar o método Put, você configura as portas de entrada, a entrada do método, a saída do método e as portas de saída.

## Exemplo

Você deseja atualizar o preço unitário da peça 501 no banco de dados de produtos da HypoStores. O serviço da Web usa a seguinte URL para a peça 501:

`http://www.HypoStores.com/products/ProductDetails/501`

Insira a seguinte URL base:

`http://www.HypoStores.com/products/ProductDetails`

A seguinte tabela mostra as portas de entrada que você pode definir:

Tipo de Porta	Nome da Porta	Valor de entrada
URL	URL_Part_No	501
Entrada	Price_USD	9.99

A tabela a seguir mostra as portas de saída que você pode definir:

Tipo de Porta	Nome da Porta	Valor de Retorno
Saída	Resposta	<Resposta retornada pelo serviço da Web>

## Método Delete HTTP

O Serviço de Integração de Dados usa o método Delete HTTP para remover dados através de um serviço da Web REST. Use o método Delete para remover um recurso ou uma coleção de recursos.

Ao configurar a transformação de Consumidor de Serviço da Web REST para usar o método Delete, você configura as portas de entrada, a entrada do método, a saída do método e as portas de saída.

## Exemplo

Você deseja excluir o número de peça 502 do banco de dados de produtos da HypoStores. O serviço da Web usa a seguinte URL para identificar uma peça:

`http://www.HypoStores.com/products/ProductDetails?Part_No=<Part_No>`

Insira a seguinte URL base:

`http://www.HypoStores.com/products/ProductDetails`

A seguinte tabela mostra a porta de entrada que você pode definir:

Tipo de Porta	Nome do Argumento	Valor de entrada
Argumento	Part_No	502

A tabela a seguir mostra portas de saída que você pode definir:

Tipo de Porta	Nome da Porta	Valor de Retorno
Saída	Resposta	<Resposta retornada pelo serviço da Web>

# Portas da Transformação de Consumidor de Serviço da Web REST

Uma transformação de Consumidor de Serviço da Web REST pode ter várias portas de entrada e várias portas de saída. Você cria portas em grupos com base na estrutura da hierarquia XML ou JSON.

Ao exibir as portas da transformação, mostre-as se você não precisar exibir a hierarquia XML ou JSON. Ao mostrar as portas, você pode definir grupos, definir portas e mapear elementos da entrada e da saída do método para portas de entrada e de saída.

Uma transformação de Consumidor de Serviço da Web REST pode ter vários grupos de entrada e de saída. Ao criar portas, crie grupos e adicione as portas aos grupos. Defina as portas em uma hierarquia de grupos com base na estrutura da hierarquia de entrada ou de saída no XML ou JSON. Adicione uma chave para relacionar um grupo filho a um grupo pai.

Todos os grupos, exceto o grupo mais baixo na hierarquia, devem ter chaves primárias. Todos os grupos na hierarquia, exceto o grupo raiz, devem ter chaves externas.

A transformação tem um grupo de entrada raiz chamado RequestInput. Você deve adicionar uma chave primária ao grupo de entrada raiz. A chave deve ser string, bigint ou integer. É possível configurar qualquer porta no grupo de entrada raiz como uma porta de passagem.

Para mapear um elemento para uma porta, clique no campo da coluna **Localização** e expanda a hierarquia na caixa de diálogo **Selecionar Localização**. Em seguida, escolha um elemento na hierarquia.

## Portas de Entrada

Portas de entrada representam dados de uma transformação upstream ou de uma origem que você deseja transmitir ao serviço da Web. É possível configurar várias portas de entrada. Cada porta de entrada é mapeada para um elemento na mensagem de solicitação.

Para adicionar uma porta de entrada, selecione um grupo de entrada, clique na seta ao lado do botão **Novo** e selecione **Campo**.

## Portas de Saída

Portas de saída representam elementos na mensagem de resposta que você deseja transmitir para uma transformação downstream ou para o destino. É possível configurar várias portas de saída. Cada porta de saída é mapeada para um elemento na mensagem de resposta.

Para adicionar uma porta de saída, selecione um grupo de saída, clique na seta ao lado do botão **Novo** e selecione **Campo**.

## Portas de Passagem

Portas de passagem transmitem dados na transformação sem alterá-los. É possível configurar qualquer porta no grupo de entrada raiz como uma porta de passagem.

Para adicionar uma porta de passagem, adicione uma porta ao grupo de entrada raiz. Em seguida, clique com o botão direito do mouse na porta e selecione **Mapear**.



## Portas de Argumentos

Portas de argumentos permitem identificar um recurso quando a URL do recurso utiliza um argumento. Adicione portas de argumentos ao grupo de entrada raiz.

Uma porta de argumentos tem um nome de porta e um nome de argumento. Se um nome de argumento contiver um caractere não permitido em um nome de porta, insira um nome de argumento que seja diferente do nome da porta. Por exemplo, você deseja transmitir o argumento "Cust-ID" ao serviço da Web, mas o Serviço de Integração de Dados não permite o caractere de traço (-) em nomes de porta. Insira "Cust-ID" como o nome do argumento, mas insira "CustID" como o nome da porta.

O Serviço de Integração de Dados acrescenta os nomes e os valores de argumento para cada porta de argumentos à URL base como pares de nome/valor. É possível configurar várias portas de argumentos. O Serviço de Integração de Dados separa vários argumentos na solicitação com um caractere de E comercial (&).

Por exemplo:

```
http://www.HypoStores.com/customers/CustDetails?Last_Name=Jones&First_Name=Mary
```

Se você definir portas de argumentos e portas de URL na transformação, o Serviço de Integração de Dados acrescentará os valores de portas de URL à URL base, seguidos pelos nomes e valores de argumentos.

Para adicionar uma porta de argumentos, clique no grupo de entrada raiz e selecione **Novo > Portas de Argumentos**. Insira o nome do argumento e o nome da porta.

## Portas de URL

Portas de URL permitem que você identifique um recurso através de uma URL estática. Para identificar um recurso, o Serviço de Integração de Dados acrescenta o valor da porta de URL à URL base.

Por exemplo:

```
http://www.HypoStores.com/products/ProductDetails/<URL_port_value>
```

Adicione portas de URL ao grupo de entrada raiz.

É possível configurar várias portas de URL. O Serviço de Integração de Dados separa os valores em cada porta de URL com um caractere de barra (/). Se você definir portas de URL e portas de argumentos na transformação, o Serviço de Integração de Dados acrescentará os valores de portas de URL à URL base, seguidos pelos nomes e valores de argumentos.

Para adicionar uma porta de URL, clique com o botão direito do mouse no grupo de entrada raiz e selecione **Novo > Portas de URL**.

## Portas de Cabeçalho HTTP

Portas de cabeçalho HTTP representam cabeçalhos HTTP na mensagem de solicitação. É possível configurar várias portas de cabeçalho HTTP.

Para transmitir informações de cabeçalho ao serviço da Web na solicitação, adicione a porta ao grupo de entrada raiz. É possível configurar uma única porta de cabeçalho HTTP para o grupo de entrada raiz. Se você adicionar um cabeçalho HTTP ao grupo de entrada raiz, poderá configurá-lo como uma porta de passagem.

Uma porta de cabeçalho HTTP tem um nome de porta e um nome de cabeçalho HTTP. Se um nome de cabeçalho HTTP contiver um caractere não permitido em um nome de porta, insira um nome de cabeçalho HTTP que seja diferente do nome da porta. Por exemplo, você deseja transmitir o nome de cabeçalho "Content-Type" para o serviço da Web, mas o Serviço de Integração de Dados não permite o caractere de

traço (-) em nomes de porta. Insira "Content-Type" como o nome do cabeçalho HTTP, mas insira "ContentType" como o nome da porta.

Para adicionar uma porta de cabeçalho HTTP, clique no grupo de entrada raiz e selecione **Novo > Cabeçalho HTTP**. Insira um nome de cabeçalho e um nome de porta.

## Portas para Cookies

É possível configurar a transformação de Consumidor de Serviço da Web REST para usar a autenticação por cookies. O servidor da Web remoto rastreia os usuários consumidores de serviços da Web com base nos cookies. Você pode melhorar o desempenho quando um mapeamento chama um serviço da Web várias vezes.

Para transmitir informações de cookies ao serviço da Web na solicitação, adicione a porta ao grupo de entrada raiz. É possível configurar uma única porta para cookies para o grupo de entrada raiz. Se você adicionar uma porta para cookies ao grupo de entrada raiz, poderá configurá-la como uma porta de passagem.

Para extrair informações de cookies da resposta, adicione uma porta para cookies a um grupo de saída. É possível configurar uma porta para cookies para cada grupo de saída.

Quando a porta para cookies é projetada para uma mensagem de solicitação de serviço da Web, o provedor do serviço da Web retorna um valor de cookie na mensagem de resposta. Você pode transmitir o valor do cookie para outra transformação downstream no mapeamento ou pode salvar o valor do cookie em um arquivo. Ao salvar o valor do cookie em um arquivo, você pode configurar esse cookie como entrada para a transformação de Consumidor de Serviço da Web REST.

Para adicionar uma porta para cookies, clique com o botão direito do mouse no grupo de entrada raiz e selecione **Novo > Outras Portas**. Em seguida, selecione **Cookie** e clique em **OK**.

## Portas de Saída XML

Portas XML de saída representam respostas do serviço da Web. Portas XML de saída são portas de string.

Adicione uma porta XML de saída a um grupo de saída. É possível configurar uma única porta XML de saída para cada grupo de saída.

clique com o botão direito do mouse no grupo de entrada raiz e selecione **Novo > Outras Portas**. Em seguida, selecione **XML de Saída** e clique em **OK**.

## Portas de Código de Resposta

Portas de código de resposta representam os códigos de resposta HTTP do serviço da Web. Portas de código de resposta são portas de números inteiros.

Adicione uma porta de código de resposta a um grupo de saída. É possível configurar uma porta de código de resposta para cada grupo de saída.

Para adicionar uma porta de código de resposta, selecione um grupo de saída, clique com o botão direito do mouse no grupo de entrada raiz e selecione **Novo > Outras Portas**. Em seguida, selecione **Código de Resposta** e clique em **OK**.

# Mapeamento de Entradas da Transformação de Consumidor de Serviço da Web REST

Ao exibir as portas de transformação, mostre o mapeamento de entrada para exibir a hierarquia de entradas de método. Ao mostrar o mapeamento de entrada, você pode definir grupos e portas de entrada, bem como mapear portas de entrada para elementos de entrada de método.

A entrada de mapeamento inclui as seguintes áreas:

## Portas

Crie os grupos e as portas de entrada da transformação na área **Portas**.

## Entrada do Método

A área **Entrada do Método** mostra os elementos na mensagem de solicitação que a transformação de Consumidor de Serviço da Web REST envia para o serviço da Web. Se você usar um objeto de esquema para criar a transformação, esse objeto definirá a hierarquia de entradas de método.

Depois de criar portas de entrada, mapeie-as da área **Portas** para os elementos na área **Entrada do Método**. Quando você mapeia uma porta de entrada para um elemento na entrada do método, a localização dessa porta aparece na coluna Localização da área **Entrada do Método**.

A ferramenta Developer mapeia elementos no primeiro nível da entrada do método para portas de entrada quando você opta por mapear o primeiro nível da hierarquia de entradas. A ferramenta Developer também cria as portas para executar o mapeamento. Se o primeiro nível da hierarquia contiver um elemento pai de ocorrência múltipla com um ou mais elementos filho de ocorrência múltipla, a ferramenta Developer não mapeará o primeiro nível da hierarquia.

É possível optar por exibir as linhas que conectam as portas de entrada aos elementos na entrada do método.

## Regras e Diretrizes para Mapear Portas de Entrada para Elementos

Analise as seguintes regras ao mapear portas de entrada para elementos na hierarquia de entradas de método:

- É possível mapear uma porta de entrada para um elemento na hierarquia. Você pode mapear a mesma porta para qualquer número de chaves na hierarquia.
- A porta de entrada e o elemento devem ter tipos de dados compatíveis.
- É possível mapear portas de um grupo de entrada para vários níveis de hierarquia na entrada do método.
- Portas de entrada devem ser mapeadas para as chaves na entrada do método. Qualquer porta mapeada para uma chave deve ter um tipo de dados de cadeia, número inteiro ou bigint. Mapeie dados para as chaves em todos os níveis da entrada de método acima do nível de hierarquia que você está incluindo na mensagem de solicitação. Inclua as chaves externas para todos os níveis acima do nível que está sendo mapeado (inclusive).

**Nota:** Você não precisará mapear portas de entrada para chaves se estiver mapeando somente o nível mais inferior da hierarquia de entradas de método.

- Você deve mapear o elemento raiz RequestInput para o elemento filho do grupo Rest\_Consumer\_input para a definição de entrada do método.
- É possível mapear várias portas de entrada do tipo cadeia, bigint ou número inteiro para uma chave na área **Entrada do Método** para criar uma chave composta. Ao clicar no campo **Localização** para uma chave composta, você pode reordenar as portas de entrada ou remover uma das portas.

- Se o serviço da Web produzir um documento JSON, certifique-se de que xmlRoot seja o primeiro nó da hierarquia de respostas. Se xmlRoot não for o primeiro nó de um serviço da Web com uma resposta JSON, poderão ser exibidos valores nulos.

## Mapeando Portas de Entrada para a Entrada do Método

Ao mostrar o mapeamento de entrada de transformação, você pode definir grupos e portas de entrada, bem como mapear portas de entrada para elementos de entrada de método.

1. Abra uma transformação de Consumidor de Serviço da Web REST.
2. Na exibição **Portas**, mostre o mapeamento de entrada.
3. Defina uma chave primária para o grupo de entrada raiz.
4. Para adicionar um grupo de entrada ou porta à área **Portas**, use um dos seguintes métodos:

Método	Descrição
Arraste um elemento.	Arraste um grupo ou um elemento filho da área <b>Entrada do Método</b> até uma coluna vazia na área <b>Portas</b> . Se você arrastar um grupo até a área <b>Portas</b> , a ferramenta Developer adicionará um grupo sem portas.
Adicione manualmente um grupo ou uma porta.	Para adicionar um grupo, clique na seta ao lado do botão <b>Novo</b> e depois clique em <b>Grupo</b> . Para adicionar uma porta, clique na seta ao lado do botão <b>Novo</b> e depois clique em <b>Campo</b> .
Arraste uma porta a partir de outra transformação.	No editor, arraste uma porta de outra transformação até a transformação de Consumidor de Serviço da Web REST.
Copie uma porta.	Selecione portas de outra transformação e copie-as para a área <b>Portas</b> . Para copiar portas, você pode usar atalhos de teclado ou os botões <b>Copiar</b> e <b>Colar</b> na ferramenta Developer.
Selecione <b>Mapear primeiro nível da hierarquia</b> .	A ferramenta Developer mapeia elementos no primeiro nível da entrada do método para portas de entrada e grupos. A ferramenta Developer também cria as portas de entrada e os grupos para executar o mapeamento.

5. Se você criar manualmente uma porta ou copiar uma porta de outra transformação, clique na coluna **Localização** na área **Entrada do Método** e escolha uma porta na lista.
6. Para mapear portas de entrada como uma chave composta, use um dos seguintes métodos:

Método	Descrição
Arraste portas de entrada.	Selecione duas ou mais portas de entrada e arraste-as até uma chave na hierarquia de entradas de método.
Selecione portas de entrada na caixa de diálogo <b>Selecionar Localização</b> .	Clique na coluna <b>Localização</b> de uma chave na hierarquia de entradas de método e selecione as portas de entrada.

7. Para limpar as localizações dos elementos, use um dos seguintes métodos:

Método	Descrição
Clique em <b>Limpar</b> .	Selecione um ou mais elementos na área <b>Entrada do Método</b> e clique em <b>Limpar</b> .
Exclua as linhas que conectam portas a elementos.	Selecione uma ou mais linhas que conectam as portas de entrada aos elementos na entrada do método e pressione <b>Delete</b> .

## Mapeamento da Saída da Transformação de Consumidor de Serviço da Web REST

Ao exibir as portas de transformação, mostre o mapeamento de saída para exibir a hierarquia de saídas de método. Ao mostrar o mapeamento de saída, você pode definir grupos e portas de saída, além de mapear elementos de saída de método para portas de saída.

O mapeamento de saída inclui as seguintes áreas:

### Saída do Método

A área **Saída do Método** mostra os elementos na mensagem de resposta que o serviço da Web retorna para a transformação de Consumidor de Serviço da Web REST. Se você usar um objeto de esquema para criar a transformação, o objeto de esquema definirá a hierarquia de saídas de método.

### Portas

Crie os grupos e portas de saída da transformação na área **Portas**.

Depois de criar portas de saída, mapeie os elementos da área **Saída do Método** para as portas na área **Portas**. Quando você mapeia um elemento da saída do método para uma porta de saída, a localização desse elemento aparece na coluna **Localização** da área **Portas**.

A ferramenta Developer mapeia elementos no primeiro nível da saída do método para portas de saída quando você opta por mapear o primeiro nível da hierarquia de saída. A ferramenta Developer também cria as portas para executar o mapeamento. Se o primeiro nível da hierarquia contiver um elemento pai de ocorrência múltipla com um ou mais elementos filho de ocorrência múltipla, a ferramenta Developer não mapeará o primeiro nível da hierarquia.

Você pode optar por exibir as portas de saída em uma hierarquia. Cada grupo filho aparece sob o grupo pai. Você também pode optar por exibir as linhas que conectam os elementos na saída do método às portas de saída.

Se o objeto de esquema associado for excluído do repositório, a ferramenta Developer manterá a localização dos elementos do método no mapeamento de saída. Quando você mostra o mapeamento de saída, a área **Portas** ainda exibe a localização dos elementos do método na coluna **Localização** para as portas de saída. Se você associar outro esquema à transformação, a ferramenta Developer verificará se cada localização é válida. A ferramenta Developer limpa a localização dos elementos do método na área **Portas** do mapeamento de saída se essa localização não for mais válida.

## Regras e Diretrizes para Mapear Elementos para Portas de Saída

Analise as seguintes regras ao mapear elementos na hierarquia de saídas de método para portas de saída:

- O elemento de saída do método e a porta de saída devem ter tipos de dados compatíveis.
- Não é possível mapear um elemento para mais de uma porta de saída em um grupo.
- Cada porta de saída deve ter uma localização válida, a menos que a porta seja uma porta de passagem.
- Se você arrastar um elemento filho de ocorrência múltipla até uma porta de saída vazia, será necessário relacionar o grupo a outros grupos de saída. Quando você seleciona um grupo, a Developer tool cria chaves para relacionar os grupos.
- Ao arrastar um elemento de ocorrência múltipla até um grupo que contém o elemento pai, você pode configurar o número de ocorrências do elemento filho para inclusão. Ou você pode substituir o grupo pai pelo grupo filho com ocorrência múltipla na saída de transformação.
- Se o serviço da Web produzir um documento JSON, certifique-se de que xmlRoot seja o primeiro nó da hierarquia de respostas. Se xmlRoot não for o primeiro nó de um serviço da Web com uma resposta JSON, poderão ser exibidos valores nulos nas portas de saída.

## Personalizar Opções de Exibição

Você pode alterar a hierarquia de saídas de método para mostrar as portas para cookies, portas de passagem e chaves na área **Saída do Método**. Você também pode mostrar construções de agrupamento que definem como ordenar os elementos.

Clique em **Personalizar Exibição** na área **Saída do Método**. Ative qualquer uma das seguintes opções:

### Sequência, Escolha e Tudo

Mostre uma linha que indica se uma definição de elemento é sequência, escolha ou tudo.

Os elementos em um grupo de sequência devem estar na ordem especificada na hierarquia.

Pelo menos um elemento em um grupo de escolha deve aparecer na mensagem de resposta.

Os elementos em um grupo Todos devem estar todos incluídos na mensagem de resposta.

### Chaves

Exiba as chaves na área **Saída do Método**. A área **Saída da Operação** inclui chaves para cada grupo. É possível adicionar uma chave a uma porta de saída na área **Portas**.

### Portas de Passagem

A área **Saída do Método** mostra as portas de passagem. As portas de passagem são portas que passam dados pela transformação sem alterá-los. É possível projetar portas de passagem da saída do método para qualquer um dos grupos de saída da transformação de Consumidor de Serviço da Web REST. Uma porta de passagem recebe dados uma única vez, para que a porta esteja no nível raiz nas mensagens de resposta.

## Mapeando a Saída do Método para Portas de Saída

Ao mostrar o mapeamento de saída de transformação, você pode definir grupos e portas de saída, além de mapear elementos de saída de método para portas de saída.

1. Abra uma transformação de Consumidor de Serviço da Web REST.
2. Na exibição **Portas**, mostre o mapeamento de saída.

3. Para adicionar um grupo ou porta de saída à área **Portas**, use um dos seguintes métodos:

Método	Descrição
Arraste um elemento.	Arraste um grupo ou um elemento filho na área <b>Saída do Método</b> até uma coluna vazia na área <b>Portas</b> . Se você arrastar um grupo até a área <b>Portas</b> , a ferramenta Developer adicionará um grupo sem portas.
Adicione manualmente um grupo ou uma porta.	Para adicionar um grupo, clique na seta ao lado do botão <b>Novo</b> e clique em <b>Grupo</b> . Para adicionar uma porta, clique na seta ao lado do botão <b>Novo</b> e depois clique em <b>Campo</b> .
Arraste uma porta a partir de outra transformação.	No editor, arraste uma porta de outra transformação até a transformação de Consumidor de Serviço da Web REST.
Copie uma porta.	Selecione portas de outra transformação e copie-as para a área <b>Portas</b> . Para copiar portas, você pode usar atalhos de teclado ou os botões <b>Copiar</b> e <b>Colar</b> na ferramenta Developer.

4. Se você criar manualmente uma porta ou copiar uma porta de outra transformação, clique na coluna **Localização** da área **Portas** e escolha um elemento na lista.
5. Para limpar as localizações das portas, use um dos seguintes métodos:

Método	Descrição
Clique em <b>Limpar</b> .	Selecione uma ou mais portas na área <b>Portas</b> e clique em <b>Limpar</b> .
Exclua as linhas que conectam elementos a portas.	Selecione uma ou mais linhas que conectam os elementos na saída do método às portas de saída e pressione <b>Delete</b> .

## Propriedades Avançadas da Transformação de Consumidor de Serviço da Web REST

Configure propriedades que ajudam a determinar como o Serviço de Integração de Dados processa dados para a transformação de Consumidor de Serviço da Web REST.

Configure as seguintes propriedades na guia **Avançado**:

### Nível de Rastreamento

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.

### Conexão

Identifica o objeto de conexão HTTP para se conectar ao serviço da Web. Crie e edite a conexão HTTP na ferramenta Developer. Ao configurar uma conexão HTTP, configure a URL base, o tipo de segurança necessário para o serviço da Web e um período de tempo limite de conexão.

A transformação de Consumidor de Serviço da Web REST se conecta a um serviço da Web usando uma URL. Você pode definir a URL nas propriedades da transformação ou na conexão HTTP.

Configure uma conexão HTTP nas seguintes circunstâncias:

- Você não usa uma porta de entrada de URL.
- O serviço da Web exige autenticação HTTP ou certificados SSL.
- Você deseja alterar o período de tempo limite de conexão padrão.

#### Validação de Esquema XML

Valida a mensagem de resposta em tempo de execução. **Selecione Erro em XML Inválido ou Nenhuma Validação.**

#### Entrada Classificada

Permite que o Serviço de Integração de Dados gere a saída sem processar todos os dados de entrada. Ative a entrada classificada quando os dados de entrada forem classificados de acordo com as chaves na hierarquia de entrada XML.

#### URL

A URL base do serviço da Web REST. A URL base na conexão HTTP substitui esse valor.

#### Formatar

O formato da resposta do serviço da Web. Selecione XML ou JSON dependendo da resposta do serviço da Web.

## Criação da Transformação de Consumidor de Serviço da Web REST

É possível criar uma transformação de Consumidor de Serviço da Web REST reutilizável ou não reutilizável. Transformações reutilizáveis podem existir em vários mapeamentos. As transformações não reutilizáveis existem dentro de um único mapeamento.

Ao criar uma transformação de Consumidor de Serviço da Web REST, você pode definir os elementos e a hierarquia XML manualmente ou pode importar os elementos e a hierarquia de um objeto de esquema. O objeto de esquema pode ser um arquivo XML ou um arquivo de texto.

## Criando uma Transformação Consumidor de Serviço da Web REST

Quando você criar uma transformação Consumidor de Serviço da Web REST, selecione um método e defina o método de entrada e de saída com base no método que você escolher.

1. Para criar uma transformação Consumidor de Serviço da Web REST, use um dos seguintes métodos:

Método	Descrição
Reutilizável	Selecione um projeto ou pasta na exibição Object Explorer. Clique em <b>Arquivo &gt; Novo &gt; Transformação</b> . Selecione a transformação Consumidor de Serviço da Web REST e clique em <b>Avançar</b> .
Não reutilizável	Em um mapeamento ou maplet, arraste uma transformação Consumidor de Serviço da Web REST da paleta de transformação para o editor de mapeamento ou maplet.

2. Digite o nome da transformação e selecione a localização e o método de HTTP.



3. Clique em **Avançar**.
4. Para definir o método de entrada, use um dos métodos a seguir:

Método	Descrição
Criar como vazio	Defina os elementos de XML e a hierarquia manualmente.
Criar a partir de um elemento em um Objeto de Esquema	Importe os elementos de XML e de hierarquia de um objeto de esquema.

A área mostra os grupos de entrada e as portas de entrada da transformação. A área **Mapeamento de entrada** mostra a hierarquia da mensagem de solicitação.

5. Defina os grupos e as portas de entrada e mapeie as portas de entrada para os elementos de entrada.
6. Clique em **Avançar**.
7. Para definir o método de saída, selecione **Criar como vazio** ou **Criar de um elemento em um Objeto de Esquema**.

A área mostra os grupos e as portas de entrada da transformação. A área mostra a hierarquia da mensagem de solicitação.

8. Defina os grupos e as portas de saída e mapeie os elementos para as portas de saída.
9. Clique em **Concluir**.

## Analizando uma mensagem de resposta JSON que contém matrizes

Quando o elemento é um filho de tipo complexo, e a ocorrência máxima desse elemento não é limitada, o esquema não é válido. O analisador JSON impede que você extraia várias instâncias de um elemento.

A ocorrência máxima de elementos filho no tipo complexo deve ser 0 ou 1 com o indicador de ordem como escolha para o tipo complexo em um esquema. Quando você altera a ocorrência máxima como 1 para validar o esquema, é possível extrair uma única instância do elemento de cada vez.

É possível usar a ocorrência máxima como não limitada no indicador de ordem de escolha de um tipo complexo no esquema.

### Exemplo de mensagem de resposta JSON

Você tem o esquema a seguir, em que o elemento de tipo complexo `xmlRoot` tem o nome de elemento `Likes`, cuja ocorrência máxima não está limitada:

```
<xs:schema attributeFormDefault="unqualified" elementFormDefault="qualified"
  xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="xmlRoot">
    <xs:complexType>
      <xs:all>
        <xs:element type="xs:byte" name="Age"/>
        <xs:element type="xs:string" name="FirstName"/>
        <xs:element type="xs:string" name="Likes" maxOccurs="unbounded" minOccurs="0"/>
        <xs:element type="xs:string" name="FamilyName"/>
      </xs:all>
    </xs:complexType>
  </xs:element>
</xs:schema>
```

```

</xs:element>
</xs:schema>

```

Você pode alterar a resposta JSON no seguinte formato:

```

<xs:schema attributeFormDefault="unqualified" elementFormDefault="qualified"
xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="xmlRoot">
    <xs:complexType>
      <xs:choice maxOccurs="unbounded">
        <xs:element type="xs:byte" name="Age"/>
        <xs:element type="xs:string" name="FirstName"/>
        <xs:element type="xs:string" name="Likes" />
        <xs:element type="xs:string" name="FamilyName"/>
      </xs:choice>
    </xs:complexType>
  </xs:element>
</xs:schema>

```

`<xs:choice maxOccurs="unbounded">` permite que o conteúdo seja repetido uma ou mais vezes, em qualquer ordem.

## Matrizes não nomeadas em uma mensagem de resposta

Uma transformação de Consumidor de Serviço da Web REST oferece suporte a matrizes não nomeadas apenas em uma mensagem de resposta, mas não em uma mensagem de solicitação. Para analisar um esquema de array não nomeado especificado na Definição de Saída do Método, o elemento pai de `complexType` ou elementos de array de tipo simples devem ter o nome `xmlRoot`.

Em uma transformação de Consumidor de Serviço da Web REST, você deve definir `xmlRoot` como o elemento filho do elemento `xmlRoot` com o máximo de ocorrências definido como sem limites e os elementos no array não nomeado como elementos filho do elemento `xmlRoot`.

A seguinte imagem mostra a saída do método definida para o array não nomeado:

☐ Ports ☐ Method input ☒ Method output

Show: ☒ Method output definition ☐ Output mapping

Method output definition

	Name	Type	Min...	Ma...	Description	>>
	<b>Rest_Consume...</b>	(Rest_Cons...				
	xmlRoot	(xmlRoot)	1	1		
	xmlRoot	(xmlRoot)	1	Un...		
	emp	xs:string	1	1		
	empid	xs:string	1	1		

## CAPÍTULO 39

# Transformação de Roteador

Este capítulo inclui os seguintes tópicos:

- [Transformação de Roteador - Visão Geral, 599](#)
- [Transformações de Roteador em Mapeamentos Dinâmicos, 600](#)
- [Trabalhando com Grupos, 601](#)
- [Trabalhando com Portas, 604](#)
- [Conectando Transformações de Roteador em um Mapeamento, 605](#)
- [Transformação de Roteador - Propriedades Avançadas, 605](#)
- [Transformação de Roteador em um ambiente não nativo, 606](#)

## Transformação de Roteador - Visão Geral

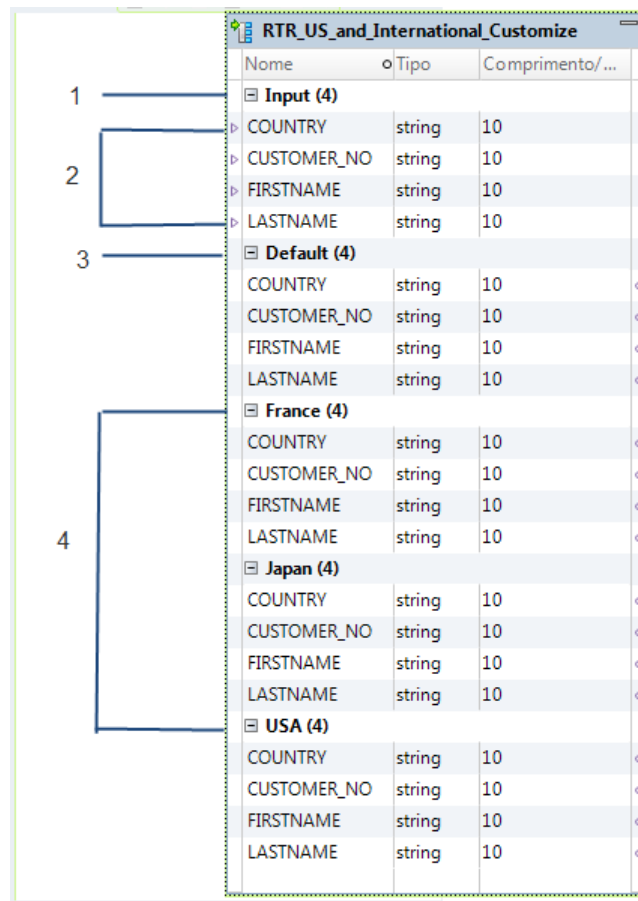
A transformação de Roteador é uma transformação ativa que encaminha dados em vários grupos de saída com base em uma ou mais condições. Encaminhe os grupos de saída para diferentes transformações ou destinos no mapeamento.

Uma transformação de Roteador é semelhante a uma transformação de Filtro, pois ambas usam uma condição para testar os dados. A transformação Filtro testa dados para uma condição e descarta as linhas de dados que não atenderem à condição. Uma transformação de Roteador testa os dados para um ou mais condições e pode rotear linhas de dados que não atendam a nenhuma das condições para um grupo de saída padrão.

Se você precisar testar os mesmos dados de entrada com base em várias condições, utilize uma transformação Roteador em um mapeamento, em vez de criar várias transformações Filtro para executar a mesma tarefa. A transformação Roteador é mais eficiente. Por exemplo, para testar dados com base em três condições, você pode usar uma transformação de Roteador, em vez de três transformações de Filtro. Quando você usa uma transformação de Roteador em um mapeamento, o Serviço de Integração de Dados processa os dados de entrada uma única vez. Quando você usa várias transformações de Filtro em um mapeamento, o Serviço de Integração de Dados processa os dados de entrada para cada transformação.

Uma transformação de Roteador consiste em grupos de entrada e saída, portas de entrada e saída, condições de filtro de grupos e propriedades avançadas que você configura na Developer tool.

A seguinte figura mostra um exemplo de transformação de Roteador e seus componentes:



1. Grupo de entrada
2. Portas de entrada
3. Grupo de saída padrão
4. Grupos de saída definidos pelo usuário

## Transformações de Roteador em Mapeamentos Dinâmicos

É possível usar uma transformação de Roteador em um mapeamento dinâmico. Você pode configurar portas dinâmicas na transformação e fazer referência às portas geradas na condição de filtro de grupo.

Quando você usa uma porta dinâmica na condição de filtro de grupo, essa porta dinâmica pode conter mais de uma porta gerada. A condição de filtro de grupo se expande para incluir cada porta gerada. Cada porta gerada deve ser um tipo válido para a expressão.

É possível parametrizar a condição de filtro de grupo. Use um parâmetro de tipo de expressão para especificar o filtro.

# Trabalhando com Grupos

Uma transformação Roteador tem os seguintes tipos de grupos:

- Entrada
- Saída

## Grupo de Entrada

A transformação de Classificação inclui um único grupo de entrada. O grupo de entrada inclui todas as portas de entrada que você adiciona à transformação.

## Grupos de Saída

A transformação de Classificação inclui os seguintes tipos de grupos de saída:

### Grupos Definidos pelo Usuário

Você cria um grupo definido pelo usuário para testar uma condição com base em dados de entrada. Um grupo definido pelo usuário é formado por portas de saída e por uma condição de filtro de grupo. Você pode criar e editar grupos definidos pelo usuário na exibição **Grupos** com a ferramenta Developer. Crie um grupo definido pelo usuário para cada condição que deseje especificar.

O Serviço de Integração de Dados utiliza a condição para avaliar cada linha de dados de entrada. Antes de processar o grupo padrão, ele testa as condições de cada grupo definido pelo usuário. O Serviço de Integração de Dados determina a ordem de avaliação de cada condição com base na ordem dos grupos de saída conectados. O Serviço de Integração de Dados processa grupos definidos pelo usuário que estejam conectados a uma transformação ou a um destino em um mapeamento.

Se uma linha atender a mais de uma condição de filtro de grupos, o Serviço de Integração de Dados passará essa linha várias vezes.

### O Grupo Padrão

A ferramenta Developer cria o grupo padrão depois que você cria um grupo definido pelo usuário. Não é permitido editar ou excluir o grupo padrão. Esse grupo não tem uma condição de filtro de grupo associada. Se todas as condições do grupo forem avaliadas como FALSE, o Serviço de Integração de Dados passará a linha para o grupo padrão. Se você quiser que o Serviço de Integração de Dados descarte todas as linhas do grupo padrão, não o conecte a uma transformação, nem a um destino no mapeamento.

A ferramenta Developer exclui o grupo padrão quando você exclui da lista o último grupo definido pelo usuário.

A ferramenta Developer copia as informações de propriedade das portas de entrada do grupo de entrada para criar um conjunto de portas de saída para cada grupo de saída. Não é possível alterar ou excluir as portas de saída ou suas propriedades.

## Usando Condições de Filtro de Grupos

Você pode testar os dados com base em uma ou mais condições de filtro de grupos. Você cria condições do filtro de grupo na guia **Grupos** usando o Editor de Expressão.

Você pode inserir qualquer expressão que retorne um único valor. Você pode ainda especificar uma constante para a condição. Uma condição de filtro de grupos retornará TRUE ou FALSE para cada linha que passar pela transformação se a linha satisfizer a condição especificada. Zero (0) equivale a FALSE. Qualquer

valor diferente de zero equivale a TRUE. É possível usar uma única porta numérica como condição de filtro. O Serviço de Integração de Dados passa as linhas de dados que são avaliadas como TRUE para cada transformação ou destino associado a cada grupo definido pelo usuário.

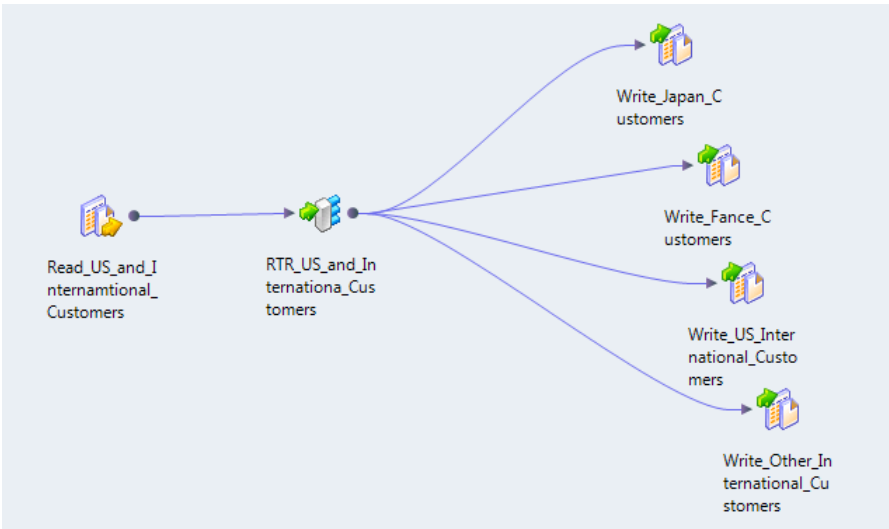
**Nota:** Não é possível usar uma única porta dinâmica para retornar um valor booleano.

Por exemplo, você tem clientes de nove países e quer realizar diferentes cálculos sobre os dados de três países. Você pode usar uma transformação de Roteador em um mapeamento para filtrar esses dados para três transformações de Expressão diferentes.

É possível usar parâmetros como elementos na condição de filtro de grupo. Você pode usar parâmetros do sistema ou parâmetros definidos pelo usuário. É possível criar parâmetros no Editor de Expressão e adicioná-los à expressão.

O grupo padrão não tem uma condição de filtro de grupos. Mas você pode criar uma transformação Expressão para realizar cálculos com base nos dados dos outros seis países.

A seguinte figura mostra um mapeamento com a transformação do Roteador, que filtra dados com base em várias condições:



Para realizar vários cálculos com base nos dados de três países diferentes, crie três grupos definidos pelo usuário e especifique três condições de filtro de grupos na guia **Grupos**.

A seguinte tabela mostra as condições do filtro de grupo que filtra dados de clientes:

Propriedades		
Visualizador de Dados		
Marcas		
Geral Portas <b>Grupos</b> Avançado	Grupos de Saída:	
	Nome do Grupo	Condição de Filtro de Grupos
	Default	
	France	COUNTRY='France'=TRUE
	Japan	COUNTRY='Japan'=TRUE
	USA	COUNTRY='USA'=TRUE
	Descrição:	

A tabela a seguir mostra as condições do filtro que filtram dados de clientes:

Nome do Grupo	Condição do Filtro de Grupo
França	customer_name='France'=TRUE
Japão	customer_name='Japan'=TRUE
EUA	customer_name='USA'=TRUE

No mapeamento, o Serviço de Integração de Dados passa as linhas de dados que avaliam como TRUE para cada transformação ou destino associado a cada grupo definido pelo usuário, como Japão, França e EUA. O Serviço de Integração de Dados passará a linha para o grupo padrão se todas as condições forem avaliadas como FALSE. Então, o Serviço de Integração de Dados passa os dados dos outros seis países para a transformação ou destino associado ao grupo padrão. Se você quiser que o Serviço de Integração de Dados descarte todas as linhas do grupo padrão, não o conecte a uma transformação, nem a um destino no mapeamento.

A transformação Roteador passa dados através de cada grupo que atenda à condição. Se os dados atenderem a três condições do grupo de saída, a transformação de Roteador os passará através de três grupos de saída.

Por exemplo, você configura as seguintes condições de grupo em uma transformação Roteador:

Nome do Grupo	Condição do Filtro de Grupo
Grupo de Saída 1	employee_salary > 1000
Grupo de Saída 2	employee_salary > 2000

Quando a transformação Roteador processa dados de uma linha de entrada com employee\_salary=3000, roteia os dados através dos grupos de saída 1 e 2.

## Portas Dinâmicas em Condições de Filtro de Grupos

É possível usar uma porta dinâmica em uma condição de filtro de grupo. O Serviço de Integração de Dados aplica a condição de filtro a cada porta gerada na porta dinâmica.

Por exemplo, uma porta dinâmica, MyDynamicPort, contém três portas decimais:

```
Salary  
Bonus  
Stock
```

Você pode configurar a seguinte condição de filtro de grupo:

```
MyDynamicPort > 100
```

A condição de filtro de grupo se expande para a seguinte expressão:

```
Salary > 100 AND Bonus > 100 AND Stock > 100
```

Cada porta gerada deve ser um tipo válido para a expressão.

## Parametrizar o Filtro de Grupo

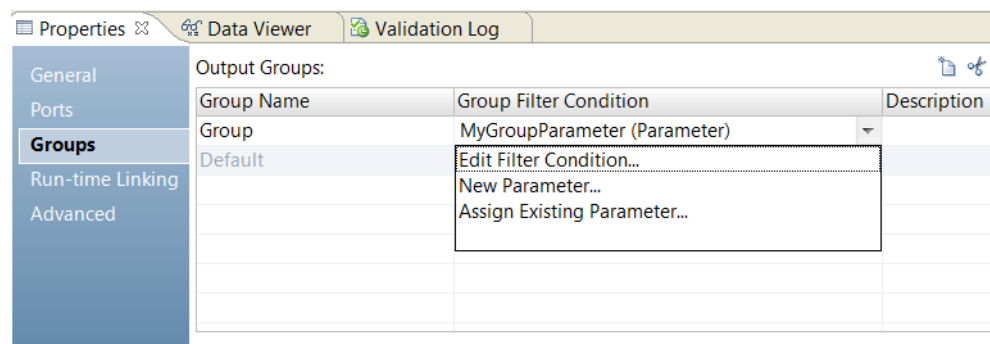
Você pode usar um parâmetro de expressão para definir um filtro de grupo. Um parâmetro de expressão é um parâmetro que contém uma expressão completa.

Por exemplo, um parâmetro de expressão pode conter a expressão

```
Employee_Salary > 1000.
```

Para parametrizar o filtro de grupo, escolha **Novo Parâmetro** no campo **Condição de Filtro de Grupos** da guia **Grupos**. Defina o parâmetro e insira a expressão no Editor de Expressão. Um parâmetro de expressão não pode conter outros parâmetros.

A seguinte imagem mostra a guia **Grupos** nas propriedades da transformação:



Quando você usa um parâmetro de expressão, a Developer tool não pode validar a expressão. Se a expressão não for válida em tempo de execução, o mapeamento poderá falhar.

## Adicionando Grupos

Quando você adiciona um grupo, a ferramenta Developer copia as informações de propriedade das portas de entrada para as portas de saída.

1. Clique na guia **Grupos**.
2. Clique no botão **Novo**.
3. Insira um nome para o grupo na seção **Nome do Grupo**.
4. Clique no campo **Condição de Filtro de Grupos** para abrir o **Editor de Expressão**.
5. Insira a condição do filtro de grupo.
6. Clique em **Validar** para verificar a sintaxe da condição.
7. Clique em **OK**.

## Trabalhando com Portas

Uma transformação de Roteador tem portas de entrada e de saída. As portas de entrada estão no grupo de entrada; as de saída estão nos grupos de saída.

Você pode criar portas de entrada copiando-as de outra transformação ou criando-as manualmente na guia **Portas**.



A ferramenta Developer cria portas de saída copiando as seguintes propriedades das portas de entrada:

- Nome da porta
- Tipo de dados
- Precisão
- Escala
- Valor padrão

Quando você faz alterações nas portas de entrada, a ferramenta Developer atualiza as portas de saída para refletir essas alterações. Não é possível editar ou excluir portas de saída.

A ferramenta Developer cria nomes de porta de saída com base nos nomes das portas de entrada. Para cada porta de entrada, a ferramenta Developer cria uma porta de saída correspondente em cada grupo de saída.

## Conectando Transformações de Roteador em um Mapeamento

Quando você conectar as transformações a uma transformação Roteador em um mapeamento, considere as regras a seguir:

- Você pode conectar um grupo a uma transformação ou destino.
- Você pode conectar uma porta de saída em um grupo a múltiplas transformações ou destinos.
- Você pode conectar múltiplas portas de saída em um grupo a múltiplas transformações ou destinos.
- Você não pode conectar mais de um grupo a um destino ou a uma única transformação de grupo de entrada.
- Você pode conectar mais de um grupo a uma transformação com vários grupos de entrada, exceto nas transformações Separador, quando conectar cada grupo de saída a um grupo de entrada diferente.

## Transformação de Roteador - Propriedades Avançadas

Configure propriedades avançadas para ajudar a determinar como o Serviço de Integração de Dados processa dados para a transformação de Roteador.

Você pode configurar níveis de rastreamento para logs.

Configure a seguinte propriedade na guia **Avançado**:

### Nível de Rastreamento

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.

# Transformação de Roteador em um ambiente não nativo

O processamento da transformação de Roteador em um ambiente não nativo depende do mecanismo que executa a transformação.

Considere o suporte para os seguintes mecanismos de tempo de execução não nativos:

- Mecanismo Blaze. Suportado sem restrições.
- Mecanismo Spark. Com suporte sem restrições em mapeamentos de fluxo e em lote.
- Mecanismo Databricks Spark. Suportado sem restrições.

## CAPÍTULO 40

# Transformação de Gerador de Sequência

Este capítulo inclui os seguintes tópicos:

- [Visão Geral da Transformação de Gerador de Sequência, 607](#)
- [Portas do Gerador de Sequência, 607](#)
- [Propriedades da Transformação Gerador de Sequência, 610](#)
- [Objeto de Dados de Sequência, 612](#)
- [Criando uma Transformação de Gerador de Sequência, 615](#)
- [Perguntas frequentes, 616](#)
- [Transformação de Gerador de Sequência em um ambiente não nativo, 617](#)

## Visão Geral da Transformação de Gerador de Sequência

A transformação de Gerador de Sequência é uma transformação passiva que gera valores numéricos. Use a transformação Gerador de Sequência para criar valores de chave primária exclusivos, substituir chaves primárias ausentes ou percorrer uma faixa sequencial de números.

O Gerador de Sequência contém portas de passagem e uma porta de saída. Você conecta a porta NEXTVAL às portas de entrada de outras transformações. O Serviço de Integração incrementa a sequência quando o mapeamento é executado.

Você pode criar uma transformação de Gerador de Sequência com base em uma nova sequência ou em um objeto de dados de sequência. Um objeto de dados de sequência é um objeto que cria e mantém uma sequência de valores.

## Portas do Gerador de Sequência

A transformação de Gerador de Sequência tem portas de passagem e uma porta de saída, NEXTVAL. Não é possível editar ou excluir a porta de saída.

## Portas de Passagem

Você pode adicionar uma porta à transformação de Gerador de Sequência como uma porta de passagem. As portas de passagem são portas de entrada e de saída que recebem dados de entrada e retornam os mesmos dados a um mapeamento sem alterá-los.

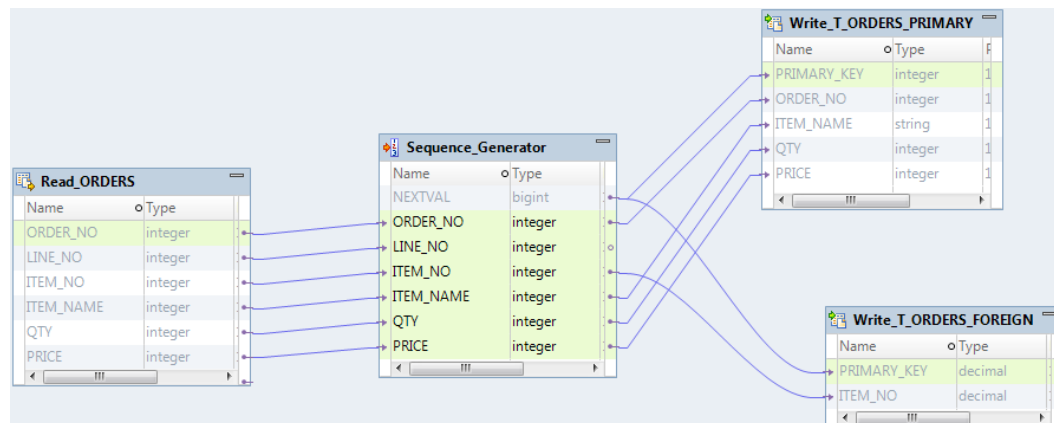
Você deve adicionar pelo menos uma porta de entrada à transformação e conectá-la a uma transformação ou origem upstream antes de vincular a porta de saída NEXTVAL aos destinos. Para adicionar uma porta de passagem à transformação, arraste uma porta de uma transformação ou origem upstream no mapeamento até a transformação de Gerador de Sequência.

## Porta NEXTVAL

Você pode conectar NEXTVAL a uma transformação para gerar valores exclusivos para cada linha na transformação. Conecte a porta NEXTVAL a um destino ou transformação downstream para gerar uma sequência de números. Se você conectar NEXTVAL a várias transformações, o Serviço de Integração gerará a mesma sequência de números para cada transformação.

Você conecta a porta NEXTVAL para gerar a sequência com base nas propriedades Valor Inicial e Valor de Incremento. Se o Gerador de Sequência não estiver configurado para percorrer a sequência, a porta NEXTVAL gerará números de sequência até o valor final configurado.

A seguinte imagem mostra um mapeamento com a porta NEXTVAL da transformação de Gerador de Sequência conectada a uma origem e a dois destinos para gerar valores de chave primária e externa:



Quando você configura a transformação de Gerador de Sequência com um Valor Inicial = 1 e um Valor de Incremento = 1, o Serviço de Integração gera os mesmos valores de chave primária para as tabelas de destino T\_ORDERS\_PRIMARY e T\_ORDERS\_FOREIGN.

## Criar Chaves

É necessário criar valores de chave primários ou estrangeiros com a transformação Gerador de Sequência conectando a porta NEXTVAL a uma transformação de destino ou de fluxo para baixo. É possível usar um intervalo de valores de 1 a 9.223.372.036.854.775,807 com o menor intervalo de 1. Você também pode criar chaves de composição para identificar cada linha na tabela.

Para criar uma chave composta, é possível configurar o Serviço de Integração para circular por um conjunto menor de valores. Por exemplo, se três lojas gerarem números de pedido, configure uma transformação de Gerador de Sequência para percorrer valores de 1 a 3, com incrementos de 1. Quando você conecta a porta ORDER\_NO à transformação de Gerador de Sequência, os valores gerados criam chaves de composição exclusivas.

O seguinte exemplo mostra chaves de composição e números de pedido:

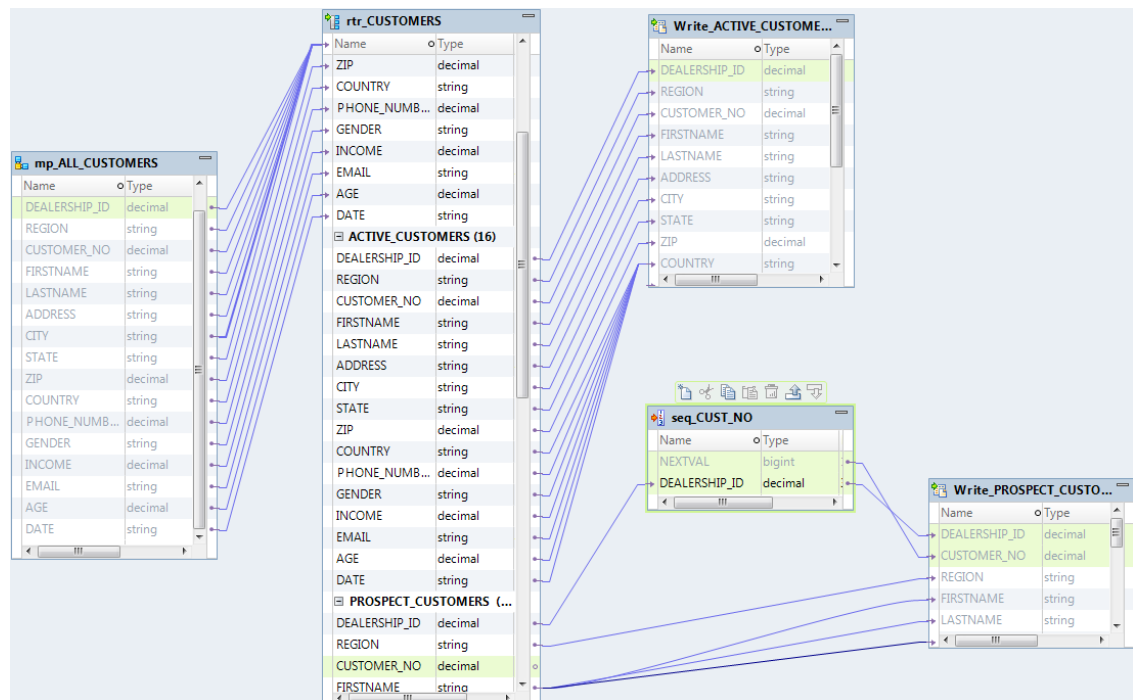
COMPOSITE_KEY	ORDER_NO
1	12345
2	12345
3	12345
1	12346
2	12346
3	12346

## Substituir Valores não Encontrados

Ao usar uma transformação de Gerador de Sequência para substituir chaves ausentes, você também pode usar uma transformação de Roteador para filtrar os valores nulos de colunas com valores atribuídos. Você conecta a transformação de Roteador à transformação de Gerador de Sequência e usa NEXTVAL para gerar uma sequência de valores numéricos para preencher os valores nulos.

Por exemplo, para substituir valores nulos em uma coluna CUSTOMER\_NO, crie um mapeamento com uma origem que contenha dados de clientes. Você adiciona uma transformação de Roteador para filtrar os clientes com números de cliente atribuídos daqueles com valores nulos. Adicione uma transformação de Gerador de Sequência para gerar valores CUSTOMER\_NO exclusivos. Você adiciona destinos de cliente nos quais gravar dados.

A figura abaixo mostra um mapeamento que substitui valores nulos na coluna CUSTOMER\_NO:



# Propriedades da Transformação Gerador de Sequência

Configure propriedades de transformação que o Serviço de Integração usa para gerar valores sequenciais.

A seguinte tabela lista as propriedades que você configura para um objeto de dados de sequência e uma nova sequência:

Propriedade	Descrição
Valor de Início	Valor inicial da sequência gerada que você deseja que o Serviço de Integração use, se estiver usando a opção Ciclo. Se você selecionar Ciclo, o Serviço de Integração voltará a esse valor quando alcançar o valor final. O padrão é 0. O valor máximo é 9.223.372.036.854.775.806.
Valor Final	Valor máximo gerado pelo Serviço de Integração. Se o Serviço de Integração atingir esse valor durante a sessão e a sequência não estiver configurada como ciclo, a sessão será encerrada. O valor máximo é 9,223,372,036,854,775,807.
Valor de Incremento	Diferença entre dois valores consecutivos da porta NEXTVAL. O padrão é 1. Deve ser um inteiro positivo. O valor máximo é 2,147,483,647.
Ciclo	Se ativado, o Serviço de Integração percorrerá o intervalo da sequência e será iniciado com o valor inicial. Se desativado, o Serviço de Integração interromperá a sequência no valor final configurado. O Serviço de Integração encerrará a sessão com erros de overflow se atingir o valor final e ainda houver linhas a serem processadas.

## Valor Inicial

Utilize Ciclo para gerar uma sequência de repetição, como números de 1 a 12 para corresponder aos meses do ano.

1. Insira o menor valor na sequência que você deseja que o Serviço de Integração use para o valor inicial.
2. Insira o maior valor a ser usado para o Valor Final.
3. Selecione Ciclo.

À medida que percorrer a sequência, o Serviço de Integração atingirá o valor final configurado. Ele continuará até o fim da sequência e iniciará o ciclo novamente, começando pelo valor inicial configurado.

## Valor final

O valor final é o valor máximo que você deseja que o Serviço de Integração gere. Se o Serviço de Integração atingir o valor final e o gerador de sequência não estiver configurado para percorrer a sequência, a execução de mapeamento falhará com um erro de overflow.

Defina o valor final como qualquer inteiro entre 1 e 9.233.372.036.854.775.807. Se você conectar a porta NEXTVAL a uma porta de inteiro downstream, defina o valor final como um valor menor que ou igual ao valor máximo de inteiro. Por exemplo, se você conectar a porta NEXTVAL a uma porta de Inteiro Curto, configure o

valor final como no máximo 32.767. Se NEXTVAL exceder o valor máximo do tipo de dados para a porta downstream, o mapeamento falhará.

## Valor de Incremento

O Serviço de Integração gera uma sequência na porta NEXTVAL com base nas propriedades Valor Atual e Incrementar em na transformação de Gerador de Sequência.

A propriedade Valor atual é o valor em que o Serviço de Integração começa a criar a sequência para cada sessão. Incrementar em é o número inteiro que o Serviço de Integração adiciona ao valor existente, para criar o novo valor da sequência. Por padrão, o Valor Atual está definido como 1, enquanto Incrementar em está definido como 1.

Por exemplo, você pode criar uma transformação de Gerador de Sequência com a um valor atual de 1000 e um incremento de 10. Se você passar três linhas através do mapeamento, o Serviço de Integração gera o seguinte conjunto de valores:

```
1000
1010
1020
```

## Percorrer um Intervalo de Valores

Você pode estabelecer um intervalo de valores para a transformação Gerador de Sequência. Se você usar a opção ciclo, a transformação Gerador de Sequência repete o intervalo quando atingir o valor final.

Por exemplo, se você definir o intervalo da sequência para iniciar em 10 e terminar em 50 e definir um valor de incremento de 10, a transformação de Gerador de Sequência criará os valores 10, 20, 30, 40, 50. A sequência é iniciada novamente em 10.

## Redefinir

Se você configurar uma transformação de Gerador de Sequência não reutilizável para usar a propriedade reset, o Serviço de Integração usará o valor inicial original para cada execução de mapeamento. Caso contrário, o Serviço de Integração incrementará o valor atual e o usará na próxima execução de mapeamento.

Por exemplo, você configura uma transformação de Gerador de Sequência para criar valores de 1 a 1.000 com um incremento de 1, um valor inicial de 1 e escolhe redefinir. Durante a primeira execução de mapeamento, o Serviço de Integração gera números de 1 a 234. Em cada execução de mapeamento subsequente, o Serviço de Integração gerará novamente números começando com o valor inicial de 1.

Se você não redefinir, o Serviço de Integração atualizará o valor atual para 235 ao final da primeira execução. Da próxima vez em que a transformação de Gerador de Sequência for usada, o primeiro valor gerado será 235.

**Nota:** Redefinir é desativado para transformações Gerador de Sequência reutilizáveis.

## Manter Ordem das Linhas

Mantenha a ordem de linhas dos dados de entrada na transformação. Selecione essa opção para o Serviço de Integração não realizar nenhuma otimização que possa alterar a ordem das linhas.

Quando o Serviço de Integração realiza otimizações, ele pode perder uma ordem estabelecida anteriormente no mapeamento. Você pode estabelecer a ordem em um mapeamento com uma origem de arquivo simples classificada, uma origem relacional classificada ou uma transformação do Classificador. Quando você configura uma transformação para manter a ordem das linhas, o Serviço de Integração considera essa

configuração quando ele executa as otimizações para o mapeamento. O Serviço de Integração realizará otimizações para a transformação se ele puder manter a ordem. O Serviço de Integração não realizará otimizações para a transformação se a otimização alterar a ordem das linhas.

## Objeto de Dados de Sequência

Um objeto de dados de sequência cria e mantém uma sequência de valores numéricos. A transformação de Gerador de Sequência usa o objeto de dados de sequência para gerar valores para a transformação.

Você pode usar um objeto de dados de sequência reutilizável em várias transformações de Gerador de Sequência. Todas as transformações de Gerador de Sequência que usam o mesmo objeto de dados de sequência usarão a mesma sequência de valores se forem executados no mesmo Serviço de Integração. Você também pode usar um objeto de dados de sequência reutilizável em uma transformação de Gerador de Sequência não reutilizável. Você pode usar um objeto de dados de sequência não reutilizável de uma transformação de Gerador de Sequência não reutilizável.

Por exemplo, crie vários mapeamentos que gravam no mesmo campo de chave primária em uma tabela relacional. Cada mapeamento usa a mesma transformação de Gerador de Sequência reutilizável, que usa o mesmo objeto de dados de sequência reutilizável e é executado no mesmo Serviço de Integração. Cada mapeamento grava valores exclusivos no campo de chave primária.

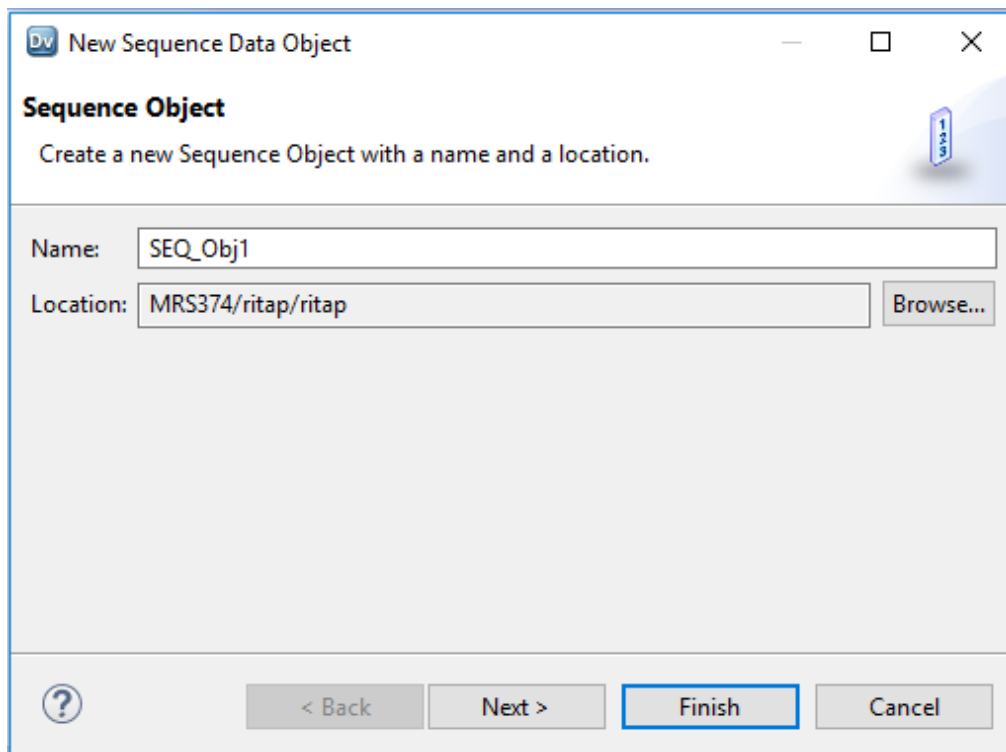
## Criando um objeto de dados de sequência

Para usar um objeto de dados de Sequência para criar uma transformação de Gerador de Sequência, crie o objeto de dados de Sequência, configure as propriedades do objeto e selecione o objeto na caixa de diálogo Transformação de Gerador de Sequência.

1. No editor de mapeamento, role para baixo na paleta de mapeamento para localizar a transformação de Gerador de Sequência e arraste-a até o mapeamento.  
O assistente de **Novas Transformações** é aberto.
2. Clique em **Novo Objeto de Dados de Sequência**.



O assistente para **Novo Objeto de Dados** é aberto.



**New Sequence Data Object**

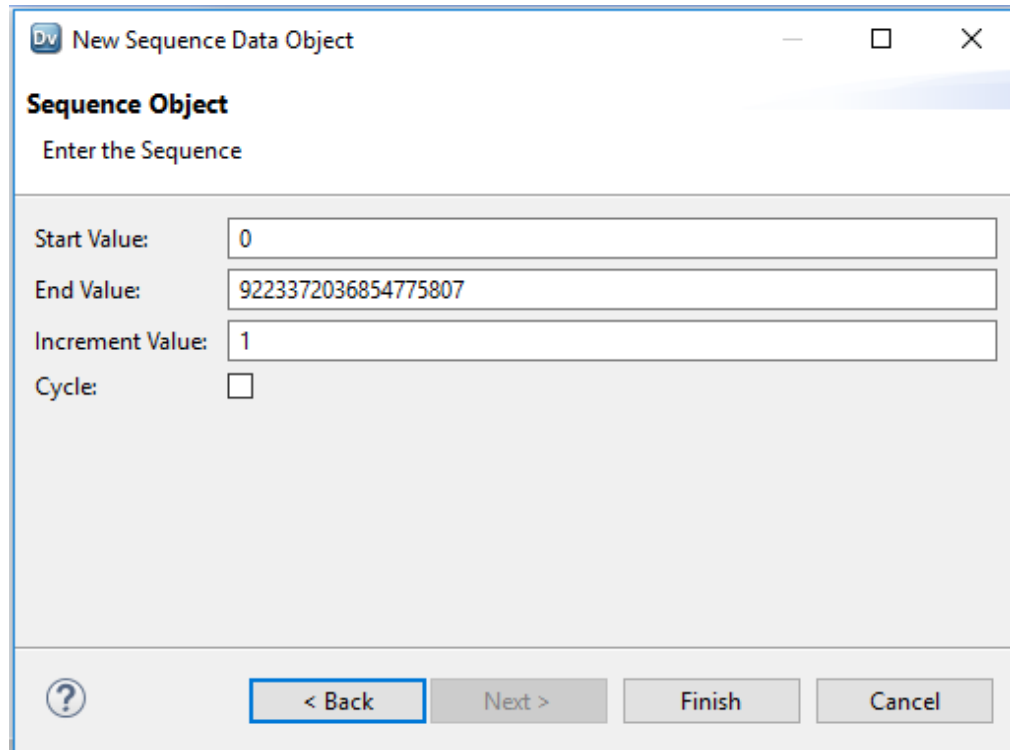
**Sequence Object**  
Create a new Sequence Object with a name and a location.

Name: SEQ\_Obj1

Location: MRS374/ritap/ritap

3. Insira um nome para o objeto de dados de Sequência.  
A convenção de nomenclatura para objetos de dados de Sequência é SEQ\_<nome do objeto de dados>.
4. Clique em **Avançar** para configurar as propriedades do objeto de dados de Sequência.

Se você criar uma transformação de Gerador de Sequência a partir do objeto, essa transformação usará as propriedades que você inserir para o objeto de dados. A seguinte imagem mostra as propriedades que você pode configurar:



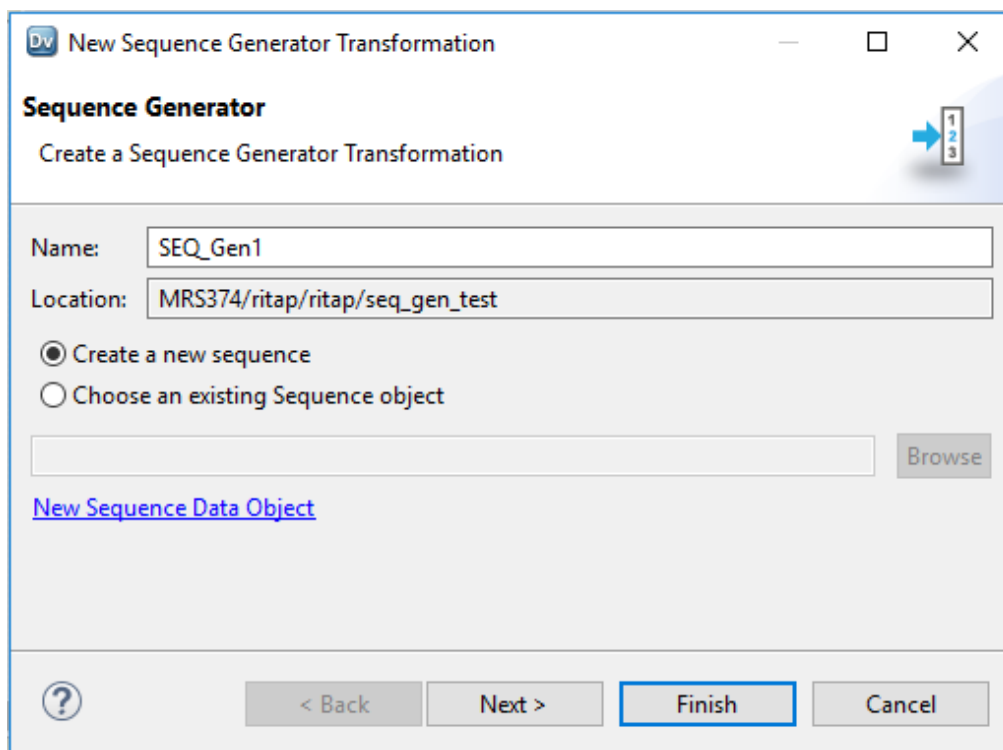
5. Depois de configurar as propriedades do objeto de dados, você pode criar uma transformação de Gerador de Sequência usando o objeto de dados de Sequência. Ao criar a transformação, nomeie a transformação de Gerador de Sequência e selecione **Escolher um objeto de Sequência existente**. Navegue até o objeto de dados e clique em **OK**.

A transformação de Gerador de Sequência aparece no editor de mapeamento com um porta somente de saída NEXTVAL. É possível conectar a porta NEXTVAL a um destino ou transformação em downstream para gerar uma sequência de números.

# Criando uma Transformação de Gerador de Sequência

Para usar uma transformação de Gerador de Sequência em um mapeamento, adicione-a ao mapeamento, configure as propriedades de transformação e conecte NEXTVAL a uma ou mais transformações.

1. No editor de mapeamento, role para baixo na paleta de mapeamento para localizar a transformação de Gerador de Sequência e arraste-a até o mapeamento.  
O assistente de **Novas Transformações** é aberto.



2. Forneça um nome para a transformação Gerador de Sequência.  
A convenção de nomenclatura para transformações de Gerador de Sequência é SEQ\_<nome da transformação>.
3. Escolha para criar uma nova sequência ou usar um objeto de Sequência existente.

- Para criar uma nova sequência, selecione **Criar uma nova sequência**. Clique em **Avançar** para configurar as propriedades da sequência. A seguinte imagem mostra as propriedades que você pode configurar:

- Para usar um objeto de Sequência existente, selecione **Escolher um objeto de Sequência existente**. Navegue até o objeto de Sequência e clique em **OK**.

A transformação de Gerador de Sequência aparece no editor de mapeamento com um porta somente de saída NEXTVAL. É possível conectar a porta NEXTVAL a um destino ou transformação em downstream para gerar uma sequência de números.

## Perguntas frequentes

### Posso alterar uma transformação de Gerador de Sequência não reutilizável para torná-la reutilizável?

Você não pode tornar a transformação reutilizável, mas pode alterar a transformação para usar um objeto de dados de sequência. O objeto de dados de sequência mantém a integridade da sequência, independentemente do número de transformações que o utilizam.

### Posso colocar uma transformação de Gerador de Sequência não reutilizável em um mapplet?

Não, não é possível. Mapplets são objetos reutilizáveis, por isso todos os objetos do mapplet também devem ser reutilizáveis. Use uma transformação Gerador de Sequência reutilizável.

# Transformação de Gerador de Sequência em um ambiente não nativo

O processamento da transformação de Gerador de Sequência em um ambiente não nativo depende do mecanismo que executa a transformação.

Considere o suporte para os seguintes mecanismos de tempo de execução não nativos:

- Mecanismo Blaze. Suportado com restrições.
- Mecanismo Spark. Com suporte com restrições em mapeamentos em lote. Sem suporte em mapeamentos de fluxo.
- Mecanismo Databricks Spark. Sem suporte.

## Transformação de Gerador de Sequência no mecanismo Blaze

Um mapeamento com uma transformação de Gerador de Sequência consome recursos significativos quando as seguintes condições são verdadeiras:

- Você define a propriedade **Manter Ordem da Linha** na transformação como *true*.
- O mapeamento é executado em uma única partição.

## Transformação de Gerador de Sequência no mecanismo Spark

A transformação do Gerador de Sequência não mantém a ordem da linha nos dados de saída. Se você ativar a propriedade **Manter Ordem da Linha** na transformação, o Serviço de Integração de Dados irá ignorar a propriedade.

## CAPÍTULO 41

# Transformação de Ordenador

Este capítulo inclui os seguintes tópicos:

- [Transformação de Classificador - Visão Geral, 618](#)
- [Transformações do Classificador em Mapeamentos Dinâmicos, 619](#)
- [Desenvolvendo uma Transformação de Ordenador, 619](#)
- [Portas de Transformação do Classificador, 620](#)
- [Guia Classificação, 620](#)
- [Configurar Chaves de Classificação, 620](#)
- [Transformação do Classificador - Propriedades Avançadas, 623](#)
- [Cache de classificador, 624](#)
- [Criando uma Transformação de Ordenador, 624](#)
- [Exemplo de Transformação de Ordenador, 625](#)
- [Transformação de Classificador em um ambiente não nativo, 627](#)

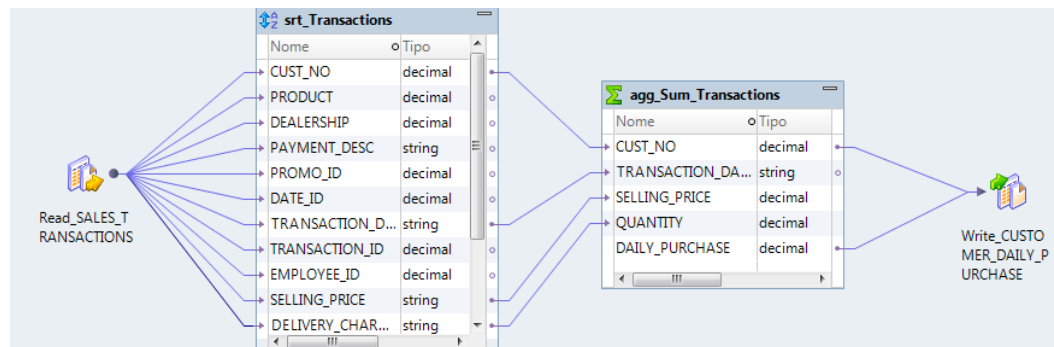
## Transformação de Classificador - Visão Geral

Use uma transformação de Classificador para classificar dados em ordem crescente ou decrescente de acordo com uma chave de classificação especificada. Você pode configurar a transformação de Classificador para classificação com distinção entre maiúsculas e minúsculas e para saída distinta. A transformação de Classificador é ativa.

Ao criar uma transformação de Classificador, você especifica portas como chaves de classificação e configura cada porta de chave de classificação para classificar em ordem crescente ou decrescente. O Serviço de Integração de Dados classifica cada porta sequencialmente quando você especifica várias portas para a chave de classificação.

Por exemplo, você precisa criar uma fatura para as vendas de suporte global de um banco de dados de cliente. Use uma transformação de Classificador na tabela de vendas de clientes para classificar os dados em ordem decrescente de acordo com o número do cliente. Use o resultado da transformação de Classificador como uma entrada para a transformação de Agregador. Você pode aumentar o desempenho da transformação de Agregador com a opção de entrada classificada.

A seguinte figura mostra o mapeamento:



## Transformações do Classificador em Mapeamentos Dinâmicos

É possível usar uma transformação do Classificador em um mapeamento dinâmico. Você pode configurar portas dinâmicas na transformação e fazer referência às portas geradas.

Você pode fazer referência a uma porta dinâmica ou a uma porta gerada na transformação do Classificador. No entanto, se a porta gerada não existir em tempo de execução, o mapeamento falhará.

Se você usar uma porta dinâmica como chave de classificação, o Serviço de Integração de Dados levará em consideração todas as portas geradas na porta dinâmica e a ordem dessas portas geradas.

É possível parametrizar as chaves de classificação. Use um parâmetro de lista de classificação para a chave de classificação.

## Desenvolvendo uma Transformação de Ordenador

Ao desenvolver uma transformação de Ordenador, é necessário considerar fatores como as portas de chave de classificação, linhas de saída distintas e os critérios de classificação com distinção entre maiúsculas e minúsculas.

Considere os seguintes fatores ao desenvolver uma transformação de Ordenador:

- As portas que você quer configurar como chaves de classificação e a direção de classificação.
- Se você quer uma classificação que diferencia maiúsculas de minúsculas.
- Se você quer considerar valores nulos como prioridade de classificação.
- Se você quer linhas de saída distintas.
- O valor de tamanho do cache de ordenador que você quer definir.

# Portas de Transformação do Classificador

Ao criar portas na transformação do Classificador, você cria portas de entrada e saída por padrão. A transformação do Classificador retorna as mesmas portas de saída como portas de entrada.

É possível definir portas dinâmicas em uma transformação do Classificador. Uma porta dinâmica pode receber diferentes colunas de dados de uma transformação em upstream de um mapeamento. Isso permite que a transformação do Classificador classifique linhas que contêm diferentes colunas.

Defina as chaves de classificação na guia **Classificar** da exibição **Propriedades**.

## Guia Classificação

Defina a chave de classificação na guia **Classificar** da exibição **Propriedades** da transformação do Classificador. Escolha uma ou mais portas que você deseja usar como critérios de classificação.

O Serviço de Integração de Dados classifica dados de acordo com a ordem das portas na guia Classificação. Configure para classificar os dados em ordem crescente ou decrescente. O padrão é a ordem crescente.

Se você configurar a transformação do Classificador para linhas de saída distintas, a Developer tool configurará todas as portas como parte da chave de classificação. O Serviço de Integração de Dados descarta linhas duplicadas durante a operação de classificação

## Configurar Chaves de Classificação

Defina a chave de classificação na guia **Classificar** da exibição **Propriedades** da transformação.

Este é o início do conceito.

A seguinte imagem mostra a guia **Classificar**:



Sort

Output:
☒ All rows
☐ Distinct rows only

Sort Keys

Specify by: Value

Ports:

Department	Ascending (A)
Employee	Ascending (A)

Add
Choose...
Delete
Move Up
Move Down

A guia **Classificar** contém as seguintes opções:

#### Saída

Escolha se deseja retornar todas as linhas classificadas ou descartar linhas duplicadas. Linhas duplicadas são linhas em que todos os valores de coluna são os mesmos.

#### Especificar por

Selecione **Valor** ou **Parâmetro**. Selecione **Valor** para usar nomes de portas. Escolha **Parâmetro** para usar um parâmetro de lista de classificação.

#### Adicionar

Aceita um nome de porta que você digita manualmente. Você deve digitar um nome de porta válido antes de clicar em **Adicionar**.

#### Escolher

Clique em **Escolher** para selecionar portas a serem adicionadas à chave de classificação. A Developer tool fornece uma lista de portas de transformação para escolher.

#### Mover para Cima e Mover para Baixo

Você pode alterar a ordem das portas no grupo. Selecione o nome da porta e, em seguida, clique em um dos botões Mover para movê-lo para cima ou para baixo na ordem de classificação.

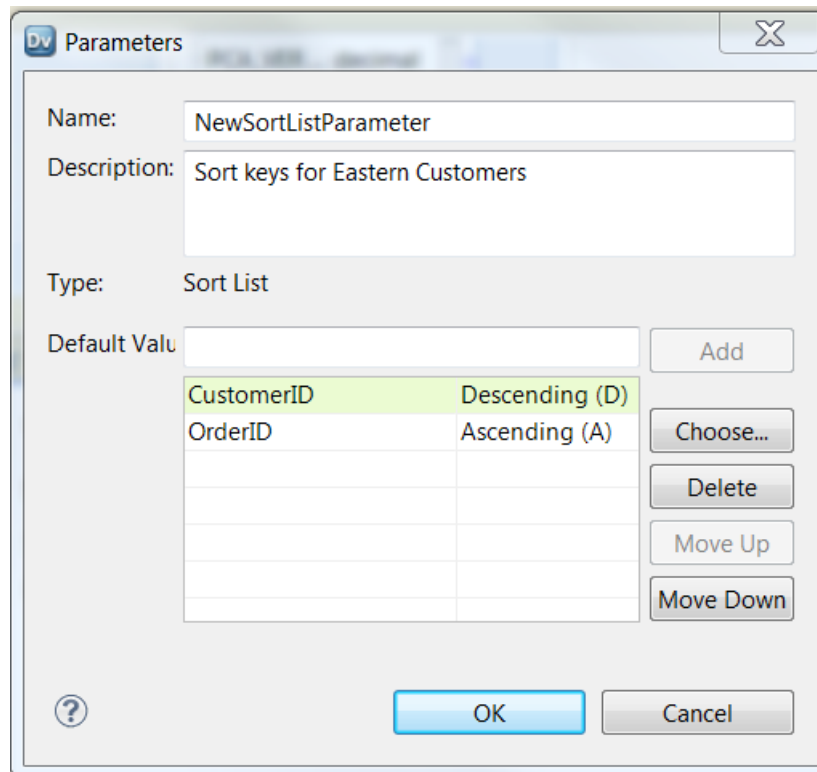
## Parametrizar as Chaves de Classificação

Você pode criar um parâmetro de lista de classificação que contém uma lista de portas para as chaves de classificação.

Se a transformação de Classificação estiver em um mapeamento dinâmico, a transformação do Classificador poderá conter portas geradas. É possível parametrizar as chaves de classificação. Crie um parâmetro de lista de classificação que contenha uma lista de portas com base nas quais classificar.

Na guia **Classificar** das propriedades da transformação, escolha **Especificar por Parâmetro**. Clique em **Novo** para criar um parâmetro.

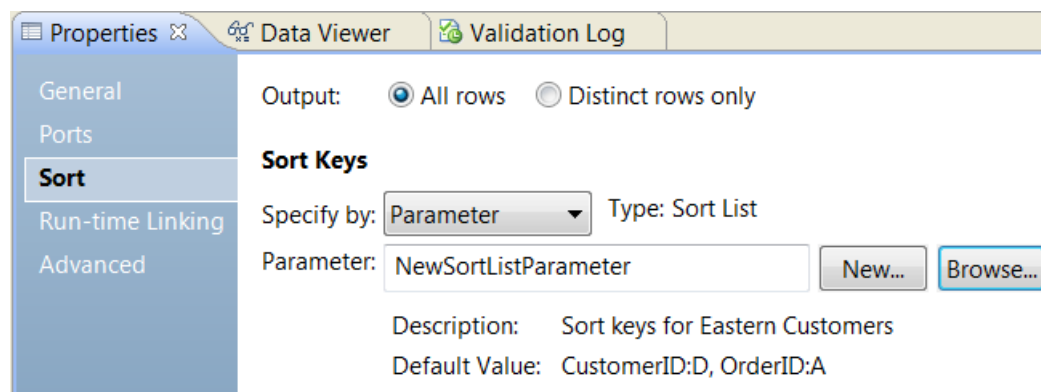
A seguinte imagem mostra a caixa de diálogo **Parâmetros**:



Escolha portas ou portas geradas para as chaves de classificação. Você pode escolher um tipo de classificação crescente ou decrescente.

É possível inserir nomes de portas manualmente. Digite o nome da porta no campo **Valor Padrão** e clique em **Adicionar**. A Developer tool adiciona o nome da porta à lista de classificação.

A seguinte imagem mostra a guia **Classificar** depois que você configura um parâmetro para as chaves de classificação:



# Transformação do Classificador - Propriedades Avançadas

Você pode especificar critérios adicionais de classificação nas propriedades avançadas da transformação do Classificador. O Serviço de Integração de Dados aplica as propriedades a todas as portas de chave de classificação. As propriedades da transformação do Classificador também determinam os recursos do sistema alocados pelo Serviço de Integração de Dados ao classificar dados.

A seguinte seção descreve as propriedades avançadas de uma transformação do Classificador:

## **Distinção Entre Maiúsculas e Minúsculas**

Determina se o Serviço de Integração de Dados faz distinção entre maiúsculas e minúsculas ao classificar dados. Quando você ativa a propriedade Distinção Entre Maiúsculas e Minúsculas, o Serviço de Integração de Dados classifica caracteres maiúsculos como sendo superiores aos minúsculos. A Developer tool define a Distinção Entre Maiúsculas e Minúsculas por padrão.

## **Nulo Tratado como Baixo**

Trata um valor nulo como inferior a qualquer outro valor. Ative a propriedade se desejar que o Serviço de Integração de Dados trate um valor nulo como inferior a qualquer outro valor ao executar a operação de classificação.

## **Tamanho do Cache do Classificador**

Quantidade de memória que o Serviço de Integração de Dados aloca no início da execução de mapeamento para realizar a operação de classificação. O Serviço de Integração de Dados passa todos os dados recebidos para a transformação do Classificador antes de executar a operação de classificação. Selecione Automático para que o Serviço de Integração de Dados calcule automaticamente os requisitos de memória no tempo de execução. Insira um valor específico em bytes quando você ajustar o tamanho do cache. O Padrão é Automático.

## **Diretório de Trabalho**

Diretório no qual o Serviço de Integração de Dados armazenará temporariamente os dados se a quantidade de dados de entrada for maior do que o tamanho do cache de Classificador. Depois que o Serviço de Integração de Dados classifica os dados, ele exclui os arquivos temporários.

Insira vários diretórios separados por ponto-e-vírgula para melhorar o desempenho durante o particionamento do cache. O particionamento do cache cria um cache separado para cada partição que processa a transformação.

O padrão é o parâmetro do sistema TempDir. Você pode configurar outro parâmetro do sistema ou parâmetro definido pelo usuário neste campo.

## **Nível de Rastreamento**

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.

## TÓPICOS RELACIONADOS:

- [“Tamanho do Cache” na página 72](#)

# Cache de classificador

O Serviço de Integração de Dados cria um cache na memória para executar a transformação do Classificador. O Serviço de Integração de Dados passa todos os dados recebidos para a transformação do Classificador antes de executar a operação de classificação. Se o Serviço de Integração de Dados exigir mais espaço que o disponível no cache de memória, ele armazenará temporariamente os dados no diretório de trabalho da transformação do Classificador.

Se você não configurar o tamanho do cache para classificar todos os dados na memória, um aviso será exibido no log da sessão, declarando que o Serviço de Integração de Dados fez várias transmissões nos dados de origem. O Serviço de Integração de Dados faz várias transmissões nos dados quando ele tem que paginar informações no disco para concluir a classificação. A mensagem especifica a quantidade de memória necessária para uma única transmissão, que é quando o Serviço de Integração de Dados lê os dados uma vez e executa a classificação na memória sem paginar no disco. Para otimizar o desempenho do mapeamento, configure o tamanho do cache para que o Serviço de Integração de Dados faça uma transmissão de dados.

Se a quantidade de dados recebidos for maior que o tamanho do cache do Classificador, o Serviço de Integração de Dados armazenará temporariamente os dados no diretório de trabalho da transformação do Classificador. Ao armazenar dados no diretório de trabalho, o Serviço de Integração de Dados requer que o espaço em disco seja pelo menos duas vezes maior que a quantidade de dados recebidos.

Para obter o melhor desempenho, configure o tamanho do cache de Classificador com um valor menor ou igual à quantidade de memória física disponível na máquina que executa o mapeamento. Para classificar dados usando uma transformação do Classificador, aloque pelo menos 16 MB (16.777.216 bytes) de memória física. O tamanho do cache do Classificador é definido como Automático por padrão.

## Otimizando o cache de classificador

O cache de classificador é otimizado para usar tamanho variável para armazenar tipos de dados binary e string que passam pela transformação de Classificador.

O tamanho variável reduz a quantidade de dados que o Serviço de Integração de Dados armazena no cache de classificador e o consumo de espaço em disco na máquina do Serviço de Integração de Dados.

Por exemplo, você armazena dados em clientes. Alguns clientes têm nomes mais longos que outros. Se o Serviço de Integração de Dados usar tamanho fixo para armazenar os dados nos nomes dos clientes, ele poderá armazenar dados em 20 caracteres para cada nome. Se o Serviço de Integração de Dados usar tamanho variável, o Serviço de Integração de Dados poderá armazenar dados com um tamanho médio de 10 caracteres.

# Criando uma Transformação de Ordenador

Você pode criar transformações de Ordenador reutilizáveis ou não reutilizáveis.

## Criando uma Transformação do Classificador Reutilizável

Crie uma Transformação do Classificador reutilizável para usar em vários mapeamentos ou mapplets.

1. Selecione um projeto ou pasta na exibição **Object Explorer**.
2. Clique em **Arquivo > Novo > Transformação**.  
A caixa de diálogo **Novo** é exibida.
3. Selecione a Transformação do Classificador.
4. Clique em **Avançar**.
5. Insira um nome para a transformação.
6. Clique em **Concluir**.  
A transformação aparece no editor.
7. Clique em **Novo** para adicionar uma porta à transformação.
8. Edite a porta para definir o nome, o tipo de dados e a precisão.
9. Na guia **Classificar**, escolha as portas a serem classificadas ou selecione um parâmetro de lista de classificação.
10. Clique na exibição **Avançado** e edite as propriedades da transformação.

## Criando uma Transformação de Ordenador Não Reutilizável

Crie uma transformação de Ordenador não reutilizável em um mapeamento ou em mapplets.

1. Em um mapeamento ou mapplet, arraste uma transformação de Ordenador da paleta de Transformação para o editor.  
A transformação aparece no editor.
2. Na exibição **Propriedades**, edite o nome da transformação e a descrição.
3. Na guia **Portas**, clique em **Novo** para adicionar portas à transformação.
4. Edite as portas para definir o nome, o tipo de dados e a precisão.
5. Selecione **Chave** para indicar a porta como chave de classificação.
6. Clique na guia **Avançado** e edite as propriedades da transformação.

## Exemplo de Transformação de Ordenador

Você tem uma tabela de banco de dados PRODUCT\_ORDERS contendo informações sobre todos os pedidos que foram feitos pelo cliente.

ORDER_ID	ITEM_ID	ITEM	QUANTITY	PRICE
43	123456	ItemA	3	3.04
41	456789	ItemB	2	12.02
43	000246	ItemC	6	34.55
45	000468	ItemD	5	0.56
41	123456	ItemA	4	3.04

ORDER_ID	ITEM_ID	ITEM	QUANTITY	PRICE
45	123456	ItemA	5	3.04
45	456789	ItemB	3	12.02

Use a transformação de Ordenador em PRODUCT\_ORDERS e especifique ORDER\_ID como a chave de classificação, com direção decrescente.

Após classificar os dados, o Data Integration Service retira as seguintes linhas da transformação de Ordenador:

ORDER_ID	ITEM_ID	ITEM	QUANTITY	PRICE
45	000468	ItemD	5	0.56
45	123456	ItemA	5	3.04
45	456789	ItemB	3	12.02
43	123456	ItemA	3	3.04
43	000246	ItemC	6	34.55
41	456789	ItemB	2	12.02
41	123456	ItemA	4	3.04

Você precisa descobrir o valor total e a quantidade de itens para cada pedido. Você pode usar o resultado da transformação de Ordenador como entrada para uma transformação de Agregador. Use a entrada classificada na transformação de Agregador para melhorar o desempenho.

Quando você não usa a entrada classificada, o Data Integration Service realiza cálculos de agregação conforme lê. O Data Integration Service armazena dados para cada grupo até ler toda a origem para garantir que todos os cálculos de agregação sejam exatos. Se você usar a entrada classificada e não pré-classificar os dados corretamente, receberá resultados inesperados.

A transformação de Agregador tem o grupo por porta ORDER\_ID, com a opção de entrada classificada selecionada. Quando você passa os dados da transformação de Ordenador, a transformação de Agregador agrupa ORDER\_ID para calcular o valor total para cada pedido.

ORDER_ID	SUM
45	54.06
43	216.42
41	36.2

# Transformação de Classificador em um ambiente não nativo

O processamento da transformação de Classificador em um ambiente não nativo depende do mecanismo que executa a transformação.

Considere o suporte para os seguintes mecanismos de tempo de execução não nativos:

- Mecanismo Blaze. Suportado com restrições.
- Mecanismo Spark. Com suporte com restrições em mapeamentos de fluxo e em lote.
- Mecanismo Databricks Spark. Suportado com restrições.

## Transformação de Classificador no mecanismo Blaze

Algumas regras de processamento para o mecanismo Blaze são diferentes das regras de processamento do Serviço de Integração de Dados.

### Validação de Mapeamento

A validação do mapeamento falha nas seguintes situações:

- O destino é configurado para manter a ordem de linha e a transformação de Classificador não está conectada diretamente a um destino de arquivo simples.

### Classificação paralela

O Serviço de Integração de Dados permite a classificação paralela com as seguintes restrições:

- O mapeamento não inclui outra transformação entre a transformação de Classificador e o destino.
- O tipo de dados das chaves de classificação não é alterado entre a transformação de Classificador e o destino.
- Cada chave de classificação na transformação de Classificador deve estar vinculada a uma coluna no destino.

### Classificação global

O mecanismo Blaze pode realizar classificações globais nas seguintes situações:

- A transformação de Classificador é conectada diretamente aos destinos de arquivo simples.
- O destino é configurado para manter a ordem da linha.
- A chave de classificação não é um tipo de dados binary.

Se alguma das condições não for verdadeira, o mecanismo Blaze executará uma classificação local.

### Otimização do cache de dados

Se uma transformação de Classificador for inserida antes de uma transformação de Agregador ou Classificação para otimizar o cache de dados de Agregador ou de Classificação, o tamanho do cache de classificação será o mesmo tamanho do cache de dados para a transformação de Agregador ou Classificação. Para configurar o cache de classificação, você deve configurar o tamanho do cache de dados para a transformação de Agregador ou Classificação.

## Transformação de Classificador no mecanismo Spark

Algumas regras de processamento para o mecanismo Spark são diferentes das regras de processamento do Serviço de Integração de Dados.

### Validação de Mapeamento

A validação de mapeamento falha quando a diferenciação de maiúsculas e minúsculas está desativada.

O Serviço de Integração de Dados registra um aviso e ignora a transformação de Classificador nas seguintes situações:

- Há uma incompatibilidade de tipo entre as chaves de classificação de transformação de Classificador e de destino.
- A transformação contém chaves de classificação que não estão conectadas ao destino.
- A transformação de Gravação não está configurada para manter a ordem da linha.
- A transformação não está diretamente upstream da transformação de Gravação.

### Valores Nulos

O Serviço de Integração de Dados trata valores nulos como baixos, mesmo se você configurar a transformação para tratar valores nulos como altos.

### Otimização do cache de dados

Você não pode otimizar o cache de classificador para armazenar dados usando o tamanho da variável.

### Classificação paralela

O Serviço de Integração de Dados permite a classificação paralela com as seguintes restrições:

- O mapeamento não inclui outra transformação entre a transformação de Classificador e o destino.
- O tipo de dados das chaves de classificação não é alterado entre a transformação de Classificador e o destino.
- Cada chave de classificação na transformação de Classificador deve estar vinculada a uma coluna no destino.

## Transformação de Classificador em um mapeamento de fluxo

Os mapeamentos de fluxo têm regras de processamento adicionais que não se aplicam a mapeamentos em lote.

### Validação de Mapeamento

A validação do mapeamento falha na seguinte situação:

- A transformação de Classificador em um mapeamento de streaming não deve ter uma transformação de Agregador upstream.

### Diretrizes gerais

Considere as seguintes diretrizes gerais:

- O mapeamento será executado no modo de saída completo se contiver uma transformação de Classificador.
- Para manter a ordem de classificação global, verifique se o destino é de partição única. A fonte pode ser de partição única ou múltipla.



# Transformação de Classificador no mecanismo Databricks Spark

Algumas regras de processamento para o mecanismo Databricks Spark são diferentes das regras de processamento do Serviço de Integração de Dados.

## Validação de Mapeamento

A validação de mapeamento falha quando a diferenciação de maiúsculas e minúsculas está desativada.

O Serviço de Integração de Dados registra um aviso e ignora a transformação de Classificador nas seguintes situações:

- Há uma incompatibilidade de tipo entre as chaves de classificação de transformação de Classificador e de destino.
- A transformação contém chaves de classificação que não estão conectadas ao destino.
- A transformação de Gravação não está configurada para manter a ordem da linha.
- A transformação não está diretamente upstream da transformação de Gravação.

## Valores Nulos

O Serviço de Integração de Dados trata valores nulos como baixos, mesmo se você configurar a transformação para tratar valores nulos como altos.

## Otimização do cache de dados

Você não pode otimizar o cache de classificador para armazenar dados usando o tamanho da variável.

## Classificação paralela

O Serviço de Integração de Dados permite a classificação paralela com as seguintes restrições:

- O mapeamento não inclui outra transformação entre a transformação de Classificador e o destino.
- O tipo de dados das chaves de classificação não é alterado entre a transformação de Classificador e o destino.
- Cada chave de classificação na transformação de Classificador deve estar vinculada a uma coluna no destino.

## CAPÍTULO 42

# Transformação SQL

Este capítulo inclui os seguintes tópicos:

- [Transformação SQL - Visão Geral, 630](#)
- [Portas da Transformação SQL, 631](#)
- [Propriedades Avançadas da Transformação SQL, 635](#)
- [Consulta da Transformação SQL, 637](#)
- [Cardinalidade de Linha de Entrada para Linha de Saída, 639](#)
- [Otimização de Filtro com a Transformação SQL, 643](#)
- [Exemplo de transformação SQL com uma Consulta SQL, 645](#)
- [Procedimentos Armazenados, 649](#)
- [Conexão de Transformação SQL, 654](#)
- [Criando Manualmente uma Transformação SQL, 655](#)
- [Criando uma Transformação SQL a partir de um Procedimento Armazenado, 656](#)

## Transformação SQL - Visão Geral

A transformação SQL processa consultas SQL midstream em um mapeamento. É possível executar consultas SQL a partir da transformação SQL ou configurar essa transformação SQL de forma a executar procedimentos armazenados de um banco de dados.

Você pode transmitir valores de portas de entrada aos parâmetros na consulta ou no procedimento armazenado. A transformação pode inserir, excluir, atualizar e recuperar linhas de um banco de dados. Você pode executar instruções DDL SQL para criar uma tabela ou descartar uma tabela midstream em um mapeamento. A transformação SQL é uma transformação ativa. A transformação pode retornar diversas linhas para cada linha de entrada.

É possível importar um procedimento armazenado de um banco de dados para a transformação SQL. Quando você importar o procedimento armazenado, a ferramenta Developer criará as portas de transformação que correspondem aos parâmetros desse procedimento. A ferramenta Developer também cria a chamada de procedimento armazenado para você.

Para configurar uma transformação SQL de forma a executar um procedimento armazenado, realize as seguintes tarefas:

1. Defina as propriedades de transformação, incluindo o tipo de banco de dados para conexão.

2. Importe um procedimento armazenado para definir as portas e criar a chamada de procedimento armazenado.
3. Defina portas manualmente para conjuntos de resultados ou procedimentos armazenados adicionais que você precisa executar.
4. Adicione as outras chamadas de procedimento armazenado no Editor SQL.

É possível configurar uma consulta SQL no Editor SQL da transformação. Quando você executa a transformação SQL, ela processa a consulta, retorna linhas e também retorna qualquer erro de banco de dados.

Para configurar uma transformação SQL de forma a executar uma consulta, realize as seguintes tarefas:

1. Defina as propriedades de transformação, incluindo o tipo de banco de dados para conexão.
2. Defina as portas de entrada e saída.
3. Crie uma consulta SQL no Editor SQL.

Depois de configurar a transformação, configure a transformação SQL em um mapeamento e conecte as portas upstream. Visualize os dados para verificar os resultados.

## Portas da Transformação SQL

Quando você cria uma transformação SQL, a ferramenta Developer cria a porta `SQLError` por padrão. Adicione portas de entrada, saída e passagem na exibição **Portas**.

A transformação SQL tem os seguintes tipos de portas:

### **Entrada**

Recebe dados de origem que podem ser usados em uma consulta SQL.

### **Saída**

Retorna dados do banco de dados de uma consulta SQL `SELECT`.

### **Passagem**

Portas de entrada-saída que passam dados de origem pela transformação sem alterá-los.

### **SQLError**

Retorna erros de SQL do banco de dados. Se nenhum erro ocorrer, retornará `NULL`.

### **NumRowsAffected**

Retorna o número total de linhas do banco de dados afetadas pelas instruções de consulta `INSERT`, `DELETE` e `UPDATE` para uma linha de entrada. A ferramenta Developer cria essa porta quando você opta por incluir as estatísticas de atualização na linha de saída.

### **Valor de Retorno**

Recebe o valor de retorno de um procedimento armazenado.

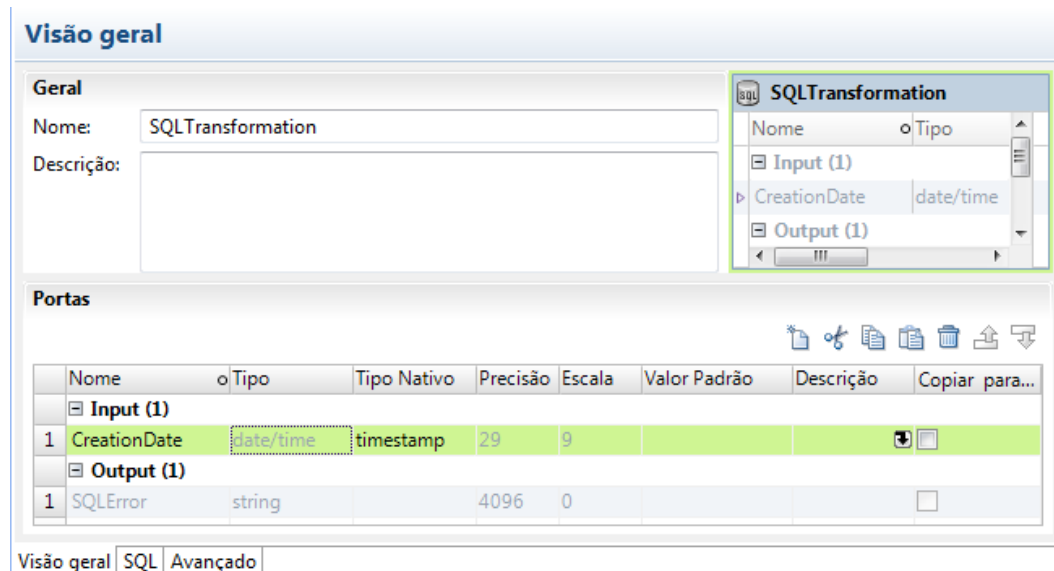
## Portas de Entrada

Você pode fazer referência a portas de entrada da transformação SQL com associação de parâmetros em qualquer tipo de instrução SQL ou procedimento armazenado. Você pode criar portas de entrada na transformação SQL para dados que não pretende passar para as portas de saída.

Será necessário adicionar portas manualmente se você estiver configurando uma consulta SQL que possui parâmetros de entrada. Quando você importa um procedimento armazenado para a transformação SQL, esta cria as portas de entrada. Você pode adicionar portas de passagem para transmitir dados pela transformação sem alterá-los.

É possível adicionar portas na exibição **Visão Geral**. Ao adicionar uma porta, insira o tipo de dados nativo dessa porta. O tipo de dados nativo é um tipo de dados que é válido para o banco de dados ao qual você está se conectando. Quando você configura um tipo de dados nativo, um tipo de dados de transformação é exibido. Se você arrasta linhas para a transformação SQL, a ferramenta Developer define o tipo de dados nativo com base nos tipos de dados válidos para o banco de dados ao qual você está se conectando. Verifique se os tipos de dados das colunas que você usa na consulta são os mesmos tipos de dados das colunas no banco de dados.

A seguinte figura mostra a porta de entrada **CreationDate** em uma transformação SQL:



Para adicionar portas de entrada, clique em **Entrada** no painel **Portas**. Clique em **Novo**.

**Nota:** Se você seleciona **Copiar para Saída** para uma porta, a porta de entrada torna-se uma porta de passagem. As portas de passagem aparecem nas seções **Entrada** e **Saída** da exibição **Portas**.

## Portas de Saída

saída de transformação SQL portas retornam valores de uma instrução de consulta ou de um procedimento armazenado.

Você deve definir as portas de saída ao configurar manualmente uma transformação SQL. Defina uma porta de saída para cada parâmetro de saída de procedimento armazenado ou para cada porta que uma instrução SELECT retorna.

Quando um procedimento armazenado é importado, a ferramenta Developer cria uma porta de saída para cada parâmetro de saída que esse procedimento retorna. Se o procedimento retornar um conjunto de resultados, você deverá definir manualmente as portas de saída no conjunto de resultados. Um

procedimento armazenado pode retornar um conjunto de resultados e também pode retornar parâmetros de saída que não fazem parte do conjunto de resultados na mesma execução. É necessário definir as portas de saída para os campos do conjunto de resultados e para os parâmetros de saída.

Ao configurar uma porta de saída, escolha o tipo de dados nativo para a porta. O tipo de dados nativo da porta de saída deve corresponder ao tipo de dados da coluna correspondente no banco de dados. Quando você configura o tipo de dados nativo, a ferramenta Developer define o tipo de dados de transformação para a porta.

Por exemplo, a transformação SQL contém a seguinte consulta SQL para um banco de dados Oracle:

```
SELECT FirstName, LastName, Age FROM EMPLOYEES
```

Você pode configurar as seguintes portas de saída e tipos de dados nativos na transformação SQL:

Porta de Saída	Tipo de Dados Nativo	Tipo de Dados de Transformação
FirstNm	varchar2	string
LastNm	varchar2	string
Idade	number	double

O número das portas de saída e sua ordem devem corresponder ao número e à ordem das colunas retornadas pela coluna ou pelo procedimento armazenado. Quando o número de portas de saída é maior que o número de colunas na coluna ou no procedimento armazenado, as portas extras retornam um valor nulo. Quando o número de portas de saída é menor que o número de colunas no SQL, o Serviço de Integração de Dados gera um erro de linha.

Se você alterar o tipo de banco de dados ao qual a transformação se conecta, a ferramenta Developer alterará os tipos nativos das portas de saída. A ferramenta Developer pode não escolher o tipo de dados correto para todas as portas. Se alterar o tipo de banco de dados, verifique se o tipo de dados nativo para cada porta de saída é o mesmo tipo de dados da coluna no banco de dados. Por exemplo, a ferramenta Developer pode escolher nVarchar2 para uma coluna de um banco de dados Oracle. Pode ser necessário mudar o tipo de dados para varchar2.

Configure portas de saída na exibição **Visão Geral** da transformação SQL.

## Portas de Passagem

As portas de passagem são portas de entrada-saída que passam dados pela transformação sem alterá-los. A transformação SQL retorna dados das portas de passagem, independentemente de uma consulta SQL retornar linhas ou não.

Quando uma linha de origem contém uma instrução de consulta SELECT, a transformação SQL retorna os dados na porta de passagem para cada linha retornada do banco de dados. Se o resultado da consulta contiver várias linhas, a transformação SQL repetirá os dados de passagem em cada linha.

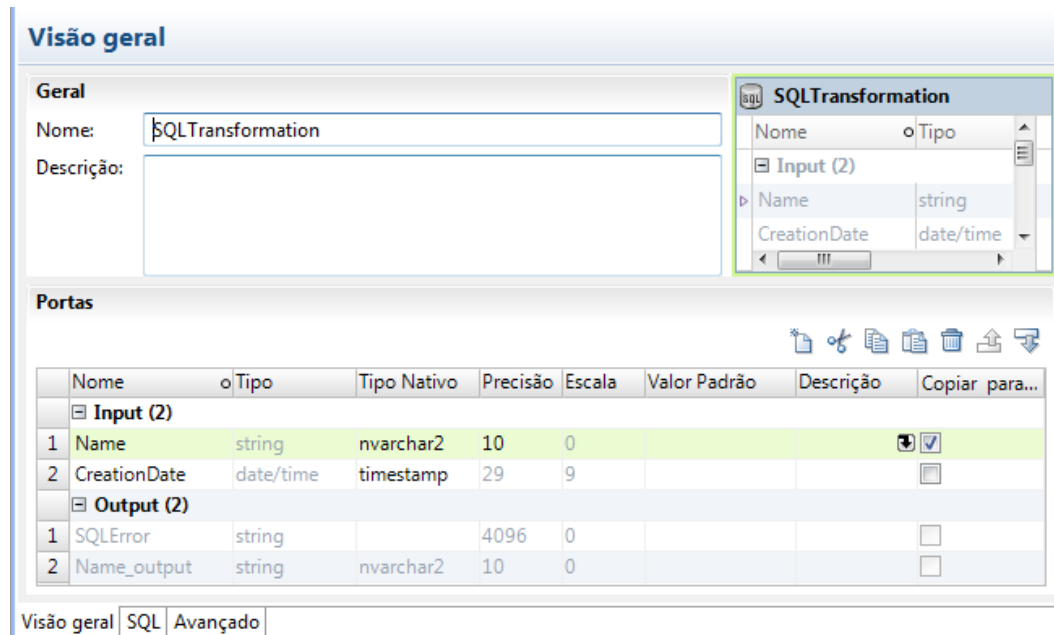
Quando uma consulta não retornar linhas, a transformação SQL retornará os dados da coluna de passagem com valores nulos nas colunas de saída. Por exemplo, as consultas que contêm instruções INSERT, UPDATE e DELETE não retornam linhas. Quando a consulta contiver erros, a transformação SQL retornará os dados da coluna de passagem, a mensagem SQLError e valores nulos nas portas de saída.

Não é possível configurar uma porta de passagem para retornar dados de uma consulta SELECT.

Para criar uma porta de passagem, crie uma porta de entrada e selecione **Copiar para Saída**. A ferramenta Developer cria uma porta de saída e adiciona um sufixo "\_output" ao nome da porta. Não é possível alterar a

porta de saída que a ferramenta Developer cria para uma porta de passagem. Você não pode criar uma porta de saída com o sufixo "\_output".

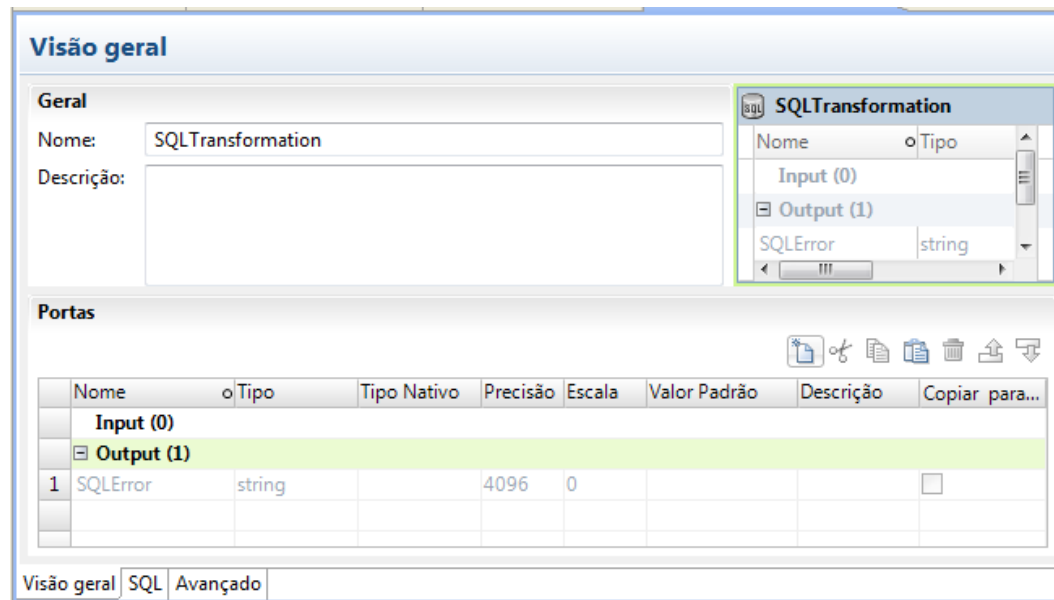
A figura a seguir mostra uma porta de passagem de Nome em uma transformação SQL reutilizável:



## Porta SQLError

A porta SQLError retorna erros de SQL do banco de dados de procedimentos armazenados ou consultas SQL.

A figura a seguir mostra a porta SQLError em uma transformação SQL:



Quando a consulta SQL contém erros de sintaxe, a porta SQLError contém o texto de erro do banco de dados. Por exemplo, a seguinte consulta SQL gera uma linha de erro em um banco de dados Oracle:

```
SELECT Product_ID FROM Employees
```

A tabela Funcionários não contém Product\_ID. O Serviço de Integração de Dados gera uma única linha. A porta SQL\_Error contém o texto do erro em uma linha:

```
ORA-0094: "Product ID": invalid identifier Database driver error... Function Name:
Execute SQL Stmt: SELECT Product_ID from Employees Oracle Fatal Error
```

É possível configurar várias instruções de consulta na consulta SQL ou chamar vários procedimentos armazenados. Quando você configura a transformação SQL para continuar com erro de SQL, ela pode retornar linhas para uma instrução de consulta, mas retornar erros de banco de dados para outra instrução de consulta. A transformação SQL retorna qualquer erro de banco de dados em uma linha separada.

## Número de Linhas Afetadas

Ative a porta de saída NumRowsAffected para retornar o número de linhas alteradas pelas instruções de consulta INSERT, UPDATE ou DELETE para cada linha de entrada. É possível configurar a porta de saída NumRowsAffected para consultas SQL.

O Serviço de Integração de Dados retorna o NumRowsAffected para cada instrução da consulta. O NumRowsAffected está desabilitado por padrão.

Quando você ativa NumRowsAffected e a consulta SQL não contém uma instrução INSERT, UPDATE ou DELETE, o valor de NumRowsAffected é zero em cada linha de saída.

Quando a consulta SQL contém várias instruções, o Serviço de Integração de Dados retorna o NumRowsAffected para cada instrução. NumRowsAffected contém a soma das linhas alteradas pelas instruções INSERT, UPDATE e DELETE para uma linha de entrada.

Por exemplo, uma consulta contém as seguintes instruções:

```
DELETE from Employees WHERE Employee_ID = '101';
SELECT Employee_ID, LastName from Employees WHERE Employee_ID = '103';
INSERT into Employees (Employee_ID, LastName, Address)VALUES ('102', 'Gein', '38 Beach Rd')
```

A instrução DELETE afeta uma linha. A instrução SELECT não afeta nenhuma linha. A instrução INSERT afeta uma linha.

O Serviço de Integração de Dados retorna uma linha da instrução DELETE. NumRowsAffected é igual a um. O Serviço de Integração de Dados retorna uma linha da instrução SELECT; NumRowsAffected é zero. O Serviço de Integração de Dados retorna uma linha da instrução INSERT com NumRowsAffected igual a um.

## Propriedades Avançadas da Transformação SQL

As propriedades da transformação SQL podem ser alteradas a qualquer momento. O tipo de banco de dados padrão é Oracle. Se o banco de dados ao qual você precisa se conectar for de outro tipo, altere o tipo de banco de dados antes de adicionar portas à transformação.

Configure as seguintes propriedades na guia **Avançado**:

### Nível de Rastreamento

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal. Quando você configura o nível de rastreamento da transformação SQL para Dados Detalhados, o Serviço de Integração de Dados grava no log de mapeamento cada consulta SQL preparada.

### Tipo de conexão

Descreve como o Serviço de Integração de Dados se conecta ao banco de dados. O tipo de conexão é estático. O Serviço de Integração de Dados se conecta uma única vez ao banco de dados. Selecione um objeto de conexão de banco de dados na transformação SQL. Somente leitura.

### Tipo de DB

Tipo de banco de dados ao qual a transformação SQL se conecta. Escolha um tipo de banco de dados na lista. É possível escolher Oracle, Microsoft SQL Server, IBM DB2 ou ODBC. O tipo de banco de dados afeta os tipos de dados que podem ser atribuídos na guia **Portas**. Quando você altera o tipo de banco de dados, a Developer tool altera os tipos de dados das portas de entrada, saída e passagem.

### Continuar com Erro Dentro da Linha

Continua processando as instruções SQL restantes em uma consulta mesmo depois de ocorrer um erro de SQL.

### Incluir Estatísticas como Saída

Adiciona uma porta de saída NumRowsAffected. A porta retorna o número total de linhas do banco de dados que as instruções de consulta INSERT, DELETE e UPDATE atualizam para uma linha de entrada.

### Contagem Máx. de Linhas de Saída

Define o número máximo de linhas que a transformação SQL pode emitir de uma consulta SELECT. Para configurar linhas ilimitadas, defina a propriedade Número de Linhas de Saída Máximas como zero.

### Descrição da Consulta

Descrição da consulta SQL que você define na transformação.

### Modo SQL

Determina se a consulta SQL é um script externo ou se a consulta é definida na transformação. O Modo SQL é Consulta. A transformação SQL executa uma consulta que você define no Editor SQL. Somente leitura.

### Consulta SQL

Exibe a consulta SQL que você configura no Editor SQL.

### Tem Efeitos Colaterais

Indica que a transformação SQL realiza uma função além de retornar linhas. A transformação SQL tem um efeito colateral quando a consulta SQL atualiza um banco de dados. Ative **Tem Efeitos Colaterais** quando a consulta SQL contiver uma instrução como CREATE, DROP, INSERT, UPDATE, GRANT ou REVOKE.

A transformação SQL também tem um efeito colateral se a transformação retorna linhas NULL para instruções SELECT que não retornam resultados. As linhas podem conter valores de porta de passagem, informações de erro de SQL ou o campo NumRowsAffected.

Desative a propriedade **Tem Efeitos Colaterais** a fim de permitir a otimização de envio ou a otimização de seleção antecipada. Por padrão, essa opção fica ativada.

### Retornar Somente Saída de Banco de Dados

A transformação SQL não gera linhas para instruções SELECT que retornam 0 resultados, linhas para outras instruções como INSERT, UPDATE, DELETE ou COMMIT, ou linhas nulas.

### Ativar Otimização de Envio

Permite que o Serviço de Integração de Dados envie a lógica de uma transformação de Filtro no mapeamento para o SQL na transformação SQL.



### **Manter Ordem das Linhas**

Mantenha a ordem de linhas dos dados de entrada na transformação. Selecione essa opção se o Serviço de Integração de Dados não realizar qualquer otimização que possa alterar a ordem de linhas.

Quando o Serviço de Integração de Dados realiza as otimizações, ele pode perder uma ordem estabelecida anteriormente no mapeamento. Você pode estabelecer a ordem em um mapeamento com uma origem de arquivo simples classificada, uma origem relacional classificada ou uma transformação do Classificador. Quando você configura uma transformação para manter a ordem de linhas, o Serviço de Integração de Dados considera essa configuração quando ele executa as otimizações para o mapeamento. O Serviço de Integração de Dados realizará otimizações para a transformação se ela puder manter a ordem. O Serviço de Integração de Dados não realizará otimizações para a transformação se a otimização alterar a ordem de linha.

### **Particionável**

A transformação pode ser processada com vários segmentos. Desmarque essa opção se quiser que o Serviço de Integração de Dados use um segmento para processar a transformação. O Serviço de Integração de Dados pode usar vários segmentos para processar as fases de pipeline de mapeamento restantes.

Desative o particionamento de uma transformação SQL quando as consultas SQL exigirem que a transformação seja processada com um segmento. Ou você pode querer desativar o particionamento de uma transformação SQL para que somente uma conexão seja feita com o banco de dados.

## Consulta da Transformação SQL

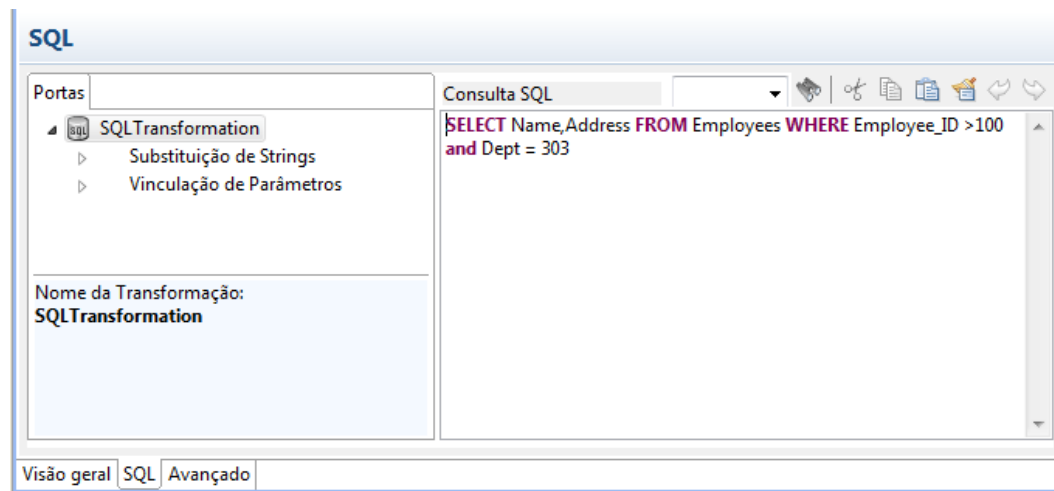
Crie uma consulta SQL no Editor SQL para recuperar linhas de um banco de dados ou para atualizar o banco de dados.

Para criar uma consulta, digite a instrução de consulta no Editor SQL na exibição SQL. Você pode inserir até 32767 caracteres em uma instrução de consulta SQL. O Editor SQL fornece uma lista das portas da transformação a que você pode fazer referência na consulta. Você pode clicar duas vezes no nome de uma porta para adicioná-la como parâmetro da consulta.

Quando você cria uma consulta, o Editor SQL valida os nomes das portas na consulta. Verifica também se as portas usadas para substituição de strings são tipos de dados de string. O Editor SQL não valida a sintaxe da consulta SQL.

É possível usar constantes na consulta SQL. Coloque cada string entre aspas simples (').

A seguinte figura mostra um exemplo de consulta SQL:



Você pode criar uma consulta SQL estática. A instrução de consulta não muda, mas você pode incluir parâmetros para alterar valores. O Serviço de Integração de Dados executa a consulta para cada linha de entrada.

## Definir a Consulta SQL

Defina uma consulta SQL que executa as mesmas instruções de consulta para cada linha de entrada. Você pode alterar a tabela ou as colunas da consulta com base no valores de porta de entrada da linha. Também é possível alterar os valores na cláusula WHERE com base nos valores de porta de entrada.

Para alterar os valores de dados na cláusula WHERE para cada linha de entrada, configure a vinculação de parâmetros.

Para alterar as colunas da consulta ou alterar a tabela com base nos valores de porta de entrada, use a substituição de strings.

## Vinculação de Parâmetros

Para alterar os dados na consulta, configure os parâmetros da consulta e vincule-os a portas de entrada na transformação. Ao vincular um parâmetro a uma porta de entrada, identifique a porta pelo nome na consulta. O Editor SQL coloca o nome da porta entre pontos de interrogação (?). Os dados da consulta mudam com base no valor dos dados na porta.

As seguintes consultas usam vinculação de parâmetros:

```
DELETE FROM Employee WHERE Dept = ?Dept?
INSERT INTO Employee(Employee_ID, Dept) VALUES (?Employee_ID?, ?Dept?)
UPDATE Employee SET Dept = ?Dept? WHERE Employee_ID > 100
```

A seguinte consulta SQL tem parâmetros de consulta que se vinculam às portas de entrada Employee\_ID e Dept de uma transformação SQL:

```
SELECT Name, Address FROM Employees WHERE Employee_Num =?Employee_ID? and Dept = ?Dept?
```

A origem poderia ter as seguintes linhas:

Employee_ID	Dept
100	Products
123	HR

Employee_ID	Dept
130	Accounting

O Data Integration Service gera as seguintes instruções de consulta a partir das linhas:

```
SELECT Name, Address FROM Employees WHERE Employee_ID = '100' and DEPT = 'Products'
SELECT Name, Address FROM Employees WHERE Employee_ID = '123' and DEPT = 'HR'
SELECT Name, Address FROM Employees WHERE Employee_ID = '130' and DEPT = 'Accounting'
```

## Substituição de Strings

Use variáveis de string para substituir componentes de instruções de consulta. Por exemplo, você pode usar a variável de string para substituir o nome da tabela em uma consulta. Ou pode substituir os nomes de coluna em uma instrução SELECT.

Para substituir o nome da tabela, configure uma porta de entrada para receber o nome da tabela de cada linha de entrada. No Editor SQL, selecione a porta da lista de portas **Substituição de Strings**. A ferramenta Developer identifica a porta de entrada pelo nome na consulta e coloca o nome entre dois sinais de til (~).

A seguinte consulta contém uma variável de string, ~Table\_Port~:

```
SELECT Emp_ID, Address from ~Table_Port~ where Dept = 'HR'
```

A origem pode passar os seguintes valores para a coluna **Table\_Port**:

### Table\_Port

Employees\_USA

Employees\_England

Employees\_Australia

O Data Integration Service substitui a variável ~Table\_Port~ pelo valor do nome da tabela na porta de entrada:

```
SELECT Emp_ID, Address from Employees_USA where Dept = 'HR'
SELECT Emp_ID, Address from Employees_England where Dept = 'HR'
SELECT Emp_ID, Address from Employees_Australia where Dept = 'HR'
```

# Cardinalidade de Linha de Entrada para Linha de Saída

Quando o Data Integration Service executa uma consulta SELECT, a transformação SQL retorna uma linha para cada linha que ela recupera. Quando a consulta não recupera dados, a transformação SQL retorna zero ou uma linha para cada linha de entrada.

### Processamento de instruções de consulta

Quando uma consulta SELECT é bem-sucedida, a transformação SQL pode recuperar várias linhas. Quando a consulta contém outras instruções, o Data Integration Service pode gerar uma linha que contém erros de SQL ou o número de linhas afetadas.

### Configuração de porta

A porta de saída NumRowsAffected contém o número de linhas alteradas por uma instrução UPDATE, INSERT ou DELETE para uma linha de entrada. A transformação SQL retorna o número de linhas afetadas para cada instrução em uma consulta. Quando a transformação SQL contém portas de passagem, ela retorna os dados da coluna pelo menos uma vez para cada linha de origem.

### Configuração de contagem máxima de linhas

A opção Contagem Máx. de Linhas de Saída limita o número de linhas que a transformação SQL retorna de consultas SELECT.

### Linhas de erro

O Data Integration Service retorna linhas de erro quando encontra erros de conexão ou de sintaxe. A transformação SQL retorna erros para a porta SQL\_Error.

### Continuar com Erro de SQL

Você pode configurar a transformação SQL para continuar processando quando há um erro em uma instrução SQL. A transformação SQL não gera um erro de linha.

## Processamento de Instruções de Consulta

O tipo de consulta SQL determina quantas linhas a transformação SQL retorna. A transformação SQL pode retornar zero, uma ou várias linhas. Quando a consulta contém uma instrução SELECT, a transformação SQL retorna cada coluna do banco de dados para uma porta de saída. A transformação retorna todas as linhas qualificadas.

A tabela a seguir lista as linhas de saída geradas pela transformação SQL para diferentes tipos de instruções de consulta quando não ocorrem erros no modo de consulta:

Instrução de Consulta	Linhas de Saída
Apenas UPDATE, INSERT, DELETE	Uma linha para cada instrução da consulta.
Uma ou mais instruções SELECT	Número total de linhas do banco de dados recuperadas.
Consultas DDL, como CREATE, DROP, TRUNCATE	Uma linha para cada instrução da consulta.

## Configuração de Porta

Quando você ativa a opção Incluir Estatísticas como Saída, a ferramenta Developer cria a porta NumRowsAffected. O Data Integration Service retorna pelo menos uma linha com o NumRowsAffected, com base nas instruções da consulta SQL.

A tabela a seguir lista as linhas de saída geradas pela transformação SQL se você ativa NumRowsAffected:

Instrução de Consulta	Linhas de Saída
Apenas UPDATE, INSERT, DELETE	Uma linha para cada instrução com o NumRowsAffected referente à instrução.
Uma ou mais instruções SELECT	Número total de linhas do banco de dados recuperadas. O valor de NumRowsAffected é zero em cada linha.
Consultas DDL, como CREATE, DROP, TRUNCATE	Uma linha com zero NumRowsAffected.

## Contagem Máxima de Linhas de Saída

Você pode limitar o número de linhas retornadas pela transformação SQL para consultas SELECT. Configure a propriedade **Contagem Máx. de Linhas de Saída** para limitar o número de linhas. Quando uma consulta contém várias instruções SELECT, a transformação SQL limita o total de linhas de todas as instruções SELECT.

Por exemplo, você define **Contagem Máx. de Linhas de Saída** como 100. A consulta contém duas instruções SELECT:

```
SELECT * FROM table1; SELECT * FROM table2;
```

Se a primeira instrução SELECT retorna 200 linhas, e a segunda instrução SELECT retorna 50 linhas, a transformação SQL retorna 100 linhas a partir da primeira instrução SELECT. A transformação SQL não retorna nenhuma linha da segunda instrução.

Para configurar linhas de saída ilimitadas, defina **Contagem Máx. de Linhas de Saída** como zero.

## Linhas de Erro

O Data Integration Service retorna erros de linha quando encontra um erro de conexão ou de sintaxe. A transformação SQL retorna erros de SQL para a porta SQLException.

Quando você configura uma porta de passagem ou a porta NumRowsAffected, a transformação SQL retorna, pelo menos, uma linha para cada linha de origem. Quando uma consulta não retorna dados, a transformação SQL retorna os dados de passagem e os valores NumRowsAffected, mas retorna valores nulos nas portas de saída. Você pode remover linhas com valores nulos passando as linhas de saída por uma transformação Filtro.

A seguinte tabela descreve as linhas que a transformação SQL gera para instruções de consulta UPDATE, INSERT ou DELETE:

Porta NumRowsAffected ou porta de passagem configurada	SQLException	Saída das linhas
Nenhuma das portas configurada	Não	Uma linha com NULL na porta SQLException.
Nenhuma das portas configurada	Sim	Uma linha com o erro na porta SQLException.
Uma das portas configurada	Não	Uma linha para cada instrução de consulta com os dados da coluna de passagem ou NumRowsAffected.
Uma das portas configurada	Sim	Uma linha com o erro na porta SQLException, a porta NumRowsAffected ou os dados da porta de passagem.

A seguinte tabela descreve o número de linhas de saída geradas pela transformação SQL para instruções SELECT:

Porta NumRowsAffected ou porta de passagem configurada	SQLException	Saída das linhas
Nenhuma das portas configurada	Não	Uma ou mais linhas, com base nas linhas retornadas de cada instrução SELECT.
Nenhuma das portas configurada	Sim	Uma linha maior que a soma das linhas de saída para as instruções bem-sucedidas. A última linha contém o erro na porta SQLException.
Uma das portas configurada	Não	Uma ou mais linhas, dependendo das linhas retornadas para cada instrução SELECT. <ul style="list-style-type: none"> <li>- Se NumRowsAffected estiver habilitada, cada linha conterá uma coluna NumRowsAffected com um valor zero.</li> <li>- Se uma porta de passagem estiver configurada, cada linha conterá os dados da coluna de passagem. Quando a consulta retorna várias linhas, os dados da coluna de passagem são duplicados em cada linha.</li> </ul>
Uma das portas configurada	Sim	Uma ou mais linhas, dependendo das linhas retornadas para cada instrução SELECT. A última linha contém o erro na porta SQLException: <ul style="list-style-type: none"> <li>- Quando a porta NumRowsAffected está ativada, cada linha contém uma coluna NumRowsAffected com o valor zero.</li> <li>- Se uma porta de passagem estiver configurada, cada linha conterá os dados da coluna de passagem. Quando a consulta retorna várias linhas, os dados da coluna de passagem são duplicados em cada linha.</li> </ul>

A seguinte tabela descreve o número de linhas de saída geradas pela transformação SQL para consultas DDL como CREATE, DROP ou TRUNCATE:

Porta NumRowsAffected ou porta de passagem configurada	SQLException	Saída das linhas
Nenhuma das portas configurada	Não	Uma linha com NULL na porta SQLException.
Nenhuma das portas configurada	Sim	Uma linha que contém o erro da porta SQLException.

Porta NumRowsAffected ou porta de passagem configurada	SQLError	Saída das linhas
Uma das portas configurada	Não	Uma linha que inclui a coluna NumRowsAffected com valor zero e os dados da coluna de passagem.
Uma das portas configurada	Sim	Uma linha com o erro na porta SQLError, a coluna NumRowsAffected com valor zero e os dados da coluna de passagem.

## Continuar com Erro de SQL

Você pode optar por ignorar um erro de SQL que ocorre em uma instrução de consulta. Ative a opção **Continuar com Erro de SQL dentro de uma Linha**. O Data Integration Service continua a executar o restante das instruções SQL para a linha.

O Data Integration Service não gera um erro de linha. Entretanto, a porta SQLError contém a instrução SQL com falha e as mensagens de erro.

Por exemplo, uma consulta pode ter as seguintes instruções:

```
DELETE FROM Persons WHERE FirstName = 'Ed';
INSERT INTO Persons (LastName, Address) VALUES ('Gein', '38 Beach Rd')
```

Se a instrução DELETE falha, a transformação SQL retorna uma mensagem de erro do banco de dados. O Data Integration Service continua processando a instrução INSERT.

Desative a opção **Continuar com Erro de SQL** para solucionar problemas de erros de banco de dados e para associar os erros às instruções de consulta que os causaram.

## Otimização de Filtro com a Transformação SQL

O Serviço de Integração de Dados pode aplicar um método de otimização de filtro com uma transformação SQL se condição de filtro faz referência somente a portas de passagem e a transformação SQL não tem efeitos colaterais.

A transformação SQL tem efeitos colaterais nas seguintes circunstâncias:

- A consulta SQL atualiza um banco de dados. A consulta SQL contém uma instrução como CREATE, DROP, INSERT, UPDATE, GRANT ou REVOKE.
- A transformação retorna linhas NULL para instruções SELECT que não retornam resultados. As linhas podem conter valores de porta de passagem, informações de erro de SQL ou o campo NumRowsAffected.

O Serviço de Integração de Dados pode aplicar a seleção antecipada e métodos de otimização de envio com a transformação SQL.

## Otimização de Seleção Antecipada com a Transformação SQL

O Data Integration Service pode realizar a otimização de seleção antecipada com uma transformação SQL quando a condição de filtro faz referência somente portas de passagem, e a transformação SQL não apresenta efeitos colaterais.

A transformação SQL apresenta efeitos colaterais nas seguintes circunstâncias:

- A consulta SQL atualiza um banco de dados. A consulta SQL contém uma instrução como CREATE, DROP, INSERT, UPDATE, GRANT ou REVOKE.
- A transformação retorna linhas NULL para instruções SELECT que não retornam resultados. As linhas podem conter valores de porta de passagem, informações de erros SQL ou o campo NUMRowsAffected.

### Habilitando a Otimização de Seleção Antecipada com a Transformação SQL

Habilite a otimização com seleção antecipada na transformação SQL quando esta última não apresenta efeitos colaterais.

1. Habilite a opção **Retornar Somente Saída de Banco de Dados** nas **Propriedades Avançadas** da transformação SQL.
2. Desmarque a opção **Tem Efeitos Colaterais** nas **Propriedades Avançadas** da transformação.
3. Se a transformação tiver uma porta **NumAffectedRows**, remova essa porta.

## Otimização de Envio com a Transformação SQL

Com a otimização de envio, o Serviço de Integração de Dados envia a lógica de filtro de uma transformação de Filtro no mapeamento para a consulta na transformação SQL.

Use as seguintes regras e diretrizes quando você habilitar a otimização de envio com a transformação SQL:

- A consulta de transformação SQL deve conter somente instruções SELECT.
- A consulta SQL da transformação deve ser uma subconsulta válida.
- A condição de filtro não pode fazer referência aos campos Erro SQL ou NumRowsAffected.
- Os nomes das portas de saída devem corresponder aos nomes das colunas na instrução SQL SELECT. Quando você faz referência a uma porta de saída em uma condição de filtro, o Serviço de Integração de Dados envia o nome da porta de entrada correspondente à consulta SQL. Você pode adicionar aliases ao SQL quando as colunas na consulta não correspondem aos nomes das portas de saída. Por exemplo, `SELECT mycolname1 AS portname1, mycolname2 AS portname2.`
- A transformação não pode ter efeitos colaterais.

### Otimização de Envio com a Transformação SQL - Exemplo

Uma transformação SQL recupera pedidos por ID de cliente. Uma transformação de Filtro que aparece após a transformação SQL retorna apenas as linhas em que a quantidade de pedidos é maior que 1000.

O Serviço de Integração de Dados envia o seguinte filtro em uma instrução SELECT na transformação SQL:

```
orderAmount > 1000
```

Cada instrução na consulta SQL se torna uma subconsulta separada da instrução SELECT que contém o filtro.

A seguinte instrução de consulta mostra a instrução de consulta original como uma subconsulta na instrução SELECT:



```
SELECT <customerID>, <orderAmount>, ... FROM (instruções de consulta originais) ALIAS WHERE  
<orderAmount> > 1000
```

Se a consulta SQL tiver várias instruções, cada uma será incluída em uma subconsulta separada. A subconsulta tem a mesma sintaxe, incluindo a cláusula WHERE.

As portas *customerID* e *orderAmount* são os nomes das portas de saída na transformação SQL. A subconsulta não inclui portas de passagem, o erro SQL ou as portas de estatísticas SQL. Se você enviar vários filtros para a transformação SQL, a cláusula WHERE conterá todos esses filtros.

## Habilitando a Otimização de Envio com a Transformação SQL

Habilite a otimização de envio configurando propriedades na guia **Propriedades Avançadas** da transformação SQL.

1. Desmarque **Tem Efeitos Colaterais**.
2. Habilite **Retornar Somente Saída de Banco de Dados**.
3. Defina **Número de Linhas de Saída Máximas** como zero.
4. Habilite a otimização de envio.

# Exemplo de transformação SQL com uma Consulta SQL

Você é um desenvolvedor do departamento de RH da Hypostores Corporation. A Hypostores mantém as informações de folha de pagamento dos funcionários em um banco de dados separado dos dados dos funcionários de recursos humanos. O departamento de Recursos Humanos precisa consultar uma única exibição dos funcionários e salários em todas as regiões.

Você quer criar um mapeamento de objetos de dados lógicos que mostre uma única exibição dos dados de funcionários e de salários em um objeto de dados lógicos de funcionários.

Crie um mapeamento de objetos de dados lógicos com a fonte de dados de funcionários. Inclua uma transformação SQL para recuperar o salário e a data de contratação do banco de dados de folha de pagamento.

## Mapeamento de Objeto de Dados Lógicos

O mapeamento de objetos de dados lógicos contém os seguintes objetos:

### Tabela Funcionário

Tabela relacional de entrada de dados de funcionários do banco de dados de Recursos Humanos.

### Tabela salário

Uma tabela no banco de dados de Folha de Pagamento que contém o salário e a data de contratação dos funcionários. O banco de dados é Oracle.

### Transformação SQL

Transformação que recupera a data de contratação e o salário para cada linha de funcionário. A transformação se conecta a um banco de dados de Folha de Pagamento e executa uma consulta SQL na tabela Salário no banco de dados.

### Objeto de dados lógicos

Contém a exibição combinada de dados do funcionário e de salário. O objeto de dados lógicos recebe a saída da transformação SQL.

### Arquivo SQLErrors

O arquivo SQLErrors é um arquivo simples que contém qualquer erro de SQL do banco de dados. O Data Integration Service grava pelo menos uma linha no arquivo SQLErrors para cada linha de entrada. Se não ocorre nenhum erro de SQL, a coluna SQLError contém o valor NULL. Examine o arquivo SQLErrors para solucionar problemas de erros.

## Tabela Salário

A tabela Salário é uma tabela relacional no banco de dados de Folha de Pagamento. A tabela contém dados de funcionários mantidos pelo departamento de Folha de Pagamento. A transformação SQL recupera a data de contratação e o salário do funcionário da tabela Salário.

A tabela a seguir mostra algumas linhas da tabela Salário:

Employee_Num	HireDate	Salário
10	3-May-97	232000
11	11-Sep-01	444000
12	17-Oct-89	656000
13	13-Aug-07	332100

## Tabela Funcionário

A origem é a tabela Funcionário do banco de dados de Recursos Humanos.

A seguinte tabela mostra exemplos de linhas da tabela Funcionário:

EmpID	LastName	FirstName	DeptId	Telefone
10	Smith	Martha	FIN	(415) 552-1623
11	Jones	Cynthia	ENG	(415) 552-1744
12	Russell	Cissy	SLS	(415) 552-1656
13	Goyal	Girish	FIN	(415) 552-1656

## Transformação SQL

A transformação SQL recupera a data de contratação e o salário do funcionário da tabela Salário do banco de dados de Folha de Pagamento. A tabela Salário está em um banco de dados Oracle.

Use as seguintes etapas para configurar a transformação SQL:

1. Configure as propriedades da transformação SQL.
2. Defina as portas.

3. Crie a consulta SQL.
4. Configure a conexão de banco de dados para a transformação SQL.

## Definir Propriedades da Transformação SQL

Configure as propriedades da transformação SQL na exibição **Propriedades Avançadas**.

Configure as seguintes propriedades:

### Tipo de banco de dados

O tipo de banco de dados é Oracle. Ao definir as portas, você pode escolher tipos de dados de porta que sejam aplicáveis para Oracle.

### Continuar com Erro Dentro da Linha

Desativar. Interromper o processamento se ocorrer um erro de SQL na linha.

### Incluir Estatísticas como Saída

Desativar. Não criar a porta de saída NumRowsAffected.

## Definir as Portas

Definir portas de entrada para cada coluna na tabela de origem de funcionários. Selecione **Copiar para Saída** para converter as portas de entrada em portas de passagem para as colunas. Quando você seleciona **Copiar para Saída**, a ferramenta Developer cria a porta de saída correspondente para cada porta que você copia.

Crie as seguintes portas de passagem de entrada:

Nome	Tipo	Tipo Nativo	Precisão	Escala	Copiar para Saída
EmpID	decimal	number(p,2)	4	0	x
LastName	string	varchar2	30	0	x
FirstName	string	varchar2	20	0	x
DeptID	string	varchar2	4	0	x
Telefone	string	varchar2	16	0	x

A transformação SQL tem as seguintes portas de saída:

Nome	Tipo	Tipo Nativo	Precisão	Escala
EmpID	decimal	number(p,s)	4	0
LastName	string	varchar2	30	0
FirstName	string	varchar2	20	0
DeptID	string	varchar2	4	0
Telefone	string	varchar2	16	0

Nome	Tipo	Tipo Nativo	Precisão	Escala
HireDate	date/time	timestamp	29	0
Salário	decimal	number(p,s)	8	2

A ferramenta Developer adiciona o sufixo "\_output" a cada porta de saída que ele cria quando você seleciona **Copiar para Saída**.

Defina manualmente as portas de saída para as colunas de data de contratação e de salário. A transformação SQL retorna as colunas de data de contratação e de salário da tabela Salário nas portas.

## Definir a Consulta SQL

Crie uma consulta SQL para selecionar a data de contratação e o salário de cada funcionário na tabela Salário.

Defina a consulta na exibição SQL da transformação SQL.

Insira a seguinte consulta no Editor SQL:

```
select HIREDATE,SALARY,from Salary where EMPLOYEE_NUM =?EmpID?
```

Hiredate, Salary e Employee\_Num são nomes de coluna na tabela Salário.

?EMPID? é um parâmetro que contém o valor da porta EmpID.

## Definir a Conexão do Banco de Dados

Na exibição **Tempo de Execução**, selecione um objeto de conexão de banco de dados para o banco de dados ao qual a transformação SQL se conecta. Selecione um objeto de conexão do banco de dados Oracle.

## Saída

Conecte a porta SQLException e a porta EmpID\_output ao arquivo simples SQLErrors. A porta SQLException contém valores nulos, a menos que ocorra um erro de SQL.

Conecte EmpID e as outras portas de saída ao objeto de dados lógicos.

A transformação SQL retorna uma linha que contém dados da tabela Funcionário e inclui a data de contratação e o salário da tabela Salário.

A seguinte tabela mostra algumas linhas do objeto de dados lógicos:

EmpID	LastName	FirstName	DeptId	Telefone	HireDate	Salário
10	Smith	Martha	FIN	(415) 552-1623	19970303 00:00:00	2320.00
11	Jones	Cynthia	ENG	(415) 552-1744	20010911 00:00:00	4440.00
12	Russell	Cissy	SLS	(415) 552-1656	19891017 00:00:00	6560.00
13	Goyal	Girish	FIN	(415) 552-1660	20070813 00:00:00	3210.00

# Procedimentos Armazenados

Você pode chamar um procedimento armazenado a partir de uma transformação SQL. Você pode usar um procedimento armazenado para automatizar tarefas em um banco de dados relacional. Procedimentos armazenados aceitam variáveis definidas pelo usuário, instruções condicionais e outros recursos sem suporte em instruções SQL padrão.

A transformação SQL se conecta a um banco de dados relacional para executar o procedimento armazenado. A transformação SQL pode chamar procedimentos armazenados do Oracle, do IBM DB2, do Microsoft SQL Server, do Sybase e do ODBC. Um procedimento armazenado é mantido e executado no banco de dados.

Crie uma conexão ODBC para chamar um procedimento armazenado de um banco de dados Sybase. Você também deverá criar uma conexão ODBC para chamar um procedimento armazenado de um banco de dados Microsoft SQL Server em sistemas operacionais não Windows.

Um procedimento armazenado é uma coleção pré-compilada de instruções Transact-SQL, PL-SQL ou outras instruções de procedimento de banco de dados. A sintaxe de um procedimento armazenado varia de acordo com o banco de dados.

Você pode usar procedimentos armazenados para realizar as seguintes tarefas:

- Verifique o status do banco de dados de destino antes de carregar dados para ele.
- Determine se há espaço suficiente no banco de dados.
- Execute um cálculo especializado.
- Recupere dados por um valor.
- Descarte e recrie índices.

É possível usar um procedimento armazenado para fazer uma consulta ou um cálculo que, de outra forma, deve ser incluído em uma transformação. Por exemplo, se você já tiver um procedimento armazenado testado para calcular impostos sobre vendas, poderá usá-los para fazer cálculos em vez de recriar o mesmo cálculo em uma transformação de Expressão.

Um procedimento armazenado pode aceitar uma entrada e depois retornar um conjunto de resultados de linhas. Um procedimento armazenado pode executar uma tarefa DDL que não exige entrada e que não retorna uma saída.

Você pode configurar a transformação SQL para executar mais de um procedimento armazenado. Para cada procedimento armazenado configurado, defina portas de transformação para corresponder os parâmetros de procedimento armazenado. Cada procedimento armazenado pode transmitir dados a portas de saída.

O banco de dados que contém o procedimento armazenado tem permissões de usuários. Você deve ter permissões para executar o procedimento armazenado no banco de dados.

**Nota:** Uma função armazenada é semelhante a um procedimento armazenado, exceto que uma função retorna um valor único. A transformação SQL pode executar funções armazenadas.

## Portas da Transformação SQL para Procedimentos Armazenados

As portas de entrada e saída da transformação SQL correspondem aos parâmetros de entrada e saída no procedimento armazenado.

Quando você importa um procedimento armazenado, a ferramenta Developer determina o tipo de banco de dados com base na conexão de banco de dados. Ela gera portas de entrada e saída na transformação SQL usando os parâmetros no procedimento armazenado. A ferramenta Developer determina o tipo de dados nativo de cada porta com base no parâmetro do procedimento armazenado.

Ao configurar manualmente a transformação SQL, você precisa configurar as portas de entrada e saída nessa transformação. Quando você configura o tipo de banco de dados, a transformação SQL altera o tipo de dados nativo de cada porta com base no tipo de banco de dados inserido.

Os seguintes tipos de dados pode ser transmitidos entre a transformação SQL e o procedimento armazenado:

#### Parâmetros de entrada e saída

A transformação SQL envia parâmetros ao procedimento armazenado e recebe parâmetros desse procedimento armazenado nas portas de entrada e saída.

#### Valor de retorno

Se o procedimento armazenado transmitir um valor de retorno, a ferramenta Developer criará uma porta de Valor de Retorno.

#### Erros de SQL

A transformação SQL retorna erros do procedimento armazenado na porta `SQL_Error`.

## Parâmetros de Entrada e Saída

Quando um procedimento armazenado é chamado em uma transformação SQL, cada campo referenciado pela instrução de chamada identifica uma porta de entrada ou saída. Quando você importa um procedimento armazenado, a ferramenta Developer gera uma instrução de chamada de procedimento armazenado. Caso contrário, você precisará configurar manualmente a instrução da chamada.

Você pode editar a instrução da chamada na exibição **SQL** da transformação.

A instrução da chamada tem o seguinte formato:

```
?RETURN_VALUE? = chamada <nome do procedimento armazenado>(?Field1?, ?Field2?, . . . )
```

Coloque os nomes de portas entre pontos de interrogação. Os nomes de porta não precisam corresponder aos nomes dos parâmetros no procedimento armazenado. As portas de saída devem estar na mesma ordem que os parâmetros em uma consulta `SELECT`.

Você pode usar um procedimento armazenado que contém parâmetros `INOUT`. A transformação SQL identifica parâmetros `INOUT` de acordo com o nome da porta de entrada. A porta de saída tem o prefixo `output_`. O Serviço de Integração de Dados associa a porta de entrada e a porta de saída ao mesmo parâmetro.

Você pode configurar uma transformação SQL para retornar um conjunto de resultados. Quando o procedimento armazenado retorna um conjunto de resultados, a ferramenta Developer não pode criar portas de saída para as colunas no conjunto de resultados. Ao importar o procedimento armazenado, você deve inserir manualmente as portas e configurar a chamada de procedimento armazenado.

## Valor de Retorno

Um valor de retorno é um código ou uma string de texto que define o status do procedimento armazenado. Quando o procedimento armazenado tem um valor de retorno, a transformação SQL tem uma porta de **Valor de Retorno**.

A maioria dos bancos de dados fornece um valor de retorno depois de executar um procedimento armazenado. Esse valor de retorno pode conter um valor integer ou pode conter um valor que você define no procedimento armazenado. Por exemplo, um procedimento armazenado pode retornar "Êxito" quando o procedimento é bem-sucedido.

Se um procedimento armazenado retornar um conjunto de resultados em vez de um único valor de retorno, a transformação SQL receberá o primeiro valor retornado desse procedimento.

## Conjuntos de Resultados de Procedimentos Armazenados

Você pode configurar uma transformação SQL para receber um conjunto de resultados de um procedimento armazenado. Um procedimento armazenado retorna várias linhas em um conjunto de resultados. A transformação SQL pode retornar cada linha para o mapeamento.

O seguinte procedimento armazenado retorna um conjunto de resultados:

```
CREATE OR REPLACE FUNCTION fetchEMPinfo
(p_State IN VARCHAR2 )
return types.cursorType
AS
my_cursor types.cursorType;
BEGIN
OPEN my_cursor FOR SELECT EMP_ID, NAME, CITY FROM EMP WHERE STATE = p_State ORDER BY
EMP_ID;
RETURN my_cursor;
END;
```

Quando você importar o procedimento armazenado, a ferramenta Developer criará uma instrução de chamada de procedimento armazenado com uma sintaxe semelhante à seguinte:

```
call FETCHEMPINFO (?P_STATE?)
```

O parâmetro de entrada é `p_state`. A ferramenta Developer não cria as portas de saída para você. Você deve criar as portas de saída manualmente com os mesmos tipos de dados dos parâmetros do procedimento armazenado.

Por exemplo, `resultSet` contém as colunas `EMP_ID`, `EMPNAME` e `CITY`. Crie portas de saída para essas colunas.

Você também deve atualizar manualmente a chamada SQL com as colunas de saída usando a seguinte sintaxe:

```
(?EMP_ID?,?EMPNAME?,?CITY?) = call FETCHEMPINFO (?P_STATE?)
```

## Conjuntos de Resultados com Diferentes Bancos de Dados

Configure procedimentos armazenados para retornar conjuntos de resultados usando uma sintaxe diferente com base no tipo de banco de dados.

### Oracle

Uma função armazenada Oracle retorna resultados com um cursor:

```
create or replace function sp_ListEmp return types.cursorType
as
l_cursor types.cursorType;
begin
open l_cursor for select ename, empno from emp order by ename;
return l_cursor;
end;
```

O Oracle só aceita cursores como parâmetros de entrada. Você pode configurar cursores como parâmetros de entrada com a transformação SQL.

## Microsoft SQL Server

Um procedimento armazenado do Microsoft SQL Server retorna um procedimento armazenado de conjuntos de resultados com uma instrução SELECT no corpo do procedimento ou com o tipo de retorno explicitamente declarado como uma tabela.

```
Create PROCEDURE InOut(  
  @inout varchar(100) OUT  
)  
AS  
BEGIN  
  set @inout = concat(@inout, '__')  
  select * from mytable;  
END
```

## IBM DB2

Um procedimento armazenado do IBM DB2 pode retornar um conjunto de resultados com um cursor aberto. O número de conjuntos de resultados retornados são declarados na cláusula RESULT SET. O procedimento armazenado abre um cursor e o retorna. O exemplo a seguir retorna 2 cursores abertos.

```
CREATE PROCEDURE TESTMULTIRS  
  (IN i_cmacct CHARACTER(5))  
RESULT SETS 2  
LANGUAGE SQL  
BEGIN  
  
  DECLARE csnum INTEGER;  
  
  --Declare serial cursors to consume less resources  
  --You do not need a rollable cursor.  
  
  DECLARE getDeptNo CHAR(50); --Be careful with the estimated length.  
  DECLARE getDeptName CHAR(200);  
  DECLARE c1 CURSOR WITH RETURN FOR s1;  
  SET getDeptNo = 'SELECT DEPTNO FROM DEPT';  
  SET getDeptName = 'SELECT DEPTNAME FROM DEPT';  
  
  PREPARE s1 FROM getDeptNo;  
  OPEN c1;  
  
  END;
```

## Sybase

Um procedimento armazenado do Sybase retorna um procedimento armazenado de conjunto de resultados com uma instrução select no corpo do procedimento ou com o tipo de retorno explicitamente declarado como uma tabela.

```
CREATE PROCEDURE FETCHEMPINFO  
(  
  @p_State VARCHAR(5),  
  @e_id INT OUTPUT,  
  @e_name VARCHAR(50) OUTPUT  
)  
AS  
BEGIN  
  SET NOCOUNT ON  
  SELECT EMP_ID, NAME FROM EMP WHERE STATE = @p_State ORDER BY EMP_ID  
  SET NOCOUNT OFF  
  SELECT @e_id AS EMP_ID, @e_name AS NAME  
  RETURN  
END  
GO  
  
--Configure the following variables to execute the procedure.  
  
DECLARE @p_State VARCHAR(5)  
DECLARE @EMPID int
```



```

DECLARE @EMPNAME varchar(50)

SET @p_State = 'CA'
exec FETCHEMPINFO @p_State, @e_id = @EMPID, @e_name = @EMPNAME
GO

```

## Linhas do Conjunto de Resultados

Alguns procedimentos armazenados retornam parâmetros de saída além de linhas de conjuntos de resultados. A transformação SQL retorna parâmetros de saída na última linha. Ela não inclui os parâmetros de saída de ocorrência única nas linhas de conjuntos de resultados.

Por exemplo, você grava um procedimento armazenado que recebe um ID de funcionário e retorna o nome do funcionário no parâmetro de saída 1 e o departamento no parâmetro de saída 2. O procedimento armazenado também retorna uma linha para cada dia de licença que o funcionário teve ao longo do ano. A linha contém a data, o número de horas e o motivo da ausência.

O conjunto de resultados contém um número diferente de linhas para cada funcionário. Cada linha do conjunto de resultados contém um nome de funcionário e um departamento vazios. A transformação SQL retorna o nome do funcionário e o departamento após o conjunto de resultados. O nome do funcionário e o departamento aparecem na última linha.

## Exemplo de Procedimento Armazenado

Você pode chamar um procedimento armazenado que retorna dados para a transformação SQL.

O seguinte procedimento armazenado recebe um número de funcionário e retorna uma linha com o número e o nome do funcionário:

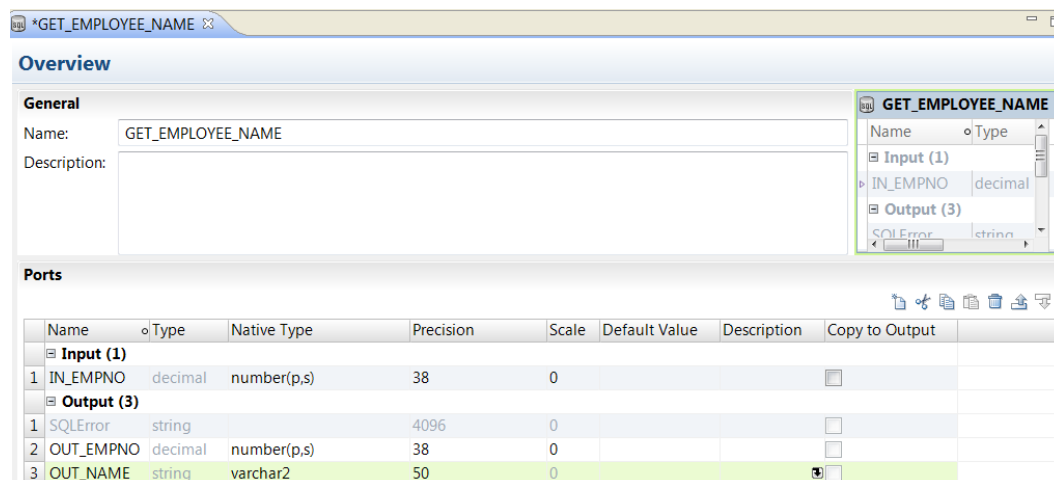
```

CREATE OR REPLACE PROCEDURE SP_GETNAME
(IN_EMPNO IN NUMBER, OUT_EMPNO NUMBER, OUT_NAME OUT STRING)
AS
BEGIN
SELECT EMP_KEY,EMP_NAME into OUT_EMPNO , OUT_NAME from EMP_TABLE where EMP_KEY=IN_EMPNO;
END;/"

```

Para criar a transformação SQL, importe o procedimento armazenado. A ferramenta Developer cria as portas de entrada e as portas de saída. Os nomes das portas são iguais aos nomes dos parâmetros no procedimento armazenado.

A seguinte figura mostra as portas para a transformação SQL:



A ferramenta Developer cria a seguinte chamada de procedimento armazenado para recuperar o nome do funcionário:

```
call SP_GETNAME (?IN_EMPNO?, ?OUT_EMPNO?, ?OUT_NAME?)
```

É possível exibir a chamada de procedimento armazenado no Editor SQL. Os erros de SQL aparecem na porta SQLError.

## Conexão de Transformação SQL

Configure a conexão de transformação SQL nas propriedades de tempo de execução da transformação. Talvez você precise configurar uma conexão de tempo de execução se não tiver especificado uma conexão ao criar a transformação.

Você pode definir um nome de conexão diferente da conexão que selecionou para criar a transformação SQL. Você deve selecionar uma conexão do mesmo tipo de banco de dados nas propriedades **Avançadas** da transformação SQL.

Você pode configurar um parâmetro para o nome da conexão da transformação SQL. Você deve definir o parâmetro na exibição **Parâmetros** do mapeamento antes atribuí-lo a uma conexão de tempo de execução.

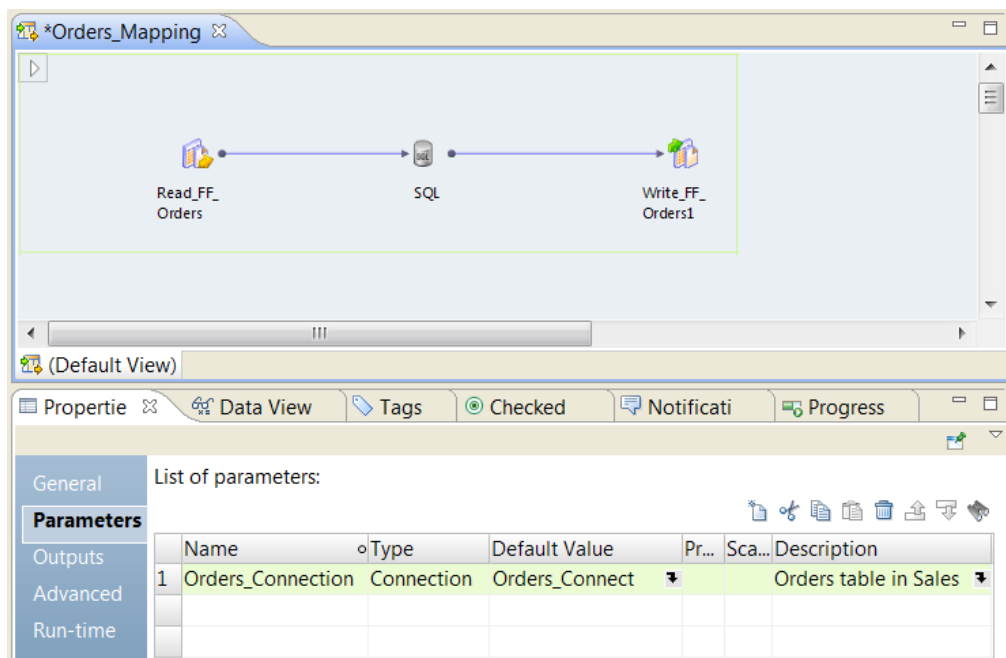
### Criando um Parâmetro de Nome da Conexão

É possível especificar um parâmetro definido pelo usuário em nome da conexão em tempo de execução para uma transformação SQL. A Developer tool cria um parâmetro de mapeamento para a conexão em vez de um parâmetro de transformação.

1. Crie um mapeamento que inclua a transformação SQL.
2. Clique na guia **Tempo de Execução** da transformação SQL.
3. Em **Nome da Conexão**, clique na seta de seleção e escolha **Atribuir Parâmetros**.
4. Na caixa de diálogo **Atribuir Parâmetro**, clique em **Novo**.
5. Na caixa de diálogo **Parâmetros**, insira um nome de parâmetro de conexão e uma descrição para esse parâmetro. O tipo de parâmetro assume conexão como padrão.
6. Clique em **OK** nas caixas de diálogo **Parâmetros** e **Atribuir Parâmetros**.

A Developer tool cria um parâmetro de mapeamento e o atribui ao Nome da Conexão. O nome do parâmetro aparece nas propriedades de Tempo de execução.

7. Para visualizar uma lista dos parâmetros de mapeamento, clique dentro do editor e depois clique na guia **Parâmetros** do mapeamento.



## Criando Manualmente uma Transformação SQL

É possível criar uma transformação SQL manualmente. Crie a transformação manualmente quando você configurar uma transformação que executa uma consulta SQL. Você também pode criar manualmente uma transformação que chama um procedimento armazenado quando o procedimento não está disponível para importação. Ao criar a transformação manualmente, você configura as portas de entrada e saída e digita as instruções SQL no Editor SQL.

1. Selecione um projeto ou pasta na exibição **Object Explorer**.
2. Clique em **Arquivo > Novo > Transformação**.  
A caixa de diálogo **Novo** é exibida.
3. Selecione a transformação SQL.
4. Clique em **Avançar**.
5. Selecione **Criar como Vazio**.
6. Insira um nome para a transformação e insira a localização do repositório para a transformação.
7. Clique em **Concluir**.
8. Clique na exibição **Visão Geral** para adicionar portas à transformação.
9. Para adicionar uma porta de entrada, clique em **Entrada** no painel **Portas** para indicar onde adicionar a porta. Clique no botão **Novo** e insira o nome da porta, o tipo nativo e a precisão.

O tipo de banco de dados padrão é Oracle. A ferramenta Developer mostra tipos nativos para bancos de dados Oracle, a menos que você altere o tipo de banco de dados na exibição **Avançado**.

10. Para adicionar uma porta de saída, clique em **Saída** no painel **Portas** antes de adicionar a porta. Clique no botão **Novo** e insira o nome da porta, o tipo nativo e a precisão.  
A porta **SQLError** é a primeira porta de saída por padrão.
11. Na exibição **Avançada**, selecione o tipo de banco de dados ao qual a transformação SQL se conecta. Configure outras propriedades avançadas para tratamento de erros e outras propriedades opcionais.  
Quando você escolhe o tipo de banco de dados, a ferramenta Developer altera os tipos de dados nativos das portas na exibição **Visão Geral**.
12. Digite a consulta SQL ou a chamada de procedimento armazenado na exibição **SQL**. Selecione portas para associação de parâmetros ou substituição de strings no **Editor SQL**.  
Se o procedimento armazenado retornar um conjunto de resultados, você deverá inserir a chamada de procedimento armazenado com uma sintaxe semelhante à seguinte: (?Field1?,?Field2?,?Field3?) = `call Stored_Procedure_Name (?Input_Parm?)`.

### TÓPICOS RELACIONADOS:

- [“Definir a Consulta SQL” na página 648](#)

## Criando uma Transformação SQL a partir de um Procedimento Armazenado

É possível configurar uma transformação SQL importando um procedimento armazenado de uma conexão de banco de dados.

1. Selecione um projeto ou pasta na exibição **Object Explorer**.
2. Clique em **Arquivo > Novo > Transformação**.  
A caixa de diálogo **Novo** é exibida.
3. Selecione a transformação SQL.
4. Clique em **Avançar**.
5. Selecione **Criar a partir de um procedimento armazenado existente**.
6. Procure e selecione uma conexão de banco de dados.
7. Procure e selecione o procedimento armazenado a ser importado.
8. Insira um nome e uma localização para a transformação.
9. Clique em **Concluir**.  
A ferramenta Developer cria as portas e a chamada de procedimento armazenado.
10. Se o procedimento armazenado retornar um conjunto de resultados, você deverá adicionar as portas de saída manualmente e, em seguida, reconfigurar a chamada de procedimento armazenado.
  - a. Na exibição **Visão Geral**, clique em **Saída** no painel **Portas**. Clique no botão **Novo** e insira o nome da porta de saída, o tipo nativo e a precisão.
  - b. Na exibição **SQL**, altere a chamada de procedimento armazenado de forma a usar a seguinte sintaxe: (?Field1?,?Field2?,?Field3?) = `call Stored_Procedure_Name (?Input_Parm?)`  
É possível selecionar os parâmetros de entrada e de saída na lista **Associação de Parâmetros** de portas no Editor SQL.

## CAPÍTULO 43

# Transformação de Padronizador

Este capítulo inclui os seguintes tópicos:

- [Transformação de Padronizador - Visão Geral, 657](#)
- [Estratégias de Padronização, 657](#)
- [Propriedades de Padronização, 658](#)
- [Configurando uma Estratégia de Padronização, 659](#)
- [Propriedades Avançadas da Transformação de Padronizador, 659](#)
- [Transformação de Padronizador em um Ambiente Não Nativo, 660](#)

## Transformação de Padronizador - Visão Geral

A transformação de Padronizador é uma transformação passiva que examina as strings de entrada e cria versões padronizadas delas.

A transformação de Padronizador cria colunas que contêm versões padronizadas das strings de entrada. A transformação pode substituir ou remover strings nos dados de entrada ao criar essas colunas.

Por exemplo, você pode usar a transformação de Padronizador para examinar uma coluna de dados de endereço que contenha as strings Avenida, Av. e AVE. Você pode substituir todas as instâncias dessas strings pela string Av.

Dentro de uma transformação de Padronizador, você pode criar várias estratégias de padronização. Cada estratégia pode conter várias operações de padronização. A transformação de Padronizador fornece um assistente usado para criar estratégias.

## Estratégias de Padronização

Use as estratégias de padronização para criar colunas com versões padronizadas das strings de entrada.

Ao configurar uma estratégia de padronização, você adiciona uma ou mais operações. Cada operação implementa uma tarefa de padronização específica.

Você pode adicionar os seguintes tipos de operações a uma estratégia de padronização:

#### **Substituir Correspondências da Tabela de Referência por Valores Válidos**

Substitua strings que correspondem valores de tabela de referência pelo valor "Válido" da tabela de referência.

#### **Substituir Correspondências da Tabela de Referência por Strings Personalizadas**

Substitua strings que correspondem a valores da tabela de referência por uma string de substituição definida pelo usuário.

#### **Remover Correspondências de Tabela de Referência**

Remova strings que correspondem a valores da tabela de referência.

#### **Substituir Strings Personalizadas**

Substitua strings definidas pelo usuário por uma string de substituição definida pelo usuário.

#### **Remover Strings Personalizadas**

Remova strings definidas pelo usuário.

**Importante:** Você pode alterar a ordem das operações. A ordem das operações pode alterar a saída de uma estratégia, porque cada operação lê os resultados da operação anterior.

## Propriedades de Padronização

Para configurar as propriedades para estratégias e operações de padronização, selecione a exibição **Estratégias** na transformação de Padronizador.

### Propriedades da Estratégia

As propriedades da estratégia aplicam-se a todas as operações dentro de uma estratégia. Você pode configurar as seguintes propriedades da estratégia:

#### **Remover vários espaços**

Substitui vários espaços consecutivos por um só.

#### **Remover espaços à direita e à esquerda**

Remove espaços do início e fim de strings de dados.

#### **Delimitadores**

Determina os delimitadores que definem os tokens de pesquisa. Por exemplo, se você escolher "Ponto e Vírgula", a transformação de Padronizador pesquisará a string "laranjas;maçãs;" e encontrará as strings "laranjas" e "maçãs". A transformação usa o delimitador de espaço por padrão.

### Propriedades da Operação

Você pode configurar propriedades para os tipos de operações de padronização a seguir.

#### **Operações da Tabela de Referência**

As operações da tabela de referência incluem as seguintes propriedades:

- **Tabela de referência.** Determina a tabela de referência que você usa para padronizar dados. Clique em **Procurar** para selecionar uma tabela de referência.
- **Distinção Entre Maiúsculas e Minúsculas.** Determina se as strings de entrada devem corresponder à distinção entre maiúsculas e minúsculas das entradas da tabela de referência.

- **Substituir por.** Substitui as strings de entrada correspondentes às entradas da tabela de referência pelo texto fornecido por você. Aplica-se somente a operações de substituição.
- **Escopo.** Especifica a parte da string de entrada que contém o valor da tabela de referência.

#### Operações de String Personalizadas

As operações de string personalizadas incluem as seguintes propriedades:

- **Corresponder tokens a.** Define as strings de pesquisa a serem localizadas em dados de entrada.
- **Substituir por.** Substitui strings de entrada que correspondem a strings de pesquisa especificadas por você. Aplica-se somente a operações de substituição.
- **Escopo.** Especifica a parte da string de entrada a ser pesquisada.

## Configurando uma Estratégia de Padronização

Para configurar uma estratégia de padronização, edite as configurações na exibição **Estratégias** da transformação de Padronizador.

1. Selecione a exibição **Estratégias**.
2. Clique em **Novo**.  
O assistente de **Nova Estratégia** é aberto.
3. Clique no campo **Entradas** para selecionar portas para a estratégia.
4. Configure as propriedades da estratégia e clique em **Avançar**.
5. Escolha uma operação e clique em **Avançar**.
6. Configure as propriedades da operação.
7. Opcionalmente, clique em **Avançar** para adicionar mais operações à estratégia.
8. Após adicionar todas as operações à estratégia, clique em **Concluir**.
9. Opcionalmente, adicione mais estratégias à transformação.
10. Opcionalmente, altere a ordem em que a transformação processa estratégias ou operações. Selecione uma estratégia ou operação e clique em **Mover para Cima** ou **Mover para Baixo**.

## Propriedades Avançadas da Transformação de Padronizador

Configure propriedades que ajudam a determinar como o Serviço de Integração de Dados processa dados para a transformação de Padronizador.

Você pode configurar níveis de rastreamento para logs.

Configure a seguinte propriedade na guia **Avançado**:

#### Nível de Rastreamento

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.

# Transformação de Padronizador em um Ambiente Não Nativo

O processamento da transformação de Padronizador em um ambiente não nativo depende do mecanismo que executa a transformação.

Considere o suporte para os seguintes mecanismos de tempo de execução não nativos:

- Mecanismo Blaze. Suportado sem restrições.
- Mecanismo Spark. Com suporte sem restrições em mapeamentos de fluxo e em lote.
- Mecanismo Databricks Spark. Sem suporte.



## CAPÍTULO 44

# Transformação de União

Este capítulo inclui os seguintes tópicos:

- [Transformação de União - Visão Geral, 661](#)
- [Grupos e Portas, 662](#)
- [Transformação de União - Propriedades Avançadas, 662](#)
- [Transformação de União - Processamento, 663](#)
- [Criando uma Transformação de União, 663](#)
- [Transformação de União em um ambiente não nativo, 664](#)

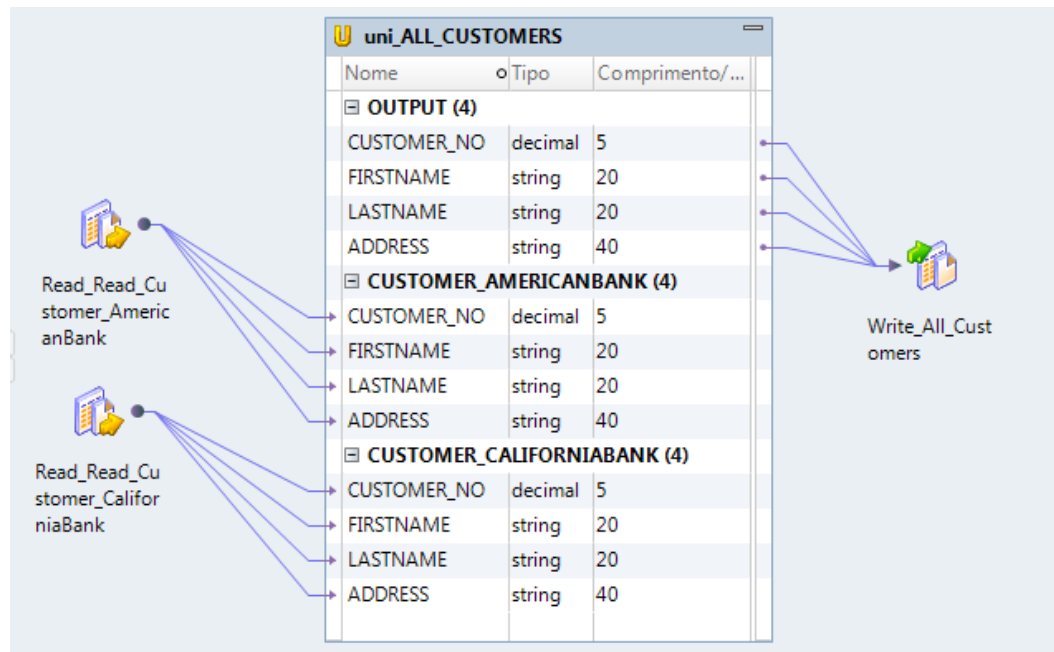
## Transformação de União - Visão Geral

Use a transformação de União para mesclar dados de vários pipelines ou ramificações em uma única ramificação de pipeline.

A transformação de União é uma transformação ativa com vários grupos de entrada e um grupo de saída. Ela mescla origens com portas correspondentes e transfere dados por meio de um grupo de saída que tenha a mesma estrutura de porta que os grupos de entrada. Use uma transformação de União na ferramenta Developer para mesclar dados de várias origens sem remover as linhas duplicadas.

Por exemplo, você deseja combinar dados de contas de clientes do American Bank e do California Bank. Você pode criar um mapeamento que contenha uma transformação de União para mesclar dados dos objetos de origem e gravar em um objeto de destino.

A figura a seguir mostra um mapeamento com uma transformação de União:



## Grupos e Portas

Uma transformação de União tem vários grupos de entrada e um grupo de saída. Você pode criar um ou mais grupos de entrada. A ferramenta Developer cria um grupo de saída. Você não pode criá-lo, nem editá-lo ou excluí-lo. Cada grupo deve ter portas correspondentes.

Para criar portas, você pode copiar portas de uma transformação ou pode criá-las manualmente. Quando você cria portas, a ferramenta Developer cria portas de entrada em cada grupo de entrada e portas de saída no grupo de saída. A ferramenta Developer usará os nomes de porta de saída que você especificar para cada porta de entrada e saída. A ferramenta Developer também usa os mesmos metadados para cada porta, como tipo de dados, precisão e escala.

Você pode conectar os grupos de entrada de diferentes ramificações em um único pipeline ou de diferentes pipelines de origem. Ao adicionar uma transformação de União a um mapeamento, você deve verificar se conectou as mesmas portas em todos os grupos de entrada. Se você conectar uma porta em um grupo de entrada, mas não conectar a mesma porta em outro grupo de entrada, o Data Integration Service passará NULLS para a porta desconectada.

## Transformação de União - Propriedades Avançadas

Configure propriedades que ajudem a determinar como o Serviço de Integração de Dados exibe os detalhes do log para a transformação de União.

Você pode configurar níveis de rastreamento para logs.

Configure a seguinte propriedade na guia **Avançado**:

### Nível de Rastreamento

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.

## Transformação de União - Processamento

Use a transformação de União para mesclar dados de vários pipelines ou ramificações em uma única ramificação de pipeline. O Data Integration Service processa todos os grupos de entrada em paralelo. Ele lê simultaneamente origens conectadas à transformação de União e transfere blocos de dados aos grupos de entrada da transformação. A transformação de União processa os blocos de dados com base na ordem em que os recebe do Data Integration Service. A transformação de União não bloqueia dados de entrada nos grupos de entrada.

## Criando uma Transformação de União

Você pode criar uma transformação de União reutilizável ou não.

### Criando uma Transformação de União Reutilizável

Crie uma transformação de União reutilizável para usar em vários mapeamentos ou mapplets.

1. Selecione um projeto ou pasta na exibição **Object Explorer**.
2. Clique em **Arquivo > Novo > Transformação**.  
A caixa de diálogo **Novo** é exibida.
3. Selecione a transformação de União.
4. Clique em **Avançar**.
5. Insira um nome para a transformação.
6. Clique em **Concluir**.  
A transformação aparece no editor.
7. Clique no botão **Novo** para adicionar uma porta à transformação.
8. Edite a porta para definir o nome, o tipo de dados e a precisão.
9. Selecione a exibição **Grupos**.
10. Clique no botão **Novo** para adicionar um grupo de entrada.
11. Clique na exibição **Avançado** e edite as propriedades da transformação.

### Criando uma Transformação de União Não Reutilizável

Crie uma transformação de União não reutilizável em um mapeamento ou mapplet.

1. Em um mapeamento ou mapplet, arraste uma transformação de União da paleta Transformação para o editor.  
A transformação aparece no editor.

2. Na guia **Geral**, edite o nome e a descrição da transformação.
3. Selecione todas as portas da transformação upstream e arraste-as para a transformação de União. As portas são exibidas como portas em um grupo de entrada e um grupo de saída da transformação de União.
4. Clique em **Novo** na guia **Grupos** da exibição **Propriedades** para adicionar um grupo de entrada. Outro grupo de entrada é exibido com portas semelhantes às do grupo de entrada existente.
5. Selecione as portas no grupo de saída da transformação de União e arraste-as para a transformação downstream no mapeamento.

## Transformação de União em um ambiente não nativo

O processamento da transformação de União em um ambiente não nativo depende do mecanismo que executa a transformação.

Considere o suporte para os seguintes mecanismos de tempo de execução não nativos:

- Mecanismo Blaze. Suportado sem restrições.
- Mecanismo Spark. Com suporte sem restrições em mapeamentos em lote. Com suporte com restrições em mapeamentos de fluxo.
- Mecanismo Databricks Spark. Suportado sem restrições.

## Transformação da União em um mapeamento de fluxo

Os mapeamentos de fluxo têm restrições de validação de mapeamento.

### Validação de Mapeamento

A validação do mapeamento falha na seguinte situação:

- A transformação é configurada para mesclar dados de pipelines de fluxo e não fluxo.

## Transformação de União no mecanismo Databricks Spark

## CAPÍTULO 45

# Transformação de Estratégia de Atualização

Este capítulo inclui os seguintes tópicos:

- [Transformação de Estratégia de Atualização - Visão Geral, 665](#)
- [Transformações de Estratégia de Atualização em Mapeamentos Dinâmicos, 666](#)
- [Sinalizando Linhas Dentro de um Mapeamento, 666](#)
- [Especificando Opções de Atualização para Destinos Individuais, 668](#)
- [Transformação de Estratégia de Atualização em um ambiente não nativo, 669](#)

## Transformação de Estratégia de Atualização - Visão Geral

A transformação de Estratégia de Atualização é uma transformação ativa que sinaliza uma linha para inserir, atualizar, excluir ou rejeitar. Use uma transformação de Estratégia de Atualização para controlar as alterações em linhas existentes de um destino com base em uma condição que você aplicar.

Como transformação ativa, a transformação de Estratégia de Atualização pode alterar o número de linhas que passam por ela. A transformação de Estratégia de Atualização testa cada linha para ver se ela atende a uma condição específica e depois sinaliza a linha de maneira compatível. A transformação passa linhas sinalizadas para inserção, atualização ou exclusão para a próxima transformação. Você pode configurar a transformação para passar linhas sinalizadas para rejeição para a próxima transformação ou para remover essas linhas.

Por exemplo, você pode usar a transformação de Estratégia de Atualização para sinalizar todas as linhas de clientes para atualização quando o endereço de correspondência tiver sido alterado. Ou, você pode sinalizar todas as linhas de funcionários para rejeição para pessoas que não trabalham mais para a organização.

Você pode usar uma transformação de Estratégia de Atualização para gravar resultados em um destino de banco de dados relacional quando o mapeamento for executado no Mecanismo Spark. O mapeamento usa uma cadeia de conexão JDBC.

## Definindo a Estratégia de Atualização

Para definir uma estratégia de atualização, conclua as etapas a seguir:

1. Para controlar a forma como as linhas são sinalizadas para inserir, atualizar, excluir ou rejeitar dentro de um mapeamento, adicione uma transformação Estratégia de Atualização ao mapeamento. Use uma transformação de Estratégia de Atualização para sinalizar as linhas que têm o mesmo destino para operações de banco de dados diferentes ou para rejeitar linhas.
2. Defina opções de inserção, atualização e exclusão para destinos individuais quando você configura o mapeamento. Em cada destino, você pode permitir ou não permitir inserções e exclusões para todas as linhas sinalizadas para inserção ou exclusão. Você pode escolher maneiras diferentes de lidar com as atualizações de todas as linhas sinalizadas para atualização.
3. Se o mapeamento for executado no mecanismo de tempo de execução do Spark, você poderá escolher Usar Mesclagem do Hive nas propriedades avançadas de transformação de Estratégia de Atualização. Quando uma consulta usa uma instrução MERGE em vez de instruções INSERT, UPDATE ou DELETE, o processamento geralmente é mais eficiente.

## Transformações de Estratégia de Atualização em Mapeamentos Dinâmicos

É possível usar uma transformação de Estratégia de Atualização em um mapeamento dinâmico. Você pode configurar portas dinâmicas na transformação e fazer referência às portas geradas.

Você pode fazer referência a uma porta dinâmica ou a uma porta gerada na transformação de Estratégia de Atualização. No entanto, se a porta gerada não existir em tempo de execução, o mapeamento falhará.

Se você usar uma porta dinâmica na expressão de estratégia de atualização, essa porta dinâmica não poderá ter mais de uma porta gerada.

É possível parametrizar a expressão de estratégia de atualização. Use um parâmetro de tipo de expressão e insira a expressão inteira como o valor do parâmetro padrão.

## Sinalizando Linhas Dentro de um Mapeamento

Adicione uma transformação de Estratégia de Atualização a um mapeamento para sinalizar linhas individuais para inserção, atualização, exclusão ou rejeição.

Defina uma expressão de estratégia de atualização para testar cada linha e ver se ela atende a uma condição específica. Em seguida, atribua a cada linha um código numérico para sinalizar as linhas para uma determinada operação de banco de dados.

A seguinte tabela lista as constantes para cada operação de banco de dados e seu equivalente numérico:

Operação	Constante	Valor Numérico
Inserir	DD_INSERT	0
Atualizar	DD_UPDATE	1

Operação	Constante	Valor Numérico
Excluir	DD_DELETE	2
Rejeitar	DD_REJECT	3

O Data Integration Service trata qualquer outro valor como uma inserção.

## Expressões de Estratégia de Atualização

Insira uma expressão de estratégia de atualização no Editor de Expressão.

A expressão de estratégia de atualização usa a função IIF ou DECODE da linguagem de transformação para testar cada linha. Por exemplo, a instrução IIF a seguir sinaliza uma linha para rejeição se a data de entrada é posterior à data de aplicação. Caso contrário, a instrução sinaliza a linha para atualização:

```
IIF( ( ENTRY_DATE > APPLY_DATE), DD_REJECT, DD_UPDATE)
```

É possível configurar parâmetros em uma expressão de estratégia de atualização. Crie ou procure parâmetros no Editor de Expressão.

Se a transformação for em um mapeamento dinâmico, os campos nela gerados poderão mudar. É possível parametrizar a expressão de estratégia de atualização completa. Se você usar um parâmetro para definir a expressão, a Developer tool não poderá validá-la. Um parâmetro de expressão não pode conter outro parâmetro.

## Transformação de Estratégia de Atualização - Propriedades Avançadas

Configure propriedades avançadas para ajudar a determinar como o Serviço de Integração de Dados processa dados para a transformação de Estratégia de Atualização.

Você pode definir as seguintes propriedades avançadas para a transformação de Estratégia de atualização na guia Avançado:

### Encaminhar Linhas Rejeitadas

Determina se a transformação de Estratégia de Atualização passa linhas rejeitadas para a próxima transformação ou se as descarta. Por padrão, o Serviço de Integração de Dados encaminha as linhas rejeitadas para a próxima transformação. O Serviço de Integração de Dados sinaliza as linhas para rejeição e as grava no arquivo de rejeição. Se você não selecionar Encaminhar Linhas Rejeitadas, o Serviço de Integração de Dados descartará as linhas rejeitadas e as gravará no arquivo de log do mapeamento.

### Usar mesclagem do Hive

Determina se a transformação de Estratégia de Atualização usa o MERGE do Hive para realizar atualizações nos destinos do Hive quando o mapeamento é executado no Spark. Quando uma consulta usa uma instrução MERGE em vez de instruções INSERT, UPDATE ou DELETE, o processamento é mais eficiente.

O mapeamento ignora a opção MERGE do Hive e o Serviço de Integração de Dados usa INSERT, UPDATE e DELETE para executar a operação nos seguintes cenários:

- O mapeamento é executado no Blaze ou Hive.

- Em cenários em que o MERGE é restrito pela implementação do Hive em distribuições específicas do Hadoop.

O log de mapeamento contém resultados da operação, incluindo se as restrições afetaram os resultados.

Quando a atualização afeta colunas de particionamento ou de intervalo, as atualizações nas colunas são omitidas.

**Nota:** A Developer tool e o Serviço de Integração de Dados não fazem a validação contra essa restrição. Se a expressão Estratégia de Atualização violar essas restrições, o mapeamento poderá produzir resultados inesperados.

#### Nível de Rastreamento

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.

## Transformações de Agregador e de Estratégia de Atualização

Ao conectar as transformações de Agregador e Estratégia de Atualização como parte do mesmo pipeline, coloque o Agregador antes da transformação de Estratégia de Atualização. Nessa ordem, o Data Integration Service realiza o cálculo de agregação e, em seguida, sinaliza as linhas que contêm os resultados desse cálculo para inserção, atualização, exclusão ou rejeição.

Se colocar a Estratégia de Atualização antes da transformação de Agregador, você deverá considerar como a transformação de Agregador lida com as linhas sinalizadas para diferentes operações. Nessa ordem, o Data Integration Service sinaliza linhas para inserção, atualização, exclusão ou rejeição antes de executar o cálculo de agregação. A maneira como você sinaliza uma linha determina como a transformação de Agregador trata os valores dessa linha usados no cálculo. Por exemplo, se você sinalizar uma linha para exclusão e depois usá-la para calcular a soma, o Data Integration Service subtrairá o valor dessa linha. Se você sinalizar uma linha para rejeição e depois usa a linha para calcular a soma, o Data Integration Service não incluirá o valor nessa linha. Se você sinalizar uma linha para inserção ou atualização e depois usar a linha para calcular a soma, o Data Integration Service adicionará o valor dessa linha à soma.

## Especificando Opções de Atualização para Destinos Individuais

Após usar uma transformação de Estratégia de Atualização para sinalizar cada linha para uma determinada operação de banco de dados, defina opções de inserção, atualização e exclusão para cada destino no mapeamento. Você pode não permitir inserções ou exclusões em linhas sinalizadas para inserção ou exclusão. Você pode escolher maneiras diferentes de lidar com as atualizações de todas as linhas sinalizadas para atualização.

Defina as opções da estratégia de atualização nas propriedades Avançado de um objeto de dados de destino em um mapeamento. Você pode definir as seguintes opções de estratégia de atualização:

#### Inserir

Inserir todas as linhas sinalizadas para inserção em um destino. Por padrão, esta opção fica ativada.

#### Excluir

Excluir todas as linhas sinalizadas para exclusão de um destino. Por padrão, esta opção fica ativada.



### Atualizar Estratégia

Estratégia de atualização para as linhas existentes. Selecione uma das seguintes estratégias:

- **Atualizar como Atualização.** Atualiza todas as linhas sinalizadas para atualização. Esse é o valor padrão.
- **Atualizar como Inserção.** Insere todas as linhas sinalizadas para atualização.
- **Atualizar e Inserir.** Atualiza todas as linhas sinalizadas para atualização caso existam no destino e depois insere as linhas restantes marcadas para inserção.

### Truncar Tabela

Trunca o destino antes de carregar dados. Por padrão, essa opção fica desativada.

## Transformação de Estratégia de Atualização em um ambiente não nativo

O processamento da transformação de Estratégia de Atualização em um ambiente não nativo depende do mecanismo que executa a transformação.

Considere o suporte para os seguintes mecanismos de tempo de execução não nativos:

- Mecanismo Blaze. Suportado com restrições.
- Mecanismo Spark. Com suporte com restrições em mapeamentos em lote. Sem suporte em mapeamentos de fluxo.
- Mecanismo Databricks Spark. Sem suporte.

**Nota:** A transformação de Estratégia de Atualização é suportada apenas nas distribuições do Hadoop que suportam o Hive ACID.

## Transformação de Estratégia de Atualização no mecanismo Blaze

Você pode usar a transformação de Estratégia de Atualização nas distribuições do Hadoop que suportam o Hive ACID.

Algumas regras de processamento para o mecanismo Blaze são diferentes das regras de processamento do Serviço de Integração de Dados.

### Restrições gerais

Se a transformação de Estratégia de Atualização receber várias linhas de atualização para o mesmo valor de chave primária, a transformação selecionará uma linha aleatória para atualizar o destino.

Se várias transformações da Estratégia de Atualização forem gravadas em instâncias diferentes do mesmo destino, os dados de destino poderão ser imprevisíveis.

O mecanismo Blaze executa as operações na seguinte ordem: exclusões, atualizações e inserções. Ele não processa linhas na mesma ordem em que a transformação de Estratégia de Atualização as recebe.

Os destinos do Hive sempre executam as operações Atualizar como Atualização. Os destinos do Hive não suportam Atualizar e Inserir ou Atualizar como Inserção.

## Validação de mapeamento e validação de compilação

A validação do mapeamento falha nas seguintes situações:

- A transformação de Estratégia de Atualização está conectada a mais de um destino.
- A transformação de Estratégia de Atualização não está localizada imediatamente antes do destino.
- O destino da Estratégia de Atualização não é um destino do Hive.
- O destino da transformação de Estratégia de Atualização é uma tabela ACID externa.
- O destino não contém uma chave primária.
- A propriedade de destino do Hive para truncar a tabela de destino no tempo de execução é ativada.
- A propriedade de destino do Hive para criar ou substituir a tabela de destino no tempo de execução é ativada.

O mapeamento falha na seguinte situação:

- O destino não é ORC com intervalos.
- O destino do Hive é modificado para ter menos linhas que a tabela real.

Ocorrem erros de validação de compilação e a execução do mapeamento é interrompida nas seguintes situações:

- A versão do Hive é anterior à 0.14.
- A tabela de destino não está ativada para transações.

## Usando tabelas de destino do Hive

Para usar uma tabela de destino do Hive com uma transformação de Estratégia de Atualização, você deve criar a tabela de destino do Hive com a seguinte cláusula na Linguagem de Definição de Dados do Hive:

```
TBLPROPERTIES ("transactional"="true").
```

Para usar uma transformação de Estratégia de Atualização com um destino do Hive, verifique se as seguintes propriedades estão configuradas no conjunto de configurações hive-site.xml associado à conexão do Hadoop:

```
hive.support.concurrency      true
hive.enforce.bucketing       true
hive.exec.dynamic.partition.mode nonstrict
hive.txn.manager               org.apache.hadoop.hive.ql.lockmgr.DbTxnManager
hive.compactor.initiator.on    true
hive.compactor.worker.threads 1
```

## Transformação de Estratégia de Atualização no mecanismo Spark

Você pode usar a transformação de Estratégia de Atualização nas distribuições do Hadoop que suportam o Hive ACID.

Você também pode usar a transformação de Estratégia de Atualização para gravar os resultados do mapeamento em destinos relacionais compatíveis com JDBC.

Algumas regras de processamento para o mecanismo Spark são diferentes das regras de processamento do Serviço de Integração de Dados.

### Restrições gerais para destinos Hive

A transformação de Estratégia de Atualização não encaminha linhas rejeitadas para a próxima transformação quando o destino é uma tabela Hive ou uma tabela compatível com JDBC.

Se a transformação de Estratégia de Atualização receber várias linhas de atualização para o mesmo valor de chave primária, a transformação selecionará uma linha aleatória para atualizar o destino.

Se várias transformações da Estratégia de Atualização forem gravadas em instâncias diferentes do mesmo destino, os dados de destino poderão ser imprevisíveis.

Se o mapeamento for executado no mecanismo Spark, você poderá escolher a opção Usar Mesclagem do Hive. A opção tem as seguintes restrições:

- Uma única linha para exclusão ou atualização não pode corresponder a várias linhas no destino. Quando o mapeamento viola essa restrição, ele falha com um erro de tempo de execução.
- Se você configurar a expressão Estratégia de Atualização para atualizar colunas de particionamento ou de intervalo, o mapeamento ignorará a opção MERGE do Hive e não atualizará as colunas.

**Nota:** A Developer tool e o Serviço de Integração de Dados não fazem a validação em relação a essas restrições. Se a expressão ou o mapeamento violar essas restrições, o mapeamento poderá ser executado, mas os resultados não serão os esperados.

Os destinos do Hive sempre executam as operações Atualizar como Atualização. Os destinos do Hive não suportam Atualizar e Inserir ou Atualizar como Inserção.

## Validação de Mapeamento

A validação do mapeamento falha nas seguintes situações.

- A transformação de Estratégia de Atualização está conectada a mais de um destino.
- A transformação de Estratégia de Atualização não está localizada imediatamente antes do destino.
- O destino da transformação de Estratégia de Atualização é uma tabela ACID externa.
- O destino não contém uma chave primária conectada.
- A propriedade para ativar o truncamento do Hive ou da tabela de destino relacional no tempo de execução é selecionada.
- Uma das seguintes estratégias de destino para o Hive ou a tabela de destino relacional no tempo de execução é selecionada:
  - Criar ou substituir a tabela de destino
  - ApplyNewColumns
  - ApplyNewSchema
  - Reprovação

O mapeamento falha nas seguintes situações quando o destino é um destino Hive:

- A tabela de destino não está ativada para transações.
- O destino não é ORC com intervalos.

## Usando tabelas de destino do Hive

Para usar uma tabela de destino do Hive com uma transformação de Estratégia de Atualização, você deve criar a tabela de destino do Hive com a seguinte cláusula na Linguagem de Definição de Dados do Hive:

```
TBLPROPERTIES ("transactional"="true").
```

Para usar uma transformação de Estratégia de Atualização com um destino do Hive, verifique se as seguintes propriedades estão configuradas no conjunto de configurações hive-site.xml associado à conexão do Hadoop:

```
hive.support.concurrency      true
hive.enforce.bucketing       true
hive.exec.dynamic.partition.mode nonstrict
hive.txn.manager              org.apache.hadoop.hive.ql.lockmgr.DbTxnManager
hive.compactor.initiator.on    true
hive.compactor.worker.threads 1
```

## CAPÍTULO 46

# Transformação de Consumidor de Serviço da Web

Este capítulo inclui os seguintes tópicos:

- [Visão Geral da Transformação de Consumidor de Serviço da Web, 672](#)
- [Seleção WSDL, 675](#)
- [Portas de Transformação de Consumidor de Serviço da Web, 676](#)
- [Mapeamento de Entrada da Transformação de Consumidor de Serviço da Web, 678](#)
- [Transformação de Consumidor de Serviço da Web - Mapeamento de Saída, 681](#)
- [Propriedades Avançadas da Transformação de Consumidor de Serviço da Web, 684](#)
- [Otimização de Filtro, 689](#)
- [Criando uma Transformação de Consumidor de Serviço da Web, 691](#)
- [Exemplo de Transformação de Consumidor de Serviço da Web, 693](#)

## Visão Geral da Transformação de Consumidor de Serviço da Web

A transformação de Consumidor de Serviço da Web se conecta a um serviço da Web como um cliente de serviços da Web para acessar ou transformar dados. A transformação de Consumidor de Serviço da Web é uma transformação de vários grupos.

Um serviço da Web usa padrões abertos, como SOAP, WSDL e XML. SOAP é o protocolo de comunicações para serviços da Web. A solicitação do cliente de serviços da Web e a resposta do serviço da Web são mensagens SOAP. Um WSDL é um esquema XML que descreve os protocolos, os formatos e as assinaturas das operações de serviços da Web.

Operações de serviços da Web incluem solicitações de informações, solicitações para atualizar dados ou solicitações para realizar tarefas. Por exemplo, a transformação de Consumidor de Serviço da Web envia uma solicitação SOAP para executar uma operação de serviço da Web denominada `getCustomerOrders`. A transformação transmite um ID de cliente na solicitação. O serviço da Web recupera as informações de pedidos e clientes. O serviço da Web retorna essas informações para a transformação em uma resposta SOAP.

A transformação de Consumidor de Serviço da Web se conecta a um serviço da Web usando um URL de ponto de extremidade definido no WSDL, em uma conexão de serviços da Web ou em uma porta de entrada

de URL de ponto de extremidade. Você habilita a segurança para serviços da Web em uma conexão de serviços da Web.

## Mensagens SOAP

A transformação de Consumidor de Serviço da Web usa o protocolo SOAP (Simple Object Access Protocol) para trocar informações com o provedor de serviços da Web e para solicitar serviços da Web. O SOAP define o formato para as mensagens de resposta e solicitação de serviço da Web.

Quando você transforma dados com uma transformação de Consumidor de Serviço da Web, a transformação gera uma solicitação SOAP e se conecta ao serviço da Web. A transformação se conecta ao serviço da Web usando um URL de ponto de extremidade definido no objeto WSDL, em uma conexão de serviços da Web ou em uma porta de entrada de URL de ponto de extremidade. A solicitação SOAP contém as informações que o serviço da Web precisa usar para executar a operação solicitada. A operação de serviço da Web retorna dados para a transformação em uma resposta SOAP. A transformação mapeia dados da resposta SOAP e retorna os dados em portas de saída.

A transformação de Consumidor de Serviço da Web codifica cabeçalhos de mensagens SOAP em ISO-8859-1.

A transformação pode processar mensagens SOAP com codificação de documento/literal. O estilo documento/literal requer um esquema XML para descrever a mensagem SOAP. Mensagens SOAP são formadas a partir do XML. Quando uma mensagem SOAP contém elementos de ocorrência múltipla, os grupos de elementos formam níveis na hierarquia XML. Os grupos são relacionados quando um nível está aninhado dentro de outro.

Uma mensagem de solicitação SOAP pode conter dados hierárquicos. Por exemplo, a transformação de Consumidor de Serviço da Web envia uma solicitação para adicionar pedidos de clientes a um banco de dados de vendas. A transformação transmite dois grupos de dados em uma mensagem de solicitação SOAP. Um grupo contém um ID de cliente e um nome, enquanto o outro grupo contém informações de pedidos. As informações de pedidos ocorrem várias vezes.

Uma mensagem de resposta SOAP pode conter dados hierárquicos. Por exemplo, a transformação de Consumidor de Serviço da Web gera uma solicitação SOAP para pedidos de clientes. O serviço da Web retorna um cabeçalho de pedido e elementos de detalhes de pedido de ocorrência múltipla na resposta SOAP.

## Arquivos WSDL

Um arquivo WSDL contém uma descrição dos dados a serem transmitidos ao serviço da Web para que o remetente e o destinatário compreendam os dados que devem ser trocados. É necessário importar um arquivo WSDL para o repositório antes de poder criar uma transformação de Consumidor de Serviço da Web.

O WSDL descreve as operações que devem ser realizadas nos dados e uma associação a um protocolo ou transporte, para que o consumidor de serviço da Web possa enviar a mensagem de solicitação no formato correto. O WSDL descreve o endereço de rede para conexão com o serviço da Web.

O WSDL inclui informações sobre como codificar mensagens de solicitação e resposta SOAP. A codificação SOAP determina o formato do corpo da mensagem SOAP. Ela descreve o formato para mensagens de solicitação e resposta que o serviço da Web usa para se comunicar com o consumidor de serviço da Web. Os desenvolvedores de serviços da Web podem usar uma variedade de kits de ferramentas para criar serviços da Web. Esses kits de ferramentas oferecem suporte a diferentes maneiras de codificação de mensagens SOAP.

A transformação de Consumidor de Serviço da Web oferece suporte ao estilo de codificação SOAP de documentos/literal. É possível usar o WSDL 1.1 com a transformação de Consumidor de Serviço da Web. Não é possível usar anexos WSDL, como mensagens MIME, DIME e MTOM.

## Operações

Um serviço da Web contém uma operação para cada ação à qual ele oferece suporte.

Por exemplo, um serviço da Web pode ter uma operação denominada `getcustomerid` que recebe um nome de cliente e responde com os detalhes do cliente. A entrada da operação inclui um elemento para o nome do cliente. A saída da operação inclui elementos para detalhes do cliente com base no nome do cliente.

Ao configurar uma transformação de Consumidor de Serviço da Web, você define como a transformação mapeia dados para a entrada da operação e como ela mapeia dados a partir da saída da operação. As seguintes informações são configuradas na transformação:

### Mapeamento de entrada

Defina como mapear as portas de entrada da transformação para os nós de entrada de operação do serviço da Web. A entrada de operação define os elementos na solicitação SOAP para a operação.

### Mapeamento de saída

Defina como mapear os nós de saída da operação do serviço da Web para as portas de saída da transformação. A saída de operação define os elementos em uma resposta SOAP para a operação.

## Segurança de Serviços da Web

Você habilita a segurança para serviços da Web em uma conexão de serviços da Web. Você pode configurar os seguintes tipos de segurança:

### Segurança de Serviços da Web

O Serviço de Integração de Dados pode incluir um cabeçalho de segurança de serviço da Web ao enviar uma solicitação SOAP para o provedor de serviços da Web. O cabeçalho de segurança de serviço da Web contém informações de autenticação, para que o provedor de serviços da Web possa autenticar o Serviço de Integração de Dados.

A transformação de Consumidor de Serviço da Web fornece o token de nome de usuário. O Serviço de Integração de Dados cria um cabeçalho SOAP de segurança separado na solicitação SOAP e transmite essa solicitação ao provedor de serviços da Web.

Você pode usar os seguintes tipos de segurança de serviços da Web em uma conexão de serviços da Web:

- **PasswordText.** O Serviço de Integração de Dados não altera a senha no cabeçalho SOAP WS-Security.
- **PasswordDigest.** O Serviço de Integração de Dados combina a senha com um valor de uso único e um carimbo de data/hora. O Serviço de Integração de Dados aplica um SHA hash da senha, codifica em codificação Base64 e usa a senha codificada no cabeçalho SOAP.

### Segurança da camada de transporte

A segurança implementada na camada de transporte (camada TCP) do TCP/IP via SSL (Secure Sockets Layer). Serviços da Web usam o protocolo HTTPS (Hypertext Transfer Protocol over SSL) como um endereço da Web para o transporte seguro de mensagens. As transformações de Consumidor de Serviço da Web podem usar o TLS 1.2, TLS 1.1 ou TLS 1.0. Você pode usar a seguinte autenticação com segurança de camada de transporte: autenticação HTTP, autenticação de servidor proxy e certificados SSL.

### Autenticação SSL

Você pode usar a autenticação SSL ao se conectar usando o protocolo HTTPS.

Os seguintes tipos de autenticação SSL podem ser usados:

- Autenticação SSL unidirecional
- Autenticação SSL bidirecional

#### **Autenticação HTTP**

Você pode usar a autenticação HTTP ao se conectar usando o protocolo HTTP.

Os seguintes tipos de autenticação HTTP podem ser usados:

- Autenticação básica
- Autenticação Digest
- Autenticação NTLM (NT LAN Manager)

## Seleção WSDL

Antes de criar uma transformação de Consumidor de Serviço da Web, você deve importar um arquivo WSDL para o repositório do modelo. Você pode importar apenas arquivos WSDL definidos com um estilo de vinculação SOAP de documento/literal.

O WSDL define a assinatura de operação do serviço da Web que você deseja executar. Quando um WSDL é importado, a ferramenta Developer cria um objeto de dados físicos que pode ser reutilizado para outras transformações.

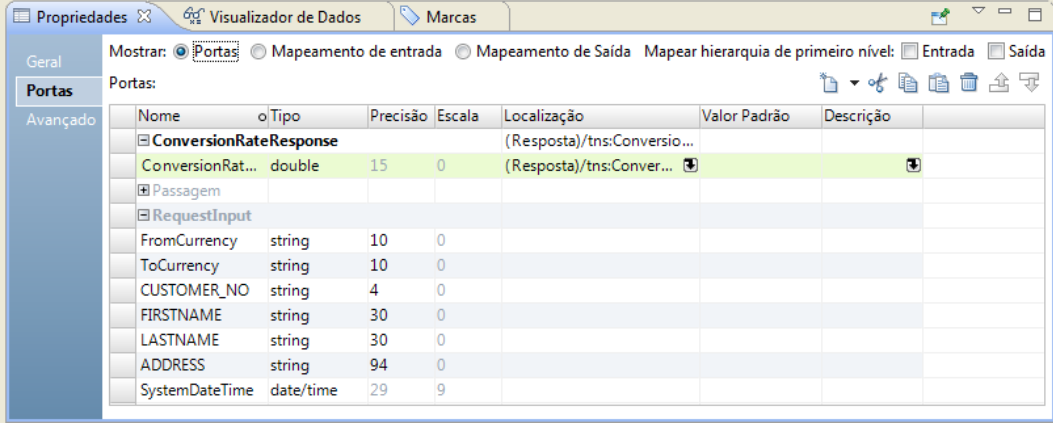
Um WSDL pode definir várias operações. Ao criar uma transformação de Consumidor de Serviço da Web, selecione qual operação você deseja executar. Você pode exibir as hierarquias de entrada e de saída de operação na transformação de Consumidor de Serviço da Web. Essas hierarquias definem a estrutura da mensagem de solicitação SOAP e da mensagem de resposta SOAP.

Você também pode importar um WSDL com a operação de entrada unidirecional. Você deve criar portas de saída fictícias ao importar um WSDL com uma operação de entrada unidirecional.

# Portas de Transformação de Consumidor de Serviço da Web

Ao exibir as portas de transformação, mostre as portas se não precisar exibir a hierarquia de operação. Quando você mostra as portas, pode definir grupos e portas, além de mapear nós da saída da operação para as portas de saída.

A seguinte figura mostra as portas para uma transformação de Consumidor de Serviço da Web não reutilizável:



Nome	o Tipo	Precisão	Escala	Localização	Valor Padrão	Descrição
ConversionRateResponse				(Resposta)/tns:Conversio...		
ConversionRat...	double	15	0	(Resposta)/tns:Conver...		
Passagem						
RequestInput						
FromCurrency	string	10	0			
ToCurrency	string	10	0			
CUSTOMER_NO	string	4	0			
FIRSTNAME	string	30	0			
LASTNAME	string	30	0			
ADDRESS	string	94	0			
SystemDateTime	date/time	29	9			

Uma transformação de Consumidor de Serviço da Web pode ter vários grupos de entrada e de saída. Ao criar portas, crie grupos e adicione as portas a esses grupos. Defina as portas em uma hierarquia de grupos com base na estrutura da hierarquia de entrada ou saída de operação. Adicione uma chave para relacionar um grupo filho a um grupo pai. Todos os grupos, exceto o grupo mais baixo na hierarquia, devem ter chaves primárias. Todos os grupos na hierarquia, exceto o grupo raiz, devem ter chaves externas.

A transformação tem um grupo de entrada raiz denominado RequestInput. Você deve adicionar uma chave primária ao grupo de entrada raiz. A chave deve ser string, bigint ou integer.

Você pode adicionar mais portas de passagem ao grupo de entrada raiz. Portas de passagem transmitem dados pela transformação sem modificá-los. A porta de passagem pode ocorrer uma vez nos dados de entrada. Você pode adicionar a porta de passagem a qualquer grupo de saída. Associe a porta de saída à porta de entrada. O valor de entrada que você transmite por meio de uma solicitação SOAP se repete nas linhas de saída da resposta SOAP.

Você também pode adicionar cabeçalhos HTTP, portas para cookies, uma porta de URL dinâmica e portas para autenticação de segurança de serviços da Web ao grupo de entrada raiz. Os dados no grupo raiz ocorrem uma única vez.

Para mapear um nó de saída de operação para uma porta de saída, clique no campo da coluna **Localização** e expanda a hierarquia na caixa de diálogo **Selecionar Localização**. Em seguida, escolha um nó da hierarquia.

## Portas de Entrada de Cabeçalho HTTP

Um serviço da Web pode exigir cabeçalhos HTTP adicionais. É possível criar portas de entrada no grupo de entrada raiz para transmitir as informações de cabeçalho adicionais ao provedor de serviços da Web.

Para adicionar um cabeçalho HTTP e uma porta HTTP, selecione o grupo de entrada raiz e clique na seta ao lado do botão **Novo**. Em seguida, clique em **Cabeçalho HTTP**. Insira um nome de cabeçalho e um nome de porta.



É possível criar vários cabeçalhos HTTP.

## Outras Portas de Entrada

É possível adicionar portas de entrada predefinidas à transformação de Consumidor de Serviço da Web.

As seguintes portas de entrada predefinidas podem ser adicionadas:

### Porta para cookies

É possível configurar a transformação de Consumidor de Serviço da Web para usar a autenticação por cookies. O servidor remoto de serviços da Web rastreia os usuários consumidores de serviços da Web com base nos cookies. Você pode melhorar o desempenho quando um mapeamento chama um serviço da Web várias vezes.

Quando a porta para cookies é projetada para uma mensagem de solicitação de serviço da Web, o provedor do serviço da Web retorna um valor de cookie na mensagem de resposta. Você pode transmitir o valor do cookie para outra transformação mais abaixo no mapeamento ou pode salvar o valor do cookie em um arquivo. Ao salvar o valor do cookie em um arquivo, você pode configurar esse cookie como entrada para a transformação de Consumidor de Serviço da Web.

É possível projetar a porta de saída de cookies para qualquer um dos grupos de saída de transformações de Consumidor de Serviço da Web.

### Porta de URL do ponto de extremidade

A transformação de Consumidor de Serviço da Web se conecta a um serviço da Web usando um URL de ponto de extremidade. Você pode definir o URL de ponto de extremidade no arquivo WSDL, em uma conexão de serviços da Web ou em uma porta de entrada de URL de ponto de extremidade. Quando a transformação recebe o URL dinamicamente em uma porta, o Data Integration Service substitui o URL definido no arquivo WSDL ou na conexão de serviços da Web.

A transformação de Consumidor de Serviço da Web pode ter um valor de porta de URL para cada solicitação de serviço da Web. Adicione uma porta de URL de ponto de extremidade ao grupo de entrada raiz.

### Portas WS-Security

Você habilita a segurança de serviços da Web em uma conexão de serviço da Web. Ao habilitar a segurança de serviços da Web, você deve definir o nome de usuário e a senha em uma conexão de serviços da Web ou em portas de entrada WS-Security.

Ao adicionar portas WS-Security, você transmite o nome de usuário e a senha por meio de portas de entrada na transformação. Quando a transformação recebe o nome de usuário e a senha dinamicamente em portas, o Data Integration Service substitui os valores definidos na conexão de serviços da Web.

**Observação:** Uma conexão de serviços da Web tem um nome de usuário e uma senha para autenticação HTTP e WS-Security.

Para adicionar portas de entrada predefinidas, clique no grupo de entrada raiz na área **Portas**. Clique na seta ao lado do botão **Novo** e depois clique em **Outras Portas**. Escolha as portas a serem adicionadas.

# Mapeamento de Entrada da Transformação de Consumidor de Serviço da Web

Ao exibir as portas de transformação, mostre o mapeamento de entrada para exibir a hierarquia de entrada da operação. Ao mostrar o mapeamento de entrada, você pode definir grupos e portas de entrada, e mapear portas de entrada para nós de entrada da operação.

A entrada de mapeamento inclui as seguintes áreas:

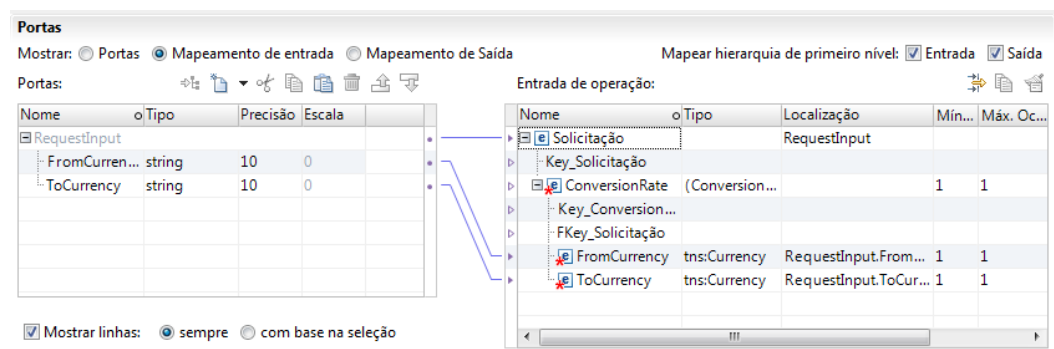
## Portas

Crie os grupos e as portas de entrada da transformação na área **Portas**.

## Entrada da Operação

A área **Entrada da Operação** mostra os nós na mensagem de solicitação SOAP que a transformação de Consumidor de Serviço da Web envia para o serviço da Web. O objeto de dados WSDL usado para criar a transformação define a hierarquia de entrada da operação.

A figura a seguir mostra o mapeamento de entrada para uma transformação de Consumidor de Serviço da Web não reutilizável:



Depois de criar portas de entrada, mapeie-as da área **Portas** para os nós na área **Entrada de Operação**. Quando você mapeia uma porta de entrada para um nó na entrada de operação, a localização da porta aparece na coluna **Localização** da área **Entrada de Operação**.

A ferramenta Developer mapeia nós no primeiro nível da entrada da operação para portas de entrada quando você opta por mapear o primeiro nível da hierarquia de entrada. A ferramenta Developer também cria as portas para realizar o mapeamento. Se o primeiro nível da hierarquia contém um nó pai de ocorrência múltipla com um ou mais nós filho de ocorrência múltipla, a ferramenta Developer não mapeará o primeiro nível da hierarquia.

Você pode mapear dados XML de uma string ou de uma porta de entrada de texto para toda a mensagem de solicitação SOAP. Quando você mapeia dados XML para toda a solicitação SOAP, não é possível mapear portas para nós na entrada de operação.

É possível optar por exibir as linhas que conectam as portas de entrada aos nós na entrada de operação.

## TÓPICOS RELACIONADOS:

- [“Gerando Visão Geral de Mensagens SOAP de Serviços da Web” na página 706](#)

## Regras e Diretrizes para Mapear Portas de Entrada para Nós

Analise as regras a seguir ao mapear portas de entrada para nós na hierarquia de entrada de operação:


- É possível mapear uma porta de entrada para um nó na hierarquia. Você pode mapear a mesma porta para qualquer número de chaves na hierarquia.
- A porta de entrada e o nó devem ter tipos de dados compatíveis.
- É possível mapear portas de um grupo de entrada para vários níveis de hierarquia na entrada de operação.
- Portas de entrada devem ser mapeadas para as chaves na entrada de operação. Qualquer porta mapeada para uma chave deve ter um tipo de dados de string, inteiro ou bigint. Mapeie dados para as chaves em todos os níveis da entrada de operação acima do nível de hierarquia que você está incluindo na mensagem SOAP. Inclua as chaves externas para todos os níveis acima do nível que está sendo mapeado (inclusive).

**Nota:** Você não precisará mapear portas de entrada para chaves se estiver mapeando somente o nível mais inferior da hierarquia de entrada de operação.

- É possível mapear várias portas de entrada do tipo string, bigint ou inteiro para uma chave na área **Entrada de Operação** para criar uma chave composta. Ao clicar no campo **Localização** para uma chave composta, você pode reordenar as portas de entrada ou remover uma das portas.

## Personalizar Opções de Exibição

Você pode alterar a hierarquia de entrada da operação para mostrar as chaves na área **Entrada da Operação**. Você também pode mostrar construções de agrupamento que definem como ordenar os nós.

Clique no botão **Personalizar Exibição** (  ) na área **Entrada de Operação**. Ative qualquer uma das seguintes opções:

### Sequência, Escolha e Tudo

Mostre uma linha que indica se uma definição de elemento é sequência, escolha ou tudo.

Os nós em um grupo "tudo" devem ser todos incluídos na mensagem SOAP.

Os nós em um grupo de sequência devem estar na ordem especificada no WSDL.

Pelo menos um nó em um grupo de escolha deve aparecer na mensagem SOAP.

### Chaves

Exiba as chaves na área **Entrada da Operação**. A área **Entrada da Operação** inclui chaves para cada grupo. Você pode adicionar uma chave a uma porta de entrada na área **Portas**.

## Mapeando Portas de Entrada para a Entrada de Operação

Ao mostrar o mapeamento de entrada de transformação, você pode definir grupos de entrada, definir portas de entrada e mapear portas de entrada para nós de entrada de operação.

1. Abra uma transformação de Consumidor de Serviço da Web.
2. Para exibir o mapeamento de entrada de transformação, use um dos métodos a seguir:

- Para uma transformação reutilizável, clique na exibição **Visão Geral**. Opte por mostrar o mapeamento de entrada.
  - Para uma transformação não reutilizável, clique na guia **Portas** da exibição **Propriedades**. Opte por mostrar o mapeamento de entrada.
3. Defina uma chave primária para o grupo de entrada raiz.
  4. Para adicionar uma porta ou grupo de entrada à área **Portas**, use um dos métodos a seguir:

Opção	Descrição
Arrastar um nó	Arraste um nó de grupo ou de um nó filho na área <b>Entrada de Operação</b> até uma coluna vazia na área <b>Portas</b> . Se o nó for um nó de grupo, a ferramenta Developer adicionará um grupo sem portas.
Adicionar um grupo ou uma porta manualmente	Para adicionar um grupo, clique na seta ao lado do botão <b>Novo</b> e clique em <b>Grupo</b> . Para adicionar uma porta, clique na seta ao lado do botão <b>Novo</b> e clique em <b>Campo</b> .
Arrastar uma porta a partir de outra transformação	No Editor, arraste uma porta a partir de outra transformação até a transformação de Consumidor de Serviço da Web.
Copiar uma porta	Selecione portas de outra transformação e copie-as para a área <b>Portas</b> . Para copiar portas, você pode usar atalhos de teclado ou pode usar os botões Copiar e Colar na ferramenta Developer.
Selecionar <b>Mapear primeiro nível da hierarquia</b>	Selecione <b>Mapear primeiro nível da hierarquia</b> . A ferramenta Developer mapeia nós no primeiro nível da entrada de operação para grupos e portas de entrada. A ferramenta Developer também cria os grupos e as portas de entrada para realizar o mapeamento.

5. Se você criar uma porta manualmente ou copiar uma porta de outra transformação, clique na coluna **Localização** da área **Entrada de Operação** e escolha uma porta na lista.
6. Para mapear portas de entrada como uma chave composta, use um dos métodos a seguir:

Opção	Descrição
Arrastar portas de entrada	Selecione duas ou mais portas de entrada e arraste-as até uma chave na hierarquia de entrada de operação.
Selecionar portas de entrada na caixa de diálogo <b>Selecionar Localização</b>	Clique na coluna <b>Localização</b> de uma chave na hierarquia de entrada de operação e selecione as portas de entrada.

7. Para limpar as localizações dos nós, use um dos métodos a seguir:

Opção	Descrição
Clicar em <b>Limpar</b>	Selecione um ou mais nós na área <b>Entrada de Operação</b> e clique em <b>Limpar</b> .
Excluir as linhas que conectam portas a nós	Selecione uma ou mais linhas que conectam as portas de entrada aos nós na entrada de operação e pressione a tecla <b>Delete</b> .

8. Se o objeto de dados WSDL associado incluir elementos anyType, elementos any, atributos anyAttribute, elementos de tipo derivado ou grupos de substituição, escolha objetos na área **Entrada de Operação**. Na coluna **Tipo** para um nó, clique em **Escolher** e, em seguida, escolha um ou mais tipos, elementos ou atributos na lista.

- Para mapear dados XML de uma string ou de uma porta de entrada de texto para a solicitação SOAP completa, clique com o botão direito na porta e selecione **Mapear como XML**.

## Transformação de Consumidor de Serviço da Web - Mapeamento de Saída

Ao exibir as portas de transformação, mostre o mapeamento de saída para exibir a hierarquia de saída da operação. Ao mostrar o mapeamento de saída, você pode definir grupos de saída e portas de saída, além de mapear nós de saída da operação para portas de saída.

O mapeamento de saída inclui as seguintes áreas:

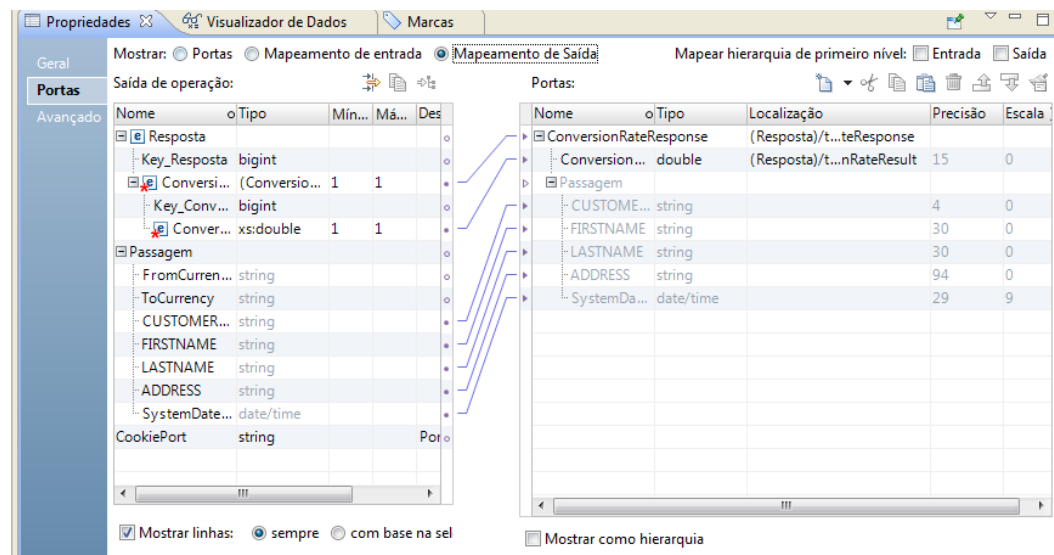
### Saída da Operação

A área **Saída da Operação** mostra os nós na mensagem de resposta SOAP que o serviço da Web retorna para a transformação de Consumidor de Serviço da Web. O objeto de dados WSDL usado para criar a transformação define a hierarquia de saída da operação.

### Portas

Crie os grupos e portas de saída da transformação na área **Portas**.

A figura a seguir mostra o mapeamento de saída para uma transformação de Consumidor de Serviço da Web não reutilizável:



Depois de criar portas de saída, mapeie os nós da área **Saída de Operação** para as portas na área **Portas**. Quando você mapeia um nó da saída de operação para uma porta de saída, a localização desse nó aparece na coluna **Localização** da área **Portas**.

A ferramenta Developer mapeia nós no primeiro nível da saída de operação para portas de saída quando você opta por mapear o primeiro nível da hierarquia de saída. A ferramenta Developer também cria as portas para realizar o mapeamento. Se o primeiro nível da hierarquia contém um nó pai de ocorrência múltipla com um ou mais nós filho de ocorrência múltipla, a ferramenta Developer não mapeará o primeiro nível da hierarquia.

Você pode optar por exibir as portas de saída em uma hierarquia. Cada grupo filho aparece abaixo do grupo pai. Você também pode optar por exibir as linhas que conectam os nós na saída de operação às portas de saída.

Se o objeto de dados WSDL associado for excluído do repositório, a ferramenta Developer manterá a localização dos nós de operação no mapeamento de saída. Quando você exibe o mapeamento de saída, a área **Portas** ainda mostra a localização dos nós de operação na coluna **Localização** para as portas de saída. Se você associar outro WSDL à transformação, a ferramenta Developer verificará se cada localização é válida. A ferramenta Developer limpará a localização dos nós de operação na área **Portas** do mapeamento de saída se essa localização não for mais válida.

## TÓPICOS RELACIONADOS:

- [“Analisando Mensagens SOAP de Serviços da Web - Visão Geral” na página 697](#)

## Regras e Diretrizes para Mapear Nós para Portas de Saída

Analise as regras a seguir ao mapear nós na hierarquia de saída de operação para portas de saída:

- O nó de saída de operação e a porta de saída devem ter tipos de dados compatíveis.
- Não é possível mapear um nó para mais de uma porta de saída em um grupo.
- Cada porta de saída deve ter uma localização válida, a menos que a porta seja uma porta de passagem.
- Se você arrastar um nó filho de ocorrência múltipla para uma porta de saída vazia, será necessário relacionar o grupo a outros grupos de saída. Quando você seleciona um grupo, a ferramenta Developer cria chaves para relacionar os grupos.
- Ao arrastar um elemento de ocorrência múltipla até um grupo que contém o elemento pai, você pode configurar o número de ocorrências do elemento filho para inclusão. Ou, você pode substituir o grupo pai pelo grupo filho de ocorrência múltipla na transformação de saída.

## Mapeando a Mensagem SOAP como XML

É possível mapear a mensagem SOAP completa como XML em vez de retornar os dados a portas de saída separadas.

Quando você mapeia a mensagem SOAP como XML, o Data Integration Service retorna a mensagem SOAP completa em uma única porta. Não crie portas de saída.

Para mapear a mensagem completa, clique com o botão direito no grupo raiz da área **Saída de Operação**. Selecione **Mapear como XML**.

A ferramenta Developer cria uma porta de saída de string. A precisão é de 65535 bytes.

## Personalizar Opções de Exibição

É possível alterar a hierarquia de saída de operação para mostrar portas para cookies, portas de passagem e chaves na área **Saída de Operação**. Você também pode mostrar construções de agrupamento que definem como ordenar nós.

Clique no botão **Personalizar Exibição** na área **Saída de Operação**. Habilite qualquer uma das opções a seguir:

### Sequência, Opção e Todos

Mostra uma linha que indica se uma definição de elemento é Sequência, Opção ou Todos.

Os nós em um grupo Todos devem todos ser incluídos na mensagem SOAP.

Os nós em um grupo de sequência deve estar na ordem especificada no WSDL.

Pelo menos um nó em um grupo de opção deve aparecer na mensagem SOAP.

#### Chaves

Exiba as chaves na área **Saída de Operação**. A área **Saída de Operação** inclui chaves para cada grupo. Você pode adicionar uma chave a uma porta de saída na área **Portas**.

#### Portas de Passagem

A área **Saída de Operação** mostra as portas de passagem. Portas de passagem são portas que transferem dados através da transformação sem alterar esses dados. Você pode projetar portas de passagem a partir da saída de operação para qualquer um dos grupos de saída da transformação de Consumidor de Serviço da Web. Uma porta de passagem recebe dados uma única vez e, portanto, essa porta está no nível raiz das mensagens SOAP.

#### Portas para Cookies

Mostra a porta para cookies. Quando você configura a autenticação por cookies, o servidor remoto de serviços da Web rastreia os usuários consumidores de serviços da Web com base nos cookies. Quando você projeta um cookie de serviço da Web na mensagem de solicitação, o serviço da Web retorna um cookie na mensagem de resposta. Você pode projetar o cookie a partir da saída de operação para qualquer um dos grupos de saída da transformação de Consumidor de Serviço da Web.

## Mapeando a Saída de Operação para Portas de Saída

Ao mostrar o mapeamento de saída de transformação, você pode definir grupos de saída, definir portas de saída e mapear nós de saída de operação para portas de saída.

1. Abra uma transformação de Consumidor de Serviço da Web.
2. Para exibir o mapeamento de saída de transformação, use um dos métodos a seguir:
  - Para uma transformação reutilizável, clique na exibição **Visão Geral**. Opte por mostrar o mapeamento de saída.
  - Para uma transformação não reutilizável, clique na guia **Portas** da exibição **Propriedades**. Opte por mostrar o mapeamento de saída.
3. Para adicionar um grupo de saída ou porta à área **Portas**, use um dos métodos a seguir:

Opção	Descrição
Arrastar um nó	Arraste um nó de grupo ou de um nó filho na área <b>Saída de Operação</b> até uma coluna vazia na área <b>Portas</b> . Se o nó for um nó de grupo, a ferramenta Developer adicionará um grupo sem portas.
Adicionar um grupo ou uma porta manualmente	Para adicionar um grupo, clique na seta ao lado do botão <b>Novo</b> e clique em <b>Grupo</b> . Para adicionar uma porta, clique na seta ao lado do botão <b>Novo</b> e clique em <b>Campo</b> .
Arrastar uma porta a partir de outra transformação	No Editor, arraste uma porta a partir de outra transformação até a transformação de Consumidor de Serviço da Web.
Copiar uma porta	Selecione portas de outra transformação e copie-as para a área <b>Portas</b> . Para copiar portas, você pode usar atalhos de teclado ou pode usar os botões Copiar e Colar na ferramenta Developer.

Opção	Descrição
Selecionar <b>Mapear primeiro nível da hierarquia</b>	Selecione <b>Mapear primeiro nível da hierarquia</b> . A ferramenta Developer mapeia nós no primeiro nível da saída de operação para grupos e portas de entrada. A ferramenta Developer também cria os grupos e as portas de saída para realizar o mapeamento.

- Se você criar uma porta manualmente ou copiar uma porta de outra transformação, clique na coluna **Localização** da área **Portas** e escolha um nó na lista.
- Para limpar as localizações das portas, use um dos métodos a seguir:

Opção	Descrição
Clicar em <b>Limpar</b>	Selecione uma ou mais portas na área <b>Portas</b> e clique em <b>Limpar</b> .
Excluir as linhas que conectam nós a portas	Selecione uma ou mais linhas que conectam os nós na saída de operação às portas de saída e pressione <b>Excluir</b> .

- Se o objeto de dados WSDL associado incluir elementos anyType, elementos any, atributos anyAttribute, elementos de tipo derivado ou grupos de substituição, escolha objetos na área **Saída de Operação**. Na coluna **Tipo** para um nó, clique em **Escolher** e, em seguida, escolha um ou mais tipos, elementos ou atributos na lista.
- Para mapear a mensagem de resposta SOAP completa como XML, clique com o botão direito no grupo raiz da área **Saída de Operação** e selecione **Mapear como XML**.

## Propriedades Avançadas da Transformação de Consumidor de Serviço da Web

As propriedades avançadas da transformação de Consumidor de Serviço da Web incluem o nível de rastreamento, as portas de falhas genéricas, a conexão de serviços da Web e as mensagens de solicitação de serviço da Web.

Você pode definir as seguintes propriedades avançadas para a transformação de Consumidor de Serviço da Web na guia **Avançado**:

### Nível de Rastreamento

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.

### Ação SOAP

Substitui o valor da ação SOAP definido no WSDL por um valor constante para a transformação de Consumidor de Serviço da Web.

### Habilitar Tratamento de Falhas SOAP Genéricas

Retorna mensagens de falha que não estão definidas no WSDL. Cria portas de saída em um grupo de saída GenericFault para lidar com mensagens e códigos de falha.



A seguinte tabela descreve as portas de saída de falha para SOAP 1.1 e SOAP 1.2:

Porta de Saída de Falha para SOAP 1.1	Porta de Saída de Falha para SOAP 1.2	Descrição
Código de Falha	Código*	Retorna um código de identificação de falha.
String de Falha	Motivo*	Retorna uma explicação do erro em uma mensagem de falha.
Detalhe da Falha	Detalhe	Retorna informações personalizadas que o provedor de serviços da Web passa para a transformação de Consumidor de Serviço da Web em uma mensagem de falha genérica.
Atuador da Falha	Função	Retorna informações sobre o objeto que causou a falha.
-	Nó	Retorna o URI do nó SOAP que gerou a falha.
* As portas de saída de Código e Motivo são hierárquicas.		

**Nota:** É possível expandir a porta de saída da falha de Código para extrair a porta de saída da falha de SubCode até um nível para cima.

#### Habilitar Tratamento de Erros HTTP

Retorna qualquer erro HTTP do serviço da Web. Cria uma porta de saída de erros HTTP no grupo de saída GenericFault.

#### Tratar Falha como Erro

Adiciona mensagens de falha ao log de mapeamento. Quando uma falha ocorre, o Serviço de Integração de Dados incrementa a contagem de erros para o mapeamento. Desabilite essa propriedade para permitir a otimização com seleção antecipada e a otimização de envio. Por padrão, esta opção fica ativada.

#### Conexão

Identifica o objeto de conexão de serviços da Web para se conectar ao serviço da Web. Crie a conexão de serviços da Web na ferramenta Developer. Edite a conexão de serviços da Web na ferramenta Developer ou na ferramenta Administrator. Quando você configurar uma conexão de serviços da Web, configure o URL do ponto de extremidade, o tipo de segurança necessário para o serviço da Web e um período de tempo limite de conexão.

A transformação de Consumidor de Serviço da Web se conecta a um serviço da Web usando um URL de ponto de extremidade. Você pode definir o URL de ponto de extremidade no arquivo WSDL, em uma conexão de serviços da Web ou em uma porta de entrada de URL de ponto de extremidade.

Use as diretrizes a seguir para determinar quando configurar uma conexão de serviços da Web:

- Configure uma conexão se quiser usar um URL de ponto de extremidade diferente do URL no arquivo WSDL e se não estiver usando uma porta de entrada de URL de ponto de extremidade.
- Configure uma conexão se o serviço da Web conectado exigir segurança de serviço da Web, autenticação HTTP ou certificados SSL.
- Configure uma conexão se quiser alterar o período de tempo limite de conexão padrão.

**Nota:** É possível associar um objeto de dados WSDL no repositório a uma conexão de serviços da Web. A conexão associada se tornará a conexão padrão para cada transformação de Consumidor de Serviço da Web que você criar a partir do WSDL.

#### Ativar Compactação

Permite a codificação de solicitações SOAP com o método de compactação GZIP e permite a decodificação de respostas SOAP com GZIP ou deflate.

#### Validação de Esquema XML

Valida a mensagem de resposta SOAP em tempo de execução. Selecione **Erro em XML Inválido** ou **Nenhuma Validação**.

#### Entrada Classificada

Permite que o Serviço de Integração de Dados gere a saída sem processar todos os dados de entrada. Habilite a entrada classificada quando os dados de entrada forem classificados pelas chaves na hierarquia de entrada de operação.

#### Otimização de Envio

Habilita a otimização de envio. Clique no botão **Abrir** na propriedade **Otimização de Envio** para selecionar portas de filtro que recebem valores de filtro. Para cada porta de filtro, escolha a porta de saída que contém a coluna filtrada na resposta do serviço da Web.

#### Tem Efeitos Colaterais

Caixa de seleção que indica que o serviço da Web realiza qualquer função além de retornar linhas. A transformação de Consumidor de Serviço da Web terá um efeito colateral se o serviço da Web, além de retornar linhas, modificar um objeto ou interagir com outros objetos ou funções. O serviço da Web pode modificar um banco de dados, adicionar a um total, gerar uma exceção, compor um e-mail ou chamar outros serviços da Web com efeitos colaterais. Desabilite a propriedade **Tem Efeitos Colaterais** para permitir a otimização com seleção antecipada ou a otimização de envio. Por padrão, esta opção fica ativada.

#### Habilitar Simultaneidade

Permite que a transformação de Consumidor de Serviço da Web crie diversas conexões simultâneas com um serviço da Web para que ele possa enviar várias solicitações de serviço da Web em paralelo. Ao habilitar a transformação de Consumidor de Serviço da Web para criar diversas conexões simultâneas com o serviço da Web, você pode definir o limite de consumo de memória total e o número de limites de conexões simultâneas.

A tabela a seguir descreve as opções:

Opções	Descrição
Habilitar simultaneidade	Cria várias conexões simultâneas com um serviço da Web.
Limite de Conexões Simultâneas	O número de conexões simultâneas de serviços da Web. O padrão é 20.
Limite de Memória de Concorrência Total (em MB)	O limite total de alocação de memória para todas as conexões simultâneas. O padrão é 100 MB.

## Tratamento de Erros de Serviços da Web

Você pode configurar a transformação de Consumidor de Serviço da Web para transmitir falhas SOAP e erros HTTP em direção descendente em um mapeamento. É possível incrementar a contagem de erros

quando uma falha ocorre. Configure o tratamento de erros de serviços da Web nas propriedades avançadas da transformação.

Um serviço da Web retorna uma mensagem de resposta ou uma falha. Uma falha é um erro. O serviço da Web pode gerar falhas diferentes com base nos erros que ocorrem.

A transformação de Consumidor de Serviço da Web pode retornar os seguintes tipos de falhas:

#### Falhas SOAP

Erros SOAP definidos pelo WSDL. Configure portas de erro de saída que retornam as falhas na mensagem de resposta do serviço da Web. Para uma associação SOAP 1.1, o Data Integration Service retorna a mensagem de falha, o código de falha, a cadeia de falha e elementos atuadores de falha para a falha. Para uma associação SOAP 1.2, o Data Integration Service retorna a mensagem de falha, o código, o motivo, o nó e os elementos de função para a falha.

#### Falhas SOAP Genéricas

O serviço da Web gera falhas SOAP genéricas em tempo de execução. Os elementos de falha são diferentes para uma associação SOAP 1.1 e uma associação SOAP 1.2. O WSDL não define falhas SOAP genéricas. Falhas SOAP genéricas incluem falhas de autenticação e erros de solicitação SOAP.

#### Erros HTTP

A ferramenta Developer adiciona a porta de saída de falha HTTP quando você habilita o tratamento de erros HTTP na transformação. O Data Integration Service retorna erros HTTP do serviço da Web em uma única porta de string. Um erro HTTP inclui um código de erro e uma mensagem.

Se a resposta SOAP do serviço da Web tiver dados XML que não são válidos, a transformação de Consumidor de Serviço da Web retornará um erro.

Você pode configurar se deseja tratar falhas SOAP como erros. Quando você habilita a opção Tratar Falha como Erro e uma falha SOAP ocorre, o Data Integration Service incrementa a contagem de erros para o mapeamento. A falha aparece no log de mensagens.

## Compactação de Mensagens

Quando você habilita a compactação de mensagens SOAP, a transformação de Consumidor de Serviço da Web compacta mensagens de solicitação de serviços da Web e recebe mensagens de resposta de serviços da Web compactadas.

A transformação de Consumidor de Serviço da Web codifica a solicitação SOAP com a compactação GZip. A transformação aceita uma mensagem de resposta codificada com a compactação GZip ou deflate.

Quando o Data Integration Service receber a resposta do serviço da Web, ele verificará o cabeçalho HTTP de Codificação de Conteúdo na mensagem SOAP e decodificará essa mensagem.

O padrão é a codificação sem compactação. O serviço da Web não compacta a resposta SOAP.

A seguinte tabela mostra os cabeçalhos nas mensagens de solicitação e de resposta quando a compactação está ativada ou desativada:

Compactação	Cabeçalho
Ativo	Cabeçalho Content-Encoding: GZip Cabeçalho Accept-Encoding: GZip, deflate
Inativo	Cabeçalho Content-Encoding Vazio Cabeçalho Accept-Encoding Vazio

Às vezes, um serviço da Web codifica uma mensagem de resposta com uma compactação padrão. A transformação de Consumidor de Serviço da Web decodificará a mensagem se ela estiver compactada com GZip ou deflate. A transformação de Consumidor de Serviço da Web registrará uma mensagem no log de mapeamento se o serviço da Web codificar a mensagem de resposta inesperadamente.

Habilite a compactação nas propriedades avançadas da transformação.

## Simultaneidade

Você pode ativar a transformação de Consumidor de Serviço da Web para criar diversas conexões simultâneas com um serviço da Web para que ele possa enviar várias solicitações de serviço da Web em paralelo.

Por exemplo, ao consultar informações bancárias, você pode configurar a transformação de Consumidor de Serviço da Web para simultaneidade, de forma que várias linhas sejam enviadas em paralelo. Se houver 20 linhas de entrada, será possível enviar 20 solicitações simultaneamente para agilizar o processamento.

Habilitando a simultaneidade na transformação de Consumidor de Serviço da Web, você pode configurar o limite de consumo de memória total.

Habilitando a simultaneidade na transformação de Consumidor de Serviço da Web, você pode configurar o número de conexões de serviço da Web simultâneas.

### Regras e Diretrizes para Simultaneidade

Use as seguintes regras e diretrizes ao usar a simultaneidade:

- A simultaneidade oferece suporte para linhas de entrada classificadas como várias conexões simultâneas com um serviço da Web. Não há suporte para linhas de saída ordenadas.
- Use a simultaneidade se o conjunto de dados tiver mais de 100 linhas.
- É aconselhável não aumentar o número de conexões de serviço da Web simultâneas. O número de conexões simultâneas de serviços da Web está vinculado ao número de soquetes usados pelo sistema operacional. É caro aumentar o número de soquetes.
- Use sistemas que possuem processadores de vários núcleos com um mínimo de 100 MB de RAM para obter o desempenho ideal ao usar o recurso de simultaneidade.
- O limite de memória de simultaneidade representa a memória consumida por fluxos de trabalho durante a invocação de serviços da Web.
- Habilitando a simultaneidade na transformação de Consumidor de Serviço da Web, você pode configurar o limite de consumo de memória. Certifique-se de que o consumo de memória ultrapasse a RAM física no servidor.

### Práticas Recomendadas para Simultaneidade

Para obter o desempenho ideal ao usar a simultaneidade, observe as seguintes práticas recomendadas:

- Evite alterar os valores padrão do limite total de memória de simultaneidade e do limite de conexão de simultaneidade.
- Evite usar a simultaneidade para conjuntos de dados com menos de 100 linhas.
- Evitar portas de passagem no mapeamento ao usar a simultaneidade.

# Otimização de Filtro

A otimização de filtro melhora o desempenho ao reduzir o número de linhas que passam pelo mapeamento. O Data Integration Service pode aplicar a otimização com seleção antecipada ou a otimização de envio.

Quando o Data Integration Service aplica um método de otimização de filtro, ele move um filtro o mais próximo possível da origem em um mapeamento. Se o Data Integration Service não conseguir mover um filtro antes de uma transformação em um mapeamento, talvez ele consiga enviar a lógica do filtro para uma transformação.

## Habilitando a Otimização de Seleção Antecipada com a Transformação de Consumidor de Serviço da Web

Habilite a otimização com seleção antecipada para a transformação de Consumidor de Serviço da Web quando essa transformação não apresenta efeitos colaterais e não trata falhas como erros.

1. Abra a exibição **Propriedades Avançadas** da transformação de Consumidor de Serviço da Web.
2. Desmarque **Tratar Falha como Erro**.
3. Desmarque **Tem Efeitos Colaterais**.

## Otimização de Envio com a Transformação de Consumidor de Serviço da Web

Com a otimização de envio, a transformação de Consumidor de Serviço da Web recebe o valor do filtro em uma porta de filtro. A porta de filtro é uma porta de entrada não conectada que você identifica como uma porta de filtro quando configura a otimização de envio. A porta de filtro tem um valor padrão que garante que o serviço da Web retorne todas as linhas quando a consulta do usuário final não contém filtros. A porta de filtro não é uma porta de passagem.

**Nota:** O campo de filtro deve fazer parte do grupo raiz na solicitação de serviço da Web.

O campo de filtro não pode ser uma porta de passagem. Quando você configura uma porta de filtro, o valor padrão da porta se transforma no valor da condição de filtro e, portanto, o valor da porta de saída de passagem é alterado. Um filtro baseado na porta de passagem de saída retorna resultados inesperados.

Você pode enviar várias expressões de filtro para a transformação de Consumidor de Serviço da Web. Cada condição de filtro deve ter o seguinte formato:

```
<Field> = <Constant>
```

As condições de filtro devem ser associadas por AND. Não é possível associar as condições com um OR.

## Otimização de Envio com a Transformação de Consumidor de Serviço da Web - Exemplo

Um serviço de dados SQL retorna pedidos para todos os clientes ou retorna pedidos para um cliente específico com base na consulta SQL que ele recebe do usuário.

O serviço de dados contém um objeto de dados lógicos com os seguintes componentes:

### Tabela Customer

Uma tabela de banco de dados Oracle que contém informações de clientes.

### Transformação de Consumidor de Serviço da Web

Uma transformação que chama um serviço da Web para recuperar os pedidos mais recentes para clientes. A transformação de Consumidor de Serviço da Web tem portas de entrada para `customerID` e `orderNum`. A transformação tem portas de passagem que contêm dados de clientes que ela recebe da tabela `Customer`. A porta `orderNum` é a porta de filtro e não está conectada. `orderNum` tem o valor padrão `"*"`. Quando o serviço da Web recebe esse valor na solicitação de serviço da Web, ele retorna todos os pedidos.

### Tabela virtual Orders

Uma tabela virtual que recebe os dados de pedidos e clientes do serviço da Web. O usuário final consulta essa tabela. A tabela `Orders` contém uma coluna `"customer"`, uma coluna `orderID` e dados de clientes e pedidos.

O usuário final transmite a seguinte consulta SQL ao serviço de dados SQL:

```
SELECT * from OrdersID where customer = 23 and orderID = 56
```

O Data Integration Service divide a consulta para otimizar o mapeamento. O Data Integration Service usa a otimização com seleção antecipada e move a lógica do filtro, `customer = 23`, para a leitura da tabela `Customer`. O Data Integration Service usa a otimização de envio e envia a lógica de filtro, `orderID = 56`, para a porta de filtro da transformação de Consumidor de Serviço da Web. A transformação de Consumidor de Serviço da Web recupera `ordersID 56` para `customer 23`.

## Habilitando a Otimização de Envio com a Transformação de Consumidor de Serviço da Web

Habilite a otimização de envio para a transformação de Consumidor de Serviço da Web quando essa transformação não apresenta efeitos colaterais e não trata falhas como erros.

1. Abra a exibição **Propriedades Avançadas** da transformação de Consumidor de Serviço da Web.
2. Desmarque **Tratar Falha como Erro**.
3. Desmarque **Tem Efeitos Colaterais**.
4. Clique no botão **Abrir** na propriedade **Otimização de Envio**.
5. Escolha o nome da porta de filtro na caixa de diálogo **Entrada Otimizada**. Você pode escolher várias portas de filtro.
6. Clique na coluna **Saída**.
7. Para cada porta de filtro, escolha a porta de saída que contém a coluna filtrada na resposta do serviço da Web.
8. Insira um valor padrão para cada porta de filtro.

**Nota:** Não é possível configurar um valor padrão para uma porta de Consumidor de Serviço da Web a menos que ela seja uma porta de filtro.

# Criando uma Transformação de Consumidor de Serviço da Web

Você pode criar uma transformação de Consumidor de Serviço da Web reutilizável ou não reutilizável. Transformações reutilizáveis podem existir em vários mapeamentos. Transformações não reutilizáveis existem dentro de um único mapeamento.

Você pode criar transformações de Consumidor de Serviço da Web para uma vinculação de SOAP 1.1 e uma vinculação de SOAP 1.2 a partir de um único objeto WSDL.

1. Para criar uma transformação, use um dos métodos a seguir:

Opção	Descrição
Reutilizável	Selecione um projeto ou pasta na exibição <b>Object Explorer</b> . Clique em <b>Arquivo &gt; Novo &gt; Transformação</b> . Selecione a transformação de Consumidor de Serviço da Web e clique em <b>Avançar</b> .
Não reutilizável	Em um mapeamento ou maplet, arraste uma transformação de Consumidor de Serviço da Web da paleta de Transformação para o editor.

A caixa de diálogo **Nova Transformação de Consumidor de Serviço da Web** é exibida.

2. Procure e selecione um objeto de dados WSDL para definir as mensagens de solicitação e resposta do serviço da Web.

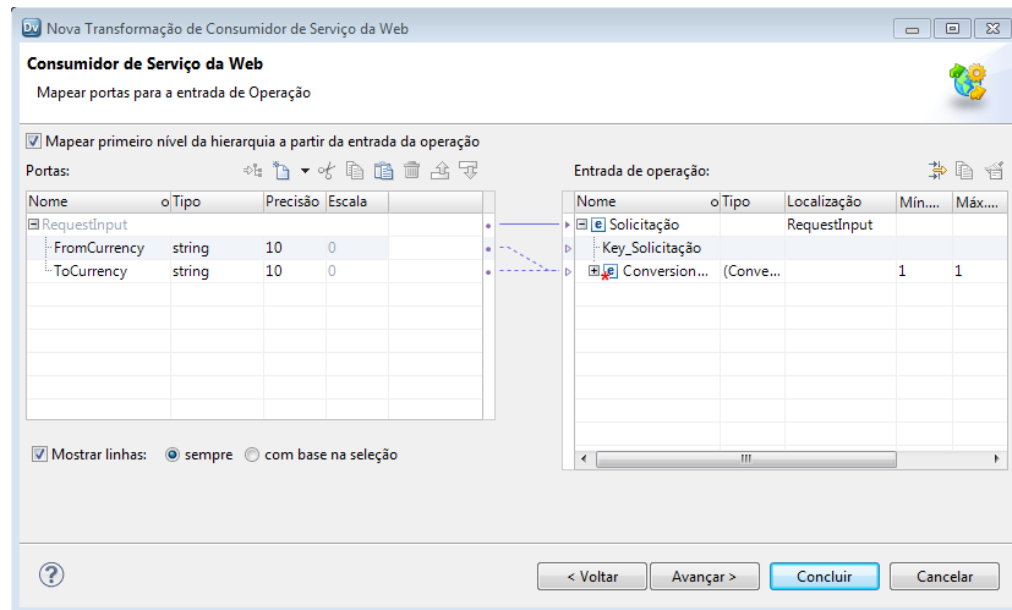
Se o WSDL não estiver no repositório, você poderá importar o WSDL da caixa de diálogo Nova Transformação de Consumidor de Serviço da Web.

3. Procure e selecione uma operação a partir do WSDL.

Você pode escolher uma operação que tem uma vinculação de SOAP 1.1 ou SOAP 1.2.

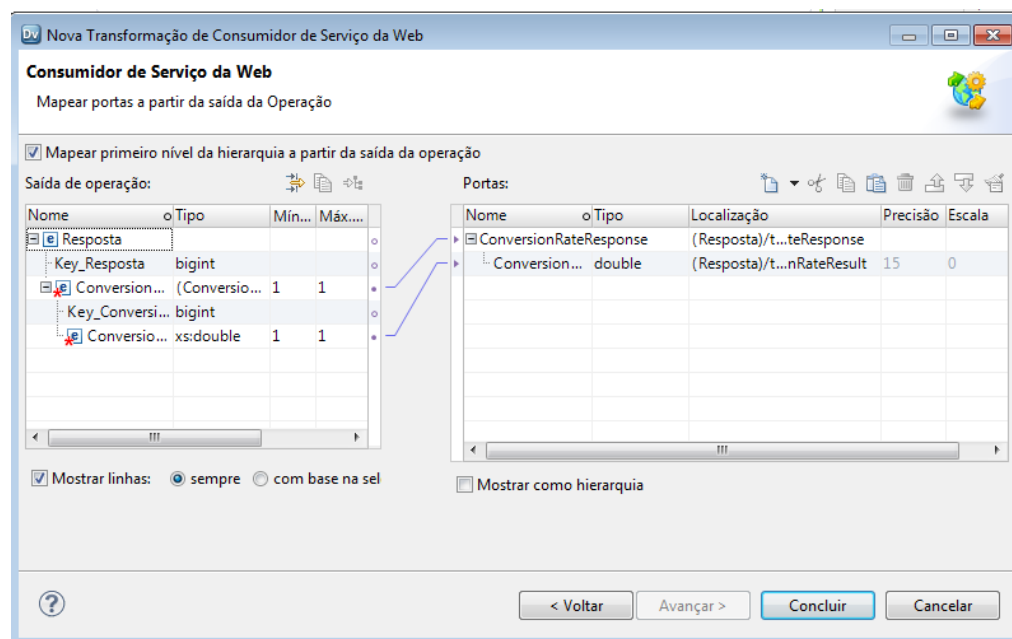
4. Clique em **Avançar**.

A tela **Mapear Portas para a Entrada de Operação** é exibida. A área **Portas** mostra os grupos de entrada da transformação e as portas de entrada. A área **Entrada de Operação** mostra a hierarquia de mensagens de solicitação.



5. Defina os grupos de entrada e as portas de entrada e mapeie as portas de entrada para nós de entrada de operação.
6. Clique em **Avançar**.

A tela **Mapear Portas a partir da Saída da Operação** é exibida. A área **Saída da Operação** mostra a hierarquia da mensagem de resposta. A área **Portas** mostra as portas e grupos de saída da transformação.



7. Defina os grupos de saída e as portas de saída e mapeie os nós de saída de operação para as portas de saída.



8. Clique em **Concluir**.
9. Clique na exibição **Avançado** para configurar as propriedades da transformação e a conexão de serviços da Web.

#### TÓPICOS RELACIONADOS:

- [“Mapeando Portas de Entrada para a Entrada de Operação” na página 679](#)
- [“Mapeando a Saída de Operação para Portas de Saída” na página 683](#)
- [“Propriedades Avançadas da Transformação de Consumidor de Serviço da Web” na página 684](#)

## Exemplo de Transformação de Consumidor de Serviço da Web

Sua organização precisa expor informações de pedidos referentes à linha de produtos RT100 para a organização de vendas. A equipe de vendas precisa consultar o resumo e os detalhes dos pedidos diariamente.

Crie um objeto de dados lógicos que exponha as informações de pedidos diárias em tabelas virtuais. O mapeamento de leitura contém uma transformação de Consumidor de Serviço da Web que retorna os pedidos mais recentes para a linha de produtos RT100. A transformação de Consumidor de Serviço da Web consome um serviço da Web que retorna informações diárias de resumo e detalhes de pedidos para a linha de produtos RT100.

### Arquivo de Entrada

O arquivo de entrada é um arquivo simples que contém o número da linha de produtos.

Crie um objeto de dados físicos para definir o arquivo de entrada. O arquivo tem um campo, Product\_Line. O valor do campo é RT100. Defina a localização do objeto de dados físicos na exibição **Propriedades do Tempo de Execução**.

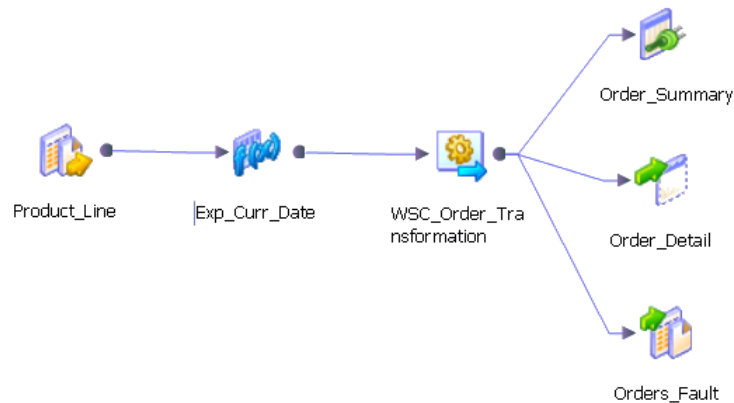
### Modelo de Objeto de Dados Lógicos

Um analista comercial na sua organização cria um modelo de dados lógicos que descreve as estruturas das tabelas de resumo e detalhes de pedidos. O modelo de dados lógicos contém os objetos de dados lógicos Order\_Summary e Order\_Detail.

O analista cria um esquema em uma ferramenta de modelagem que define o modelo de dados lógicos. Você importa o modelo de dados lógicos a partir do esquema e cria os objetos de dados lógicos Order\_Summary e Order\_Detail.

## Mapeamento de Objeto de Dados Lógicos

O mapeamento de objeto de dados lógicos descreve como acessar dados através do objeto de dados lógico.



O mapeamento de leitura contém os seguintes objetos:

### **Product\_Line**

Arquivo simples de entrada que contém o número da linha de produtos.

### **Transformação Exp\_Curr\_Date**

Transformação de Expressão que retorna a data atual e uma chave primária para o grupo de entrada de nível raiz da transformação de Consumidor de Serviço da Web.

### **Transformação WSC\_Order**

Transformação de Consumidor de Serviço da Web que consome um serviço da Web para recuperar informações de pedidos. A transformação transmite a linha de produtos e a data atual ao serviço da Web na mensagem de solicitação. A transformação recebe informações de pedidos do serviço da Web na mensagem de resposta.

### **Tabela Order\_Summary**

Um objeto de dados lógicos que contém informações de pedidos, como Order\_No, Customer\_Id, Qty e Order\_Date.

### **Tabela Order\_Detail**

Um objeto de dados lógicos que contém informações de detalhes de pedidos, como Order\_No, Product\_Id, Qty e Status.

### **Orders\_Fault**

Arquivo simples de saída que recebe mensagens de falha genérica.

## Transformação de Consumidor de Serviço da Web

A transformação de Consumidor de Serviço da Web recebe uma linha de produtos, uma data e um número de sequência como entrada. A transformação consome a operação do serviço da Web Get\_Order\_Info para recuperar informações de pedidos.

Quando você criar a transformação de Consumidor de Serviço da Web, escolha um objeto de dados WSDL que descreva as mensagens de solicitação e resposta do serviço da Web. Uma mensagem de serviço da Web contém grupos hierárquicos de elementos XML. Um elemento pode conter outros elementos. Alguns elementos ocorrem várias vezes. Crie a transformação a partir do objeto WSDL Order\_Info no repositório.

Configure as portas de entrada da transformação e mapa essas portas para a hierarquia de entrada de operação. Mapeie nós da hierarquia de saída de operação para as portas de saída. Defina a conexão de serviços da Web e as propriedades de tempo de execução.

## Mapeamento de Entrada da Transformação

Ao mostrar o mapeamento de entrada na exibição **Portas**, você pode definir portas de entrada e mapeá-las para nós na entrada de operação.

A área **Portas** da transformação tem um grupo raiz e um grupo Order. O grupo raiz é o grupo de entrada Request. Adicione uma porta ao grupo de entrada Request para representar a chave primária.

O grupo Order tem as portas de entrada **Select\_Date** e **Select\_Product\_Line**.

Mapeia as portas de entrada para os nós **Order\_Date** e **Product\_Line** na área **Entrada de Operação**.

A área **Entrada de Operação** define a mensagem de solicitação transmitida pela transformação de Consumidor de Serviço da Web ao serviço da Web. Os nós são exibidos na área **Entrada de Operação** por padrão.

## Mapeamento de Saída da Transformação

Ao mostrar o mapeamento de saída na exibição **Portas**, você pode definir as portas de saída mapeando nós a partir da saída de operação para os grupos de saída da transformação.

O serviço da Web retorna a seguinte hierarquia em uma mensagem de resposta de serviço da Web:

```
Response
  Orders
    Order
      Key_Order
      Order_ID
      Order_Date
      Customer_ID
      Total_Qty
      Order_Details
        Order_Detail
          Product_ID
          Description
          Qty
          Status
```

O serviço da Web retorna vários pedidos. Order é um nó de ocorrência múltipla no nível Orders. Para cada pedido, o serviço da Web pode retornar vários detalhes de pedidos. Order\_Detail é um nó de ocorrência múltipla no nível Order\_Details.

**Nota:** A ferramenta Developer adiciona o nó Key\_Order na interface do usuário. É possível mapear a chave para grupos de saída de forma a definir relacionamentos entre grupos. Por exemplo, Order\_ID é a chave primária em Order e é a chave externa em Order\_Details.

Crie os seguintes grupos de saída na área **Portas**:

```
Order
  Order_ID
  Order_Date
  Customer_ID
  Total_Qty

Order_Detail
  Order_ID
  Product_ID
  Description
  Qty
  Status
```

O Data Integration Service grava uma linha do grupo Order sempre que o valor de Order\_ID é alterado.

O Data Integration Service grava uma linha do grupo Order\_Detail sempre que os valores de Order\_ID e Product\_ID são alterados.

## Propriedades Avançadas da Transformação

Configure as seguintes propriedades avançadas para a transformação de Consumidor de Serviço da Web:

### **Habilitar Tratamento de Falhas SOAP Genéricas**

Adiciona portas de saída que recebem mensagens de falha SOAP.

### **Conexão**

Escolha uma conexão de serviços da Web para acessar o serviço da Web.

### **Ativar Compactação**

A transformação de Consumidor de Serviço da Web compacta as mensagens da Web com GZIP.

## CAPÍTULO 47

# Analizando Mensagens SOAP de Serviços da Web

Este capítulo inclui os seguintes tópicos:

- [Analizando Mensagens SOAP de Serviços da Web - Visão Geral, 697](#)
- [Interface do Usuário de Transformações, 698](#)
- [Configuração da Saída de Ocorrência Múltipla, 699](#)
- [Analizando Elementos anyType, 701](#)
- [Analizando Tipos Derivados, 702](#)
- [Analizando Elementos QName, 703](#)
- [Analizando Grupos de Substituição, 704](#)
- [Analizando Construções XML em Mensagens SOAP, 704](#)

## Analizando Mensagens SOAP de Serviços da Web - Visão Geral

O Serviço de Integração de Dados gera dados de linha quando analisa uma mensagem SOAP em uma transformação de serviço da Web.

A transformação de Consumidor de Serviço da Web e a transformação de Entrada de serviço da Web são transformações de serviço da Web que analisam mensagens SOAP.

Para configurar uma transformação de forma a analisar uma mensagem SOAP, crie portas de saída em uma estrutura semelhante à hierarquia da mensagem SOAP. Mapeie os nós na hierarquia da mensagem SOAP para essas portas.

Você pode configurar grupos normalizados de portas de saída, grupos desnormalizados e grupos dinâmicos de portas. Quando a mensagem SOAP contém tipos derivados, elementos anyType ou grupos de substituição, é possível configurar diferentes grupos de saída com base em quais tipos podem ocorrer na instância da mensagem SOAP.

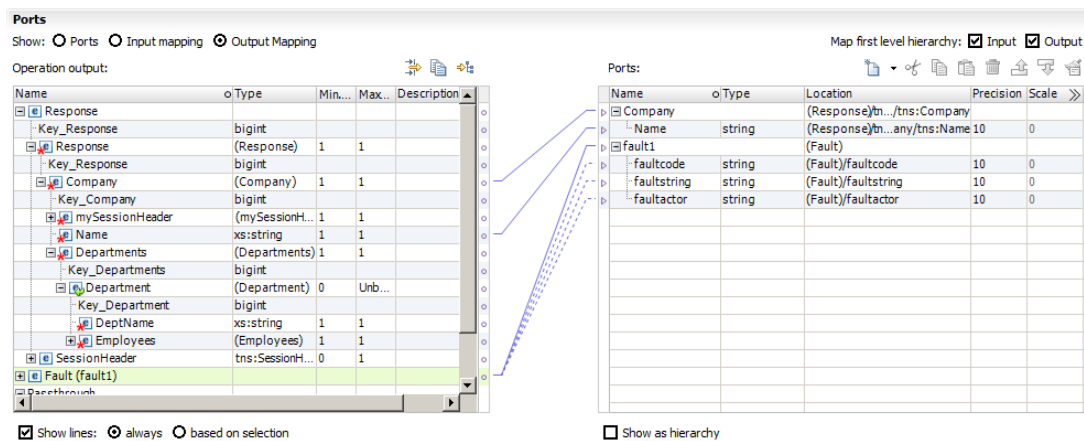
## TÓPICOS RELACIONADOS:

- [“Transformação de Consumidor de Serviço da Web - Mapeamento de Saída” na página 681](#)

# Interface do Usuário de Transformações

A transformação de Consumidor de Serviço da Web e a transformação de Entrada de serviços da Web fornecem uma interface de usuário que pode ser usada para mapear dados a partir da mensagem SOAP para as portas de saída da transformação.

A figura a seguir mostra um mapeamento entre nós de mensagem SOAP 1.1 e portas de saída em uma transformação de Consumidor de Serviço da Web:



## Área Operação

A área Operação contém a hierarquia da mensagem SOAP. Nós complexos ou nós de ocorrência múltipla definem níveis de hierarquia na estrutura. A ferramenta Developer adiciona chaves aos níveis que definem os relacionamentos de pai/filho entre eles.

Na figura anterior, a hierarquia da mensagem SOAP tem os seguintes níveis:

### Resposta ou Solicitação

Nível que representa a raiz da resposta ou da mensagem de solicitação.

### Companhia

Nível superior dos dados da solicitação.

### Departamentos

Departamentos de ocorrência múltipla dentro da empresa.

### Funcionários

Funcionário é um elemento complexo dentro de um departamento.

### Grupo de Falhas

Grupo de mensagens de falha que recebe mensagens de erro.

## Área Portas

Você pode mapear dados de níveis da mensagem SOAP para portas de saída. Cada grupo de portas de saída pode estar relacionado a outros grupos de saída com relacionamentos de chaves primárias/chaves externas.

Na figura anterior, a transformação tem grupos de portas de saída que correspondem aos grupos de nós na mensagem SOAP.

# Configuração da Saída de Ocorrência Múltipla

Quando uma transformação de Entrada ou uma transformação de Consumidor de Serviço da Web retorna dados de ocorrência múltipla, é possível configurar portas de saída em diferentes configurações.

Você pode configurar dados de saída normalizados, dados de saída dinamizados ou dados de saída desnormalizados.

Por exemplo, uma mensagem SOAP contém os elementos complexos Departamentos e Funcionários. Cada departamento contém vários funcionários. Departamentos é o pai de Funcionários.

A mensagem SOAP contém a seguinte hierarquia de elementos:

```
Departments
  Department_ID
  Department_Name
  Employees
    Employee_ID
    Employee_Name
```

## Saída Relacional Normalizada

Quando dados de saída normalizados são criados, os valores de dados não se repetem em um grupo de saída. Você cria um relacionamento de um-para-um entre os níveis da hierarquia na mensagem SOAP e os grupos de portas de saída.

Quando a mensagem SOAP contém um nível de hierarquia pai Departamentos e um nível de hierarquia filho Funcionários, é possível criar os seguintes grupos de portas:

```
Departments
  Department_Key
  Department_ID
  Department_Name

Employees
  Department_Key
  Employee_ID
  Employee_Name
```

Department\_Key é uma chave gerada que relaciona o grupo de saída Funcionários a um grupo Departamento.

## Chaves Geradas

Quando você adiciona um grupo de saída, a ferramenta Developer o relaciona a outro grupo de saída com uma chave gerada. A ferramenta Developer adiciona uma chave bigint ao grupo pai e ao grupo filho. Em tempo de execução, o Data Integration Service cria os valores de chaves para as chaves geradas.

## Exemplo

A hierarquia SOAP tem os seguintes nós:

```
Departments
  Dept_Key
  Dept_Num
  Dept_Name

  Employees
    Dept_FK
    Employee_Num
    Employee_Name
```

Ao criar um grupo de saída de portas para Departamentos, você mapeia o nó Departamentos para um campo vazio na área Portas. A ferramenta Developer cria o seguinte grupo de saída:

```
Departments
  Dept_Num
  Dept_Name
```

Quando você mapeia o nó Funcionários para um campo vazio na área Portas, a ferramenta Developer solicita que o grupo Funcionários seja relacionado ao grupo Departamentos. É possível relacionar o grupo Funcionários a mais de um grupo. A ferramenta Developer adiciona uma chave a cada grupo.

A ferramenta Developer cria os seguintes grupos e chaves geradas:

```
Departments
  Key_Departments
  Dept_Num
  Dept_Name

  Employees
    Key_Departments
    Employee_Num
    Employee_Name
```

**Nota:** Você não precisa mapear nós para as chaves geradas. O Data Integration Service cria os valores de chaves em tempo de execução.

A ferramenta Developer pode criar chaves geradas em vários níveis de um grupo de saída. O grupo Funcionários pode conter as seguintes portas:

```
Employees
  Key_Employees
  Key_Departments
  Key_Managers
  Employee_Num
  Employee_Name
```

Key\_Departments e Key\_Managers são as chaves geradas que apontam para grupos pai. Key\_Employees é uma chave gerada para o grupo Funcionários. Key\_Employees aparece quando você relaciona um grupo filho ao grupo Funcionários.

## Saída Relacional Desnormalizada

É possível desnormalizar uma saída relacional. Quando você desnormaliza os dados de saída, os valores de elementos do grupo pai se repetem para cada elemento filho.

Para desnormalizar dados de saída, mapeie nós a partir do nível de hierarquia pai para o grupo filho de portas de saída.

O exemplo a seguir mostra Department\_ID e Department\_Name no grupo de saída Funcionários:

```
Employees
  Department_ID
  Department_Name
```



```
Employee_ID
Employee_Name
```

Department\_ID e Department\_Name se repetem para cada funcionário do departamento:

Department_ID	Department_Name	Employee_ID	Employee_Name
100	Contabilização	56500	Kathy Jones
100	Contabilização	56501	Tom Lyons
100	Contabilização	56509	Bob Smith

## Saída Relacional Dinamizada

Você pode incluir um número específico de elementos de ocorrência múltipla em um grupo de saída.

Para dinamizar elementos de ocorrência múltipla, mapeie o elemento filho de ocorrência múltipla para o grupo pai de portas de saída. A ferramenta Developer solicita que você defina o número de elementos filho a serem incluídos no pai.

O exemplo a seguir mostra duas instâncias de Employee\_ID no grupo pai Departamentos:

```
Departments
  Department_ID
  Department_Name
  Employee_ID1
  Employee_ID2
```

## Analisando Elementos anyType

O elemento anyType representa uma opção de todos os tipos globais em um WSDL ou esquema. Ao mapear nós para portas na ferramenta Developer, você escolhe quais tipos devem aparecer na mensagem SOAP para o elemento anyType. É necessário substituir um elemento anyType na mensagem SOAP por um tipo complexo ou xs:string. Crie grupos de portas para cada tipo que você escolher.

É necessário escolher um tipo para mapear dados para portas de saída. Se o WSDL ou o esquema não contiver um tipo global, a ferramenta Developer substituirá o elemento anyType por xs:string.

Para escolher um tipo de elemento na área Operação, clique em **Escolher** na coluna **Tipo** para o elemento anyType. Uma lista de tipos complexos e xs:string disponíveis é exibida.

Quando você substitui um elemento anyType por tipos derivados, o Data Integration Service preenche elementos para um tipo de cada vez. A mensagem SOAP não contém dados para o tipo base e o tipo derivado ao mesmo tempo.

### Exemplo de Tipos Derivados

O WSDL contém um elemento anyType. Substitua o elemento por AddressType e um tipo derivado denominado USAddressType. A hierarquia da mensagem SOAP possui os seguintes grupos:

```
Address:AddressType (base type)
  Address: AddressType
    Street
    City

Address:USAddressType (derived type)
  Street
  City
  State
```

ZipCode

A mensagem SOAP contém os seguintes dados:

```
<address xsi:type="AddressType">
  <street>1002 Mission St.</street>
  <city>san jose</city>
</address>

<address xsi:type="USAddressType">
  <street>234 Fremont Blvd</street>
  <city>Fremont</city>
  <zip>94556</zip>
  <state>CA</state>
</address>
```

O Data Integration Service retorna uma linha para xsi: AddressType:

Rua	City
1002 Mission St.	San Jose

O Data Integration Service retorna uma linha para o tipo derivado xsi: USAddressType:

Rua	City	Estado	Zip
234 Fremont Blvd.	Sunnyvale	AC	94556

O Data Integration Service não preencherá AddressType se o tipo for xsi: USAddressType.

## Analisando Tipos Derivados

Você pode analisar mensagens SOAP que contêm tipos derivados. Ao definir as portas que recebem dados da mensagem SOAP, escolha quais tipos podem aparecer em uma mensagem SOAP. Os elementos nos tipos escolhidos determinam as portas que você precisa criar.

Por exemplo, o WSDL pode conter AddressType e um tipo derivado denominado USAddressType. É possível criar os seguintes grupos na área Operação da ferramenta Developer:

```
Address
  Address: AddressType
    Street
    City

  Address:USAddressType
    Street
    City
    State
    ZipCode
```

A mensagem SOAP pode conter os seguintes dados:

```
<address>
  <street>1002 Mission St.</street>
  <city>san jose</city>
</address>

<address xsi:type="USAddressType">
  <street>234 Fremont Blvd</street>
  <city>Fremont</city>
  <zip>94556</zip>
```

```

<state>CA</state>
</address>

<address xsi:type="USAddressType">
<street>100 Cardinal Way</street>
<city>Redwood City</city>
<zip>94536</zip>
<state>CA</state>
</address>

<address>
<street>100 El Camino Real</street>
<city>Sunnyvale</city>
</address>

```

O Data Integration Service retorna as seguintes linhas para o tipo base, Address:

Rua	City
1002 Mission St.	San Jose
234 Fremont Blvd	Sunnyvale
100 Cardinal Way	Redwood City
100 El Camino Real	Sunnyvale

O Data Integration Service retorna as seguintes linhas para o tipo derivado, USAddress:

Rua	City	Estado	Zip
234 Fremont Blvd.	Sunnyvale	AC	94556
100 Cardinal Way	Redwood City	AC	94536

O Data Integration Service retorna todos os endereços no tipo base. O Data Integration Service retorna endereços dos EUA no tipo derivado. O tipo derivado inclui os elementos Street e City que USAddressType herda do tipo base.

## Analizando Elementos QName

Quando o Data Integration Service analisa elementos QName na mensagem SOAP, ele atualiza valores QName que pertencem ao espaço de nome do esquema para usar o prefixo do espaço de nome definido nesse esquema. Caso contrário, o Data Integration Service não atualiza o valor do elemento.

Por exemplo, o esquema tem o prefixo de espaço de nome `tns` definido para o espaço de nome `"http://user/test"`. A mensagem SOAP tem o prefixo de espaço de nome `mytns` definido para o mesmo espaço de nome. Quando o Data Integration Service analisa o valor QName `mytns:myelement`, ele altera esse valor para `tns:myElement`.

Quando o Data Integration Service gera elementos QName na mensagem SOAP, ele não atualiza o valor do elemento.

## Analizando Grupos de Substituição

Um grupo de substituição substitui um elemento por outro do mesmo grupo. Grupos de substituição são similares a tipos derivados, com a diferença de cada definição de elemento inclui um nome de grupo de substituição.

Você pode configurar um grupo de saída de portas que recebe elementos de um tipo específico em um grupo de substituição. Você pode criar um grupo de saída de portas diferente que recebe um elemento de outro tipo no grupo de substituição.

## Analizando Construções XML em Mensagens SOAP

Uma mensagem SOAP pode conter construções XML, como elementos de opção, lista e união

Com algumas limitações, as transformações de serviços da Web podem analisar mensagens SOAP contendo essas construções.

### Elemento de Opção

Um elemento de opção restringe um elemento filho a um dos elementos na declaração <choice>.

O texto a seguir mostra um elemento de pessoa que é um funcionário ou um fornecedor:

```
<xs:element name="person">
  <xs:complexType>
    <xs:choice>
      <xs:element name="employee" type="employee"/>
      <xs:element name="contractor" type="contractor"/>
    </xs:choice>
  </xs:complexType>
</xs:element>
```

É possível mapear elementos de opção usando os seguintes métodos:

- Crie portas de saída para cada elemento de opção em um grupo de saída. Alguns elementos terão valores nulos na linha de saída.
- Crie um grupo de saída para cada opção. Para o exemplo acima, crie um grupo de funcionários e um grupo de fornecedores. O Data Integration Service gera uma linha com base em qual elemento aparece na mensagem SOAP.

### Elemento de Lista

Uma lista é um elemento XML que pode conter vários valores de tipo simples, como "Monday Tuesday Wednesday".

O Data Integration Service pode retornar uma lista como um valor de string. Quando a mensagem SOAP contém uma lista, não é possível mapear itens dessa lista para linhas de saída separadas. Você poderá configurar uma transformação de Expressão para separar os elementos na lista se precisar deles separados em um mapeamento.

## Elemento de União

O elemento de união é um tipo simples que é uma união de mais de um tipo.

O texto a seguir mostra um elemento Size que é uma união de dois tipos simples, size\_no e size\_string:

```
<xs:element name="Size">
  <xs:simpleType>
    <xs:union memberTypes="size_no size_string" />
  </xs:simpleType>
</xs:element>
```

Para mapear Size para uma porta de saída, crie uma porta para Size. Configure a porta de saída como uma string. Você pode configurar outra transformação no mapeamento para converter os dados para outro tipo.

## CAPÍTULO 48

# Gerando Mensagens SOAP de Serviços da Web

Este capítulo inclui os seguintes tópicos:

- [Gerando Visão Geral de Mensagens SOAP de Serviços da Web, 706](#)
- [Interface do Usuário de Transformações, 707](#)
- [Relacionamentos de Níveis de Hierarquia e Portas , 708](#)
- [Chaves, 709](#)
- [Mapear Portas, 711](#)
- [Dinamizando Portas de Ocorrência Múltipla , 713](#)
- [Mapear Dados Desnormalizados, 714](#)
- [Tipos Derivados e Substituição de Elementos, 715](#)
- [Gerando Construções XML em Mensagens SOAP, 717](#)

## Gerando Visão Geral de Mensagens SOAP de Serviços da Web

O Serviço de Integração de Dados gera dados XML a partir de grupos de dados de entrada ao gerar uma mensagem SOAP. Ao criar uma transformação de Consumidor de Serviço da Web, um transformação de Saída de serviço da Web ou uma transformação de Falha, você configura quais portas de entrada devem ser mapeadas para a hierarquia da mensagem SOAP.

Para configurar uma transformação de forma a gerar uma mensagem SOAP, crie grupos de portas de entrada e mapeie cada grupo para um grupo na hierarquia da mensagem SOAP. Um WSDL ou esquema define a estrutura da mensagem SOAP.

Você pode configurar grupos de dados na mensagem SOAP a partir de dados de entrada desnormalizados. Você também pode dinamizar dados de entrada de ocorrência múltipla para nós de ocorrência múltipla na mensagem SOAP.

Você pode mapear dados para tipos derivados, elementos anyType ou grupos de substituição em uma mensagem SOAP. Você deve escolher quais tipos podem ocorrer na mensagem SOAP ao definir uma transformação. Os tipos escolhidos determinam as portas de entrada que você precisa criar.

Quando a hierarquia da mensagem SOAP é exibida na ferramenta Developer, ela contém chaves. Essas chaves não aparecem na mensagem SOAP. O Serviço de Integração de Dados usa chaves para definir

relacionamento de pai/filho entre grupos na mensagem SOAP. Para configurar valores de chaves, mapeie dados de entrada para chaves na mensagem SOAP.

## TÓPICOS RELACIONADOS:

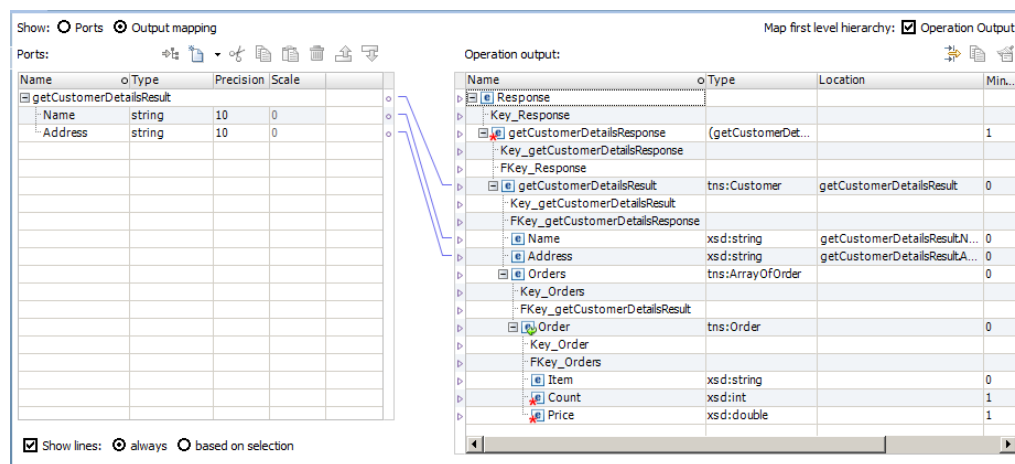
- [“Mapeamento de Entrada da Transformação de Consumidor de Serviço da Web” na página 678](#)

# Interface do Usuário de Transformações

A transformação de Saída de serviços da Web, a transformação de Falha e a transformação de Consumidor de Serviço da Web contêm uma interface de usuário que é usada para configurar a mensagem SOAP.

Para configurar uma transformação de forma a gerar uma mensagem SOAP, crie portas de entrada em uma estrutura semelhante à hierarquia da mensagem SOAP. O WSDL ou esquema determina a estrutura da hierarquia. Mapeie cada porta de entrada para um nó na mensagem SOAP.

A figura a seguir mostra um mapeamento entre as portas de entrada e os nós da mensagem SOAP em uma transformação de Saída de serviço da Web:



## Área Portas de Entrada

Crie grupos de portas de entrada na área **Portas de Entrada**. Inclua portas de entrada para cada nível na hierarquia da mensagem SOAP que você precisa mapear.

Você deve criar um grupo de entrada de Resposta ou Solicitação e os grupos filho que recebem os dados.

Ao criar grupos de portas de entrada, defina uma chave primária em cada grupo pai. Defina uma chave externa em cada grupo filho. A chave externa relaciona o grupo a um grupo pai.

Você não precisa definir chaves para o nível Resposta ou para o nível raiz do WSDL, a menos que esteja transmitindo dados no nível raiz do WSDL. Por exemplo, o nível raiz pode conter cabeçalhos HTTP.

Você pode criar grupos de portas semelhantes aos grupos a seguir para clientes e pedidos:

```

Response
  Response_Key

Customer_Details_Root
  Key_Cust_Det
  
```

```

FK_Response_Key

Customer
  Customer_ID
  FK_Cust_Det
  Name
  Address

Orders
  Order_Num
  FK_Cust_ID

Order_Items
  Order_Num
  Item
  Count
  Price

```

## Área Operação

A área **Operação** mostra os elementos na hierarquia da mensagem SOAP conforme definido pelo WSDL ou esquema. A mensagem SOAP não precisa contém todos os elementos do WSDL ou do esquema. A mensagem contém os dados que você mapeia a partir das portas de entrada.

Nós de ocorrência múltipla e nós complexos definem níveis de hierarquia na estrutura da mensagem SOAP. A ferramenta Developer adiciona chaves aos níveis para criar relacionamentos de pai/filho entre eles. Todos os níveis da hierarquia, exceto os níveis folha, possuem uma chave primária. Cada nível filho possui uma chave externa para um nível pai. As chaves que aparecem na hierarquia da mensagem SOAP não aparecem na instância da mensagem SOAP. O Data Integration Service precisa de valores nas chaves para relacionar níveis de dados ao gerar a mensagem SOAP.

A coluna **Localização** contém o nome do grupo e a porta de entrada que contém os dados para um elemento na mensagem SOAP. A coluna **Localização** ficará vazia até que você mapeie uma porta de entrada para o nó.

Na figura anterior, a mensagem SOAP contém uma única instância dos detalhes e pedidos do cliente. O grupo Orders contém um elemento de ocorrência múltipla denominado Order. A hierarquia da mensagem SOAP tem os seguintes níveis relacionados por chave:

```

Response
  GetCustomerDetailsResponse
    GetCustomerDetailsResult
      Orders
        Order

```

O nível Resposta representa a raiz da mensagem de resposta. O Data Integration Service precisa desse nível para anexar cabeçalhos à mensagem SOAP.

O nível GetCustomerDetailsResponse é a raiz da mensagem.

## Relacionamentos de Níveis de Hierarquia e Portas

Ao mapear portas de entrada para a hierarquia da mensagem SOAP, você mantém um relacionamento entre um grupo de entrada e um nível de hierarquia da mensagem SOAP. Por exemplo, você pode ter dois grupos de entrada, Departamento e Funcionário.



O grupo de entrada Departamento recebe as seguintes linhas:

Dept_num	Nome	Localização
101	HR	Nova York
102	Produto	Califórnia

O grupo de entrada Funcionário recebe as seguintes linhas:

Dept_num	Funcionário
101	Alice
101	Bob
102	Carol
102	Dave

Mapeie o número do departamento mapa no grupo Funcionário como uma chave externa que estabelece o relacionamento entre os grupos Departamento e Funcionário. O número do departamento ocorre no nível de hierarquia de departamento, mas não no nível de funcionário.

A mensagem SOAP contém a seguinte estrutura XML:

```
<department>
  <dept_num>101</dept_num>
  <name>HR</name>
  <location>New York</location>

  <employee>
    <name>Alice</name>
  </employee>

  <employee>
    <name>Bob</name>
  </employee>
</department>

<department>
  <dept_num>102</dept_num>
  <name>Product</name>
  <location>California</location>

  <employee>
    <name>Carol</name>
  </employee>

  <employee>
    <name>Dave</name>
  </employee>
</department>
```

## Chaves

Uma hierarquia de mensagem SOAP inclui chaves. O Serviço de Integração de Dados precisa de valores de chave para construir a hierarquia XML na mensagem SOAP.

Você deve mapear dados de portas de entrada para as chaves na hierarquia da mensagem SOAP. Mapeie dados para as chaves em cada nível para o qual que está fornecendo dados. Quando você tiver um nó de ocorrência múltipla, será necessário relacionar esse nó a um pai.

As chaves aparecem na mensagem SOAP sem tipos. Qualquer porta mapeada para uma chave deve ter um tipo de dados de string, inteiro ou bigint. A chave primária no grupo pai e a chave externa em cada grupo

filho devem ter o mesmo tipo de dados, precisão e escala. Você pode mapear as chaves geradas para as chaves da mensagem SOAP.

Você pode mapear uma porta para um nó e para uma chave no mesmo nível de hierarquia. Por exemplo, é possível mapear Employee\_ID para um nó na mensagem SOAP e mapeá-lo para uma chave no nível de Funcionário.

Se dois nós de grupo na hierarquia tiverem um relacionamento de pai/filho, conclua as seguintes tarefas:

- Mapeie uma porta para a chave primária no grupo de nós pai.
- Mapeie uma porta para a chave externa no grupo de nós filho.

Você também pode mapear chaves primárias para portas de entrada a fim de remover linhas com uma chave primária nula ou com chaves primárias duplicadas.

É possível uma chave composta em uma mensagem SOAP mapeando várias portas para a mesma chave. Use chaves compostas quando precisar desnormalizar dados e manter chaves exclusivas para algumas combinações de ocorrência múltipla de valores. Você pode criar chaves compostas que contêm valores de string, bigint ou inteiros.

**Nota:** Você pode incluir uma transformação de Expressão no mapeamento de operação para gerar valores de chave.

### Exemplo de Chave Composta

Configure uma chave exclusiva de divisão-departamento a partir dos seguintes grupos de portas:

```
Company
  Company_Num
  Company_Name

  Division
    Company_Num
    Division_Num
    Division_Name

    Department
      Division_Num
      Dept_Num
      Dept_Name
      Location
```

Dept\_Num é exclusivo em uma divisão, mas Dept\_Num não é exclusivo para todas as divisões da empresa.

Você pode configurar um grupo Departamento que contém as informações do departamento e da divisão. Configure o número da divisão e o número do departamento como parte da chave composta:

```
Department
  Division_Num + Dept_Num (key)
  Dept_Name
  Location
```

A ordem em que você mapeia as portas determina o valor da chave.

# Mapear Portas

Depois de criar portas de entrada, mapeie cada porta de entrada para a hierarquia da mensagem SOAP. A localização da porta aparecerá para o nó na área **Operação**.

Você pode mapear portas para os seguintes tipos de nós:

## **Nó atômico**

Um elemento simples ou um atributo simples que não tem filhos e não pode ser dividido.

## **Nó atômico de ocorrência múltipla**

Um elemento simples ou atributo que ocorre várias vezes na mesma localização da hierarquia.

## **Nó complexo**

Um elemento que contém outros elementos.

Se o nó pai não tiver uma localização, ele receberá o nome do grupo de entrada como a localização. Quando o nó pai tem uma localização, cada nó no nível de hierarquia deve ter uma localização de saída a partir da mesma localização.

É possível mapear um nome de grupo de entrada para um nó pai em um nível de hierarquia. A ferramenta Developer atualiza o campo de localização para o nó pai na hierarquia. A ferramenta Developer não atualiza os nós filhos que pertencem ao grupo na hierarquia. Quando você mapeia portas de entrada para os nós filho, a localização de cada uma dessas portas deve ser idêntica à do nó pai.

Você pode alterar um grupo de entrada depois de mapeá-lo para um nível de hierarquia. Você pode clicar em **Limpar** ou pode excluir as linhas entre as áreas Portas e Operação. Para excluir as linhas, arraste o ponteiro das linhas para selecioná-las. Clique em **Excluir**.

## Mapear uma Porta

Quando você mapeia uma porta para um nó na mensagem SOAP, a ferramenta Developer fornece resultados diferentes com base no tipo de nó para o qual a porta é mapeada.

A seguinte tabela descreve os resultados quando você mapeia uma única porta para diferentes nós no destino na área **Operação**:

Nó de Destino	Resultados
Nó atômico	Quando você mapeia uma única porta para um nó, e o nó pai não tem uma localização, esse nó recebe a localização da porta. A localização do nó pai recebe a localização do grupo de entrada para a única porta. Quando uma única porta é mapeada para um nó, e o nó pai já tem uma localização, você pode alterar a localização para o nó pai e limpar a localização para os outros nós filho no mesmo nível. A localização do nível de hierarquia muda para o nome de grupo da porta.
Nó atômico de ocorrência múltipla ou chave primária do nó atômico de ocorrência múltipla	Quando você mapeia uma única porta para o nó atômico de ocorrência múltipla, a ferramenta Developer define a localização do nó atômico como sendo o grupo da porta selecionada.
Nó complexo	Quando você mapeia uma única porta para um nó complexo, a ferramenta Developer define a localização desse nó complexo como a localização do grupo que contém a porta. A ferramenta Developer solicita que você especifique o nó atômico de ocorrência única ao qual a porta deve ser atribuída.  Se todos os nós atômicos de ocorrência única tiverem uma localização, não será possível mapear o nó complexo.

## Mapear um Grupo

Quando você mapeia um grupo de entrada para um nó na mensagem SOAP, a ferramenta Developer fornece resultados diferentes com base no tipo de nó para o qual a porta é mapeada.

A seguinte tabela descreve os resultados quando você mapeia um grupo para um nó na área **Operação**:

Nó de Destino	Resultados
Nó atômico	Não é possível mapear um grupo para um nó atômico.
Nó atômico de ocorrência múltipla	É solicitado que você escolha uma porta no grupo de entrada para atualizar a localização do nó e da chave primária.
Nó complexo de ocorrência múltipla	A ferramenta Developer define a localização para o nó complexo como a localização do grupo.

## Mapear Várias Portas

Quando você mapeia várias portas para um nó na mensagem SOAP, a ferramenta Developer fornece resultados diferentes com base no tipo de nó para o qual essas portas são mapeadas. Será possível mapear várias portas ao mesmo tempo se você as mapear a partir do mesmo grupo.

A seguinte tabela descreve os resultados para o nó quando você mapeia várias portas para nós:

Nó de Destino	Resultados
Nó atômico único	Ao mapear várias portas para um único nó, você atualiza a localização para mais de um nó atômico único na área <b>Operação</b> . Se a hierarquia não tiver um número suficiente de nós no nível para a atualização, a ferramenta Developer mapeará portas apenas para os nós disponíveis.
Nó atômico de ocorrência múltipla	Ao mapear várias portas para o nó atômico de ocorrência múltipla, você dinamiza essas portas em várias ocorrências do nó. A ferramenta Developer cria instâncias do nó com base no número de portas que são mapeadas. É exibida uma mensagem que descreve o número de portas que você projetou.
Nó complexo de ocorrência múltipla	Ao mapear várias portas para um nó complexo, você deve selecionar quais nós atômicos de ocorrência única serão atualizados. Você dinamiza as portas em várias ocorrências do nó. A ferramenta Developer cria instâncias do nó com base no número de portas que são mapeadas.

## Dinamizando Portas de Ocorrência Múltipla

Você pode mapear várias portas de entrada para um nó de ocorrência múltipla na mensagem SOAP. A ferramenta Developer dinamiza os dados de entrada em vários nós na mensagem SOAP.

Para alterar o número de elementos a serem dinamizados, escolha **Substituir dinamização existente** na caixa de diálogo **Opções de Mapa**.

Se você remover uma das instâncias de portas dinamizadas da área **Portas**, a ferramenta Developer removerá todas as instâncias da área **Operação**.

### Exemplo de Dinamização

Um grupo de entrada pode ter as seguintes linhas:

Num	Nome	Localização	emp_name1	emp_name2	emp_name3
101	HR	Nova York	Alice	Tom	Bob
102	Produto	Califórnia	Carol	Tim	Dave

Cada linha contém um número de departamento e três nomes de funcionários.

Funcionário é um nó de ocorrência múltipla na hierarquia da mensagem SOAP. É possível mapear todas as instâncias de Funcionário a partir da linha de entrada para a hierarquia da mensagem SOAP. Selecione todas as ocorrências de Funcionário. Clique em **Mapear**. A caixa de diálogo **Opções de Mapeamento** solicita que você escolha um nó na lista.

A ferramenta Developer altera o nó Funcionário de forma a incluir o nome de vários nós na hierarquia da mensagem SOAP:

```
Department
  num
  name
  location
```

```

Employee      (unbounded)
  emp_name1
  emp_name2
  emp_name3

```

A mensagem SOAP retorna a seguinte hierarquia:

```

<department>
  <num>101</num>
  <name>HR</name>
  <location>New York</location>
  <employee>
    <emp_name>Alice</name>
  </employee>
  <employee>
    <emp_name>Tom</name>
  </employee>
  <employee>
    <emp_name>Bob</name>
  </employee>
</department>

<department>
  <num>102</num>
  <name>Product</name>
  <location>California</location>
  <employee>
    <emp_name>Carol</name>
  </employee>
  <employee>
    <emp_name>Tim</name>
  </employee>
  <employee>
    <emp_name>Dave</name>
  </employee>
</department>

```

## Mapear Dados Desnormalizados

É possível mapear dados desnormalizados e transmiti-los para nós normalizados em uma mensagem SOAP.

Ao mapear dados desnormalizados, você transmite dados de um grupo de entrada para vários nós na hierarquia da mensagem SOAP. Você pode criar relacionamentos de grupos na mensagem SOAP semelhantes aos seguintes tipos de relacionamentos:

### Relacionamento Linear de Nós

O Nó A é pai do Nó B. O Nó B é pai do Nó C. O Nó C é pai do Nó D.

### Relacionamento Hierárquico de Nós

O Nó A é pai do Nó B. O Nó A também é pai do Nó C. Os Nós B e C não estão relacionados.

A seguinte tabela mostra linhas de entrada que contêm dados desnormalizados de divisão e departamento:

Divisão	Dept_Num	Dept_Name	Telefone	Employee_Num	Employee_Name
01	100	Contabilização	3580	2110	Amir
01	100	Contabilização	3580	2113	Robert
01	101	Engenharia	3582	2114	Stan
01	101	Engenharia	3582	2115	Jim
02	102	Instalações	3583	2116	Jose

Os dados de entrada contêm nomes e números exclusivos de funcionários. Os dados do departamento e da divisão se repetem para cada funcionário no mesmo departamento e divisão.

### Relacionamento Linear de Grupos

Ao configurar portas, você pode configurar um grupo separado para Divisão, Departamento e Funcionário. Divisão é o pai de Departamento, e Departamento é o pai de Funcionário. É possível configurar grupos na seguinte estrutura linear:

```
Division
  Division_Key
  Division_Num
  Division Name

  Department
    Department_Key
    Division_FKey
    Dept_Num
    Dept_Name
    Phone

    Employee
      Department_Fkey
      Employee_Num
      Employee_Name
```

A mensagem SOAP contém instâncias exclusivas de Divisão e Departamento, embora Division\_Num e Dept\_Num se repitam nos dados de entrada. Defina Division\_Num como a chave primária no grupo Divisão. Defina Dept\_Num como a chave primária no grupo Departamento.

### Relacionamento Hierárquico de Grupos

Você pode criar uma hierarquia de grupos que contém o grupo pai Divisão e os grupos filho Departamento e Funcionário. Departamento e Funcionário não tem um relacionamento de chaves primárias e chaves externas. Departamento e Funcionário são filhos de Divisão. É possível configurar os grupos na seguinte estrutura:

```
Division
  Division_Key
  Division_Num
  Division_Name

  Department
    Division_FKey
    Dept_Num
    Dept_Name

    Employee
      Division_FKey
      Employee_Num
      Employee_Name
```

## Tipos Derivados e Substituição de Elementos

É possível mapear portas de entrada para tipos complexos derivados, elementos anyType e grupos de substituição em uma mensagem SOAP. A mensagem SOAP pode incluir elementos para o tipo base e os tipos derivados.

Em um relacionamento de tipos, o tipo base é o tipo a partir do qual você deriva outro tipo. Um tipo derivado herda elementos do tipo base. Um tipo complexo estendido é um tipo derivado que herda elementos de um

tipo base e inclui elementos adicionais. Um tipo complexo restrito é um tipo derivado que restringe alguns elementos do tipo base.

## Gerando Tipos Derivados

Quando um WSDL ou esquema inclui tipos derivados, você deve escolher os tipos que deseja incluir na mensagem SOAP.

Por exemplo, o WSDL define um tipo base `AddressType`. O WSDL também contém `USAddressType` e `UKAddressType`, que são derivados de `AddressTypes`.

Cada tipo contém os seguintes elementos:

- `AddressType`: `street`, `city`
- `USAddressType` (estende `AddressType`): `state`, `zipCode`
- `UKAddressType` (estende `AddressType`): `postalCode`, `country`

Quando você escolhe um `USAddressType` na área Operação, a ferramenta Developer cria um grupo para o elemento `USAddressType` na mensagem SOAP. Esse grupo inclui "street" e "city" a partir do endereço base e "state" e "zipCode" para `USAddress`. Os tipos derivados que estendem tipos base sempre incluem os elementos do tipo base.

Se você escolher todos os tipos derivados disponíveis para a mensagem SOAP, a ferramenta Developer criará grupos semelhantes aos seguintes na hierarquia SOAP:

```
Address
  Address: Address
    Street
    City

  Address:USAddressType
    Street
    City
    State
    ZipCode

  Address: UKAddressType
    Street
    City
    PostalCode
    Country
```

É necessário definir grupos de porta de entrada para `Address`, `USAddress` e `UKAddress`.

## Gerando Elementos e Atributos anyType

Alguns elementos e atributos de esquema permitem qualquer tipo de dados em uma mensagem SOAP.

O elemento `anytype` representa uma opção de todos os tipos globalmente conhecidos. Antes de mapear uma porta para um elemento `anyType` em uma mensagem SOAP, escolha um tipo complexo disponível ou `xs:string`. Se o WSDL ou esquema não contiver um tipo complexo, a ferramenta Developer substituirá o tipo de elemento `anyType` por `xs:string`.

Para escolher um tipo de elemento na área Operação, clique em **Escolher** na coluna **Tipo** para o elemento `anyType`. Uma lista de tipos complexos e `xs:string` disponíveis é exibida.

Os seguintes elementos e atributos permitem qualquer tipo de dados:

### Elemento `anyType`

Permite que um elemento seja qualquer tipo de dados no arquivo XML associado.



### Elemento anySimpleType

Permite que um elemento seja qualquer simpleType no arquivo XML associado.

### Elementos de conteúdo ANY

Permite que um elemento seja qualquer elemento global definido no esquema.

### Atributo anyAttribute

Permite que um elemento seja qualquer atributo já definido no esquema.

## Gerando Grupos de Substituição

Use grupos de substituição para substituir um elemento por outro em uma mensagem SOAP. Grupos de substituição funcionam de maneira semelhante a tipos derivados, com a diferença de que as definições de elementos incluem um nome de grupo de substituição.

Por exemplo, você pode ter um tipo base Address e os tipos derivados USAddress e UKAddress:

```
xs:element name="Address" type="xs:string"/>
<xs:element name="USAddress" substitutionGroup="Address"/>
<xs:element name="UKAddress" substitutionGroup="Address"/>
```

Ao configurar a hierarquia da mensagem SOAP, você pode escolher qual elemento substituir para Address na mensagem SOAP.

## Gerando Construções XML em Mensagens SOAP

Um WSDL ou esquema pode conter elementos de opção, lista ou união. Transformações de serviços da Web podem gerar mensagens SOAP que contêm esses elementos.

### Elemento de Opção

Um elemento de opção restringe um elemento filho a um dos elementos na declaração <choice>.

Para mapear portas para uma mensagem SOAP que contém elementos de opção, crie um grupo de entrada que inclua todos os elementos na construção de opção. Por exemplo, uma descrição de item pode ser uma dimensão ou uma ponderação:

```
item: description, choice {dimension, weight}
```

Quando a descrição é uma dimensão, ela é um tipo complexo que contém comprimento largura e altura.

Quando a descrição é uma ponderação, ela é um tipo de string simples.

Os dados de entrada têm as seguintes colunas e linhas:

descrição	length	width	height	weight
box	20cm	18cm	15cm	NULL
coffee	NULL	NULL	NULL	500g

A mensagem SOAP contém um grupo de Itens que inclui dimensões ou descrições de ponderações:

```
Item
  Description
    Dimension
      Length
      Width
```

```
Height
Weight
```

Os valores NULL nos dados de entrada se tornam elementos ausentes na saída XML.

A mensagem SOAP contém os seguintes dados:

```
<item>
  <desc>box</desc>
  <dimension>
    <length>20cm</length>
    <width>18cm</width>
    <height>15cm</height>
  </dimension>
</item>

<item>
  <desc>coffee</desc>
  <weight>500g</weight>
</item>
```

## Elemento de Lista

Uma lista é um elemento XML que pode conter vários valores de tipo simples no mesmo elemento ou atributo. O Data Integration Service poderá processar uma lista nos dados de entrada se essa lista for representada como uma string consolidada de dados.

Se cada item da lista for um elemento separado, como ClassDates1, ClassDates2 e ClassDates3, o Data Integration Service não poderá processar os itens como uma lista. É possível usar uma transformação de Expressão para combinar esses itens em uma string se você precisar retornar uma lista em uma mensagem SOAP.

As seguintes linhas de entrada contêm um elemento de lista denominado ClassDates que contém os dias da semana:

CourseID	Nome	ClassDates
Matemática 1	Introdução à Álgebra	Seg Qua Sex
História 1	História do Mundo	Ter Qui

O Data Integration Service pode retornar uma mensagem SOAP com a seguinte estrutura XML:

```
<class>
  <courseId>Math 1</courseId>
  <name>Beginning Algebra</name>
  <classDates>Mon Wed Fri</classDates>
</class>
<class>
  <courseId>History 1</courseId>
  <name>World History</name>
  <classDates>Tue Thu</classDates>
</class>
```

## Elemento de União

O elemento de união é um tipo simples que é uma união de mais de um tipo. Quando uma mensagem SOAP contém um elemento de união, você deve mapear uma única porta de entrada que contenha os dados em uma string.

Por exemplo, a mensagem SOAP contém um elemento denominado "tamanho". "Tamanho" é uma união de inteiro e string:

```
<xs:element name="size">
  <xs:simpleType>
```

```

        <xs:union memberTypes="size_no size_string" />
    </xs:simpleType>
</xs:element>

```

As linhas de entrada contêm itens com uma descrição e um tamanho. Um item pode ter um tamanho de numérico, como 42. Ou, um item pode ter um tamanho que é um valor de string, como grande, médio ou pequeno.

A seguinte tabela mostra linhas de entrada com um tamanho numérico e um tamanho de string:

Descrição	Tamanho
sapatos	42
camisa	grande

Crie uma porta para o tamanho do item. Mapeie a porta como uma string. A mensagem SOAP contém os seguintes elementos:

```

<item>
  <desc>shoes</desc>
  <size>42</size>
</item>

<item>
  <desc>shirt</desc>
  <size>large</size>
</item>

```

## CAPÍTULO 49

# Transformação de Média Ponderada

Este capítulo inclui os seguintes tópicos:

- [Transformação de Média Ponderada - Visão Geral, 720](#)
- [Configurando uma Transformação de Média Ponderada, 720](#)
- [Exemplo de Pontuações de Correspondência Ponderada, 721](#)
- [Propriedades Avançadas da Transformação de Média Ponderada, 721](#)
- [Transformação de média ponderada em um ambiente não nativo, 722](#)

## Transformação de Média Ponderada - Visão Geral

A transformação de Média Ponderada é uma transformação passiva que lê pontuações de correspondência de várias operações de correspondência e produz uma única pontuação de correspondência.

Você pode aplicar um peso numérico a cada pontuação inserida na transformação de Média Ponderada. Um peso é um valor entre zero e 1. Você pode editar o peso aplicado a cada pontuação de entrada para aumentar ou reduzir sua contribuição para a pontuação de saída. Aplique pesos que reflitam a importância relativa de cada coluna de dados na análise de duplicata.

Use a transformação de Média Ponderada quando adicionar transformações de Comparação a um mapeamento ou mapplet.

**Nota:** Você também pode atribuir pesos em uma transformação de Correspondência. Use a transformação de Correspondência para configurar estratégias de correspondência e atribuir pesos em uma única transformação. Você pode incorporar um mapplet correspondente em uma transformação de Correspondência.

## Configurando uma Transformação de Média Ponderada

Use a transformação de Média Ponderada para ajustar a pontuação de correspondência geral que um mapeamento gera para uma série de operações de análise de correspondência. Edite os pesos relativos de cada porta de entrada para refletir as prioridades comparações de dados que você definiu para o conjunto de

dados de origem. Cada porta de entrada na transformação de Média Ponderada representa uma saída de pontuação de correspondência de uma estratégia de transformação de Comparação.

As etapas a seguir descrevem o processo para configurar uma transformação de Média Ponderada não reutilizável em um mapplet ou mapeamento que usa transformações de Comparação.

1. Abra um mapplet ou mapeamento de análise de correspondência e adicione um downstream de transformação de Média Ponderada das transformações de Comparação.
2. Conecte a saída da pontuação de uma transformação de Comparação para uma porta de entrada de Média Ponderada.  
Repita essa etapa para outras transformações de Comparação no mapplet ou mapeamento.
3. Selecione a guia **Portas** na transformação de Média Ponderada.
4. Clique duas vezes no campo **Peso** para cada entrada e insira um valor de peso entre 0,001 e 1. O valor de peso deve refletir a importância relativa da pontuação de entrada quando em comparação com outras entradas na transformação.
5. Salve o mapplet ou mapeamento.

## Exemplo de Pontuações de Correspondência Ponderada

Crie um mapeamento de análise de correspondência para determinar o número de duplicação de nomes de clientes em um banco de dados de clientes. Adicione duas transformações de Comparação para gerar pontuações de correspondência para as colunas `Código Postal` e `Sobrenome` no conjunto de dados.

Muitos registros têm códigos postais correspondentes, mas uma quantidade muito menor de registros tem sobrenomes correspondentes. Para obter a média dessas pontuações de correspondência, você precisa enfatizar a importância de correspondências mais exclusivas.

Para Enfatizar a importância das pontuações de correspondência de sobrenome, aplique um peso mais alto à pontuação de correspondência `Sobrenome`.

Por exemplo, defina o valor **Peso** da entrada de pontuação `Sobrenome` como 0,8 e o valor **Peso** da entrada de pontuação do código de CEP como 0,4.

## Propriedades Avançadas da Transformação de Média Ponderada

Configure propriedades que ajudam a determinar como o Serviço de Integração de Dados processa dados para a transformação de Média Ponderada.

Você pode configurar níveis de rastreamento para logs.

Configure a seguinte propriedade na guia **Avançado**:

### Nível de Rastreamento

Quantidade de detalhes que aparecem no log para esta transformação. Você pode escolher conciso, normal, inicialização detalhada ou dados detalhados. O padrão é normal.

# Transformação de média ponderada em um ambiente não nativo

O processamento da transformação de Média Ponderada em um ambiente não nativo depende do mecanismo que executa a transformação.

Considere o suporte para os seguintes mecanismos de tempo de execução não nativos:

- Mecanismo Blaze. Suportado sem restrições.
- Mecanismo Spark. Com suporte sem restrições em mapeamentos em lote. Sem suporte em mapeamentos de fluxo.
- Mecanismo Databricks Spark. Sem suporte.

# Transformação de Janela

Use a transformação de Janela quando quiser acumular dados transmitidos em grupos de dados e, em seguida, processar os conjuntos de dados. A transformação de Janela é uma transformação passiva.

Ao ler de fontes não delimitadas, convém acumular os dados em grupos de dados delimitados para processamento adicional. Para introduzir intervalos limitados em dados ilimitados, use uma transformação de Janela.

Ao configurar uma transformação de Janela, defina o tipo de janela e os limites de dados por tempo. Para especificar os limites de dados, configure o tamanho da janela e o intervalo de slides da janela. O tamanho da janela define o intervalo de tempo para o qual os dados são acumulados como um grupo de dados. O intervalo do slide define o intervalo de tempo após o qual o grupo de dados acumulados é processado posteriormente. O atraso de marca d'água define o tempo limite para que um evento atrasado seja acumulado em um grupo de dados.

Você pode executar apenas a transformação de Janela para fazer o mapeamento de streaming no mecanismo Spark.

Para obter mais informações sobre a transformação de Janela, consulte o *Guia do Usuário do Data Engineering Streaming*.

# Transformação de Gravação

Este capítulo inclui os seguintes tópicos:

- [Visão Geral da Transformação de Gravação, 724](#)
- [Propriedades da transformação de Gravação , 724](#)
- [Criar uma Transformação de Gravação, 730](#)

## Visão Geral da Transformação de Gravação

Uma transformação de Gravação é uma transformação passiva. O mapeamento usa a transformação de Gravação para gravar dados em um destino. A transformação de Gravação é não reutilizável.

É possível criar uma transformação de Gravação a partir de um objeto de dados físicos, de um objeto de dados lógicos ou de um parâmetro. Se quiser criar uma transformação de Gravação a partir de um objeto de dados físicos importado de uma origem de adaptador do PowerExchange, o editor de mapeamento poderá solicitar que você especifique uma operação de gravação antes de criar uma transformação de Gravação a partir desse objeto de dados.

As propriedades que podem ser configuradas para uma transformação de Gravação dependem do tipo de objeto de dados que você usou para criar essa transformação.

Transformações de Gravação podem representar destinos dinâmicos. É possível configurar uma transformação de Gravação para atualizar dinamicamente suas portas, metadados e outras propriedades. Para obter informações sobre como configurar destinos dinâmicos, consulte o *Guia de Mapeamentos do Informatica Developer*.

## Propriedades da transformação de Gravação

Após criar uma transformação de Gravação, você poderá configurar propriedades para ela.

Configure as propriedades da transformação de Gravação nas guias da exibição **Propriedades** da transformação. As guias que você pode configurar dependem do tipo de destino que a transformação de Gravação representa.



A seguinte tabela descreve cada guia de propriedades e identifica o tipo de destino para o qual você usa a guia:

Guia Propriedade	Descrição	Tipo de Destino
Geral	Especifique as propriedades e o comportamento da transformação. Para fontes de objetos de dados relacionais e personalizados, sincronize portas de entrada de transformação com a origem.	Tudo
Objeto de Dados	Especifique a fonte de dados da transformação. Para origens de objetos de dados relacionais e personalizadas, obter as colunas de objetos de dados da fonte de dados em tempo de execução.	Arquivo simples Relacional Objeto de dados personalizados
Formato	Configurações de entrada para uma fonte de dados de arquivo simples	Arquivo simples
Portas	Definir a porta por objeto de dados associado ou por fluxo de mapeamento.	Tudo
Tempo de execução	Propriedades que o Serviço de Integração de Dados usa ao gravar dados no destino em tempo de execução, por exemplo, para onde enviar arquivos rejeitados. Para um destino de arquivo simples, nomes e diretórios de arquivos rejeitados.	Arquivo simples Relacional
Parâmetros de Objeto de Dados	Exibir os parâmetros de objetos de dados. Configurar valores de parâmetros para o mapeamento ou vincular os parâmetros aos parâmetros de um mapeamento.	Arquivo simples Objeto de dados personalizados Objetos de dados lógicos
Vinculação em tempo de execução	Criar novos links em tempo de execução e exibir as propriedades deles.	Tudo
Avançado	Definir o nível de rastreamento e a ordem das linhas. Para um destino relacional ou Hive, defina a estratégia do esquema de destino.	Arquivo simples Relacional Objetos de dados personalizados Objetos de dados lógicos

## Propriedades Gerais

É possível configurar o nome e a descrição da transformação de Gravação. Você também pode configurar as seguintes propriedades:

### Quando os Metadados da Coluna Forem Alterados

Disponível para destinos relacionais e personalizados. Selecione uma das seguintes opções:

- Sincronizar portas de entrada. A Developer tool atualiza portas de porta de entrada de transformações de Gravação com alterações de metadados que o repositório do Modelo armazena para o objeto de dados.
- Não sincronizar. A Developer tool não mostra alterações de metadados no objeto de dados.

A opção padrão é Sincronizar portas de entrada.

### Objeto de Dados Físicos

Disponível para destinos de arquivo simples e personalizados. O objeto usado para criar a transformação.

Você pode selecionar o nome do objeto de dados e configurar suas propriedades.

## Propriedades do Objeto de Dados

Na guia Objeto de Dados, é possível especificar ou alterar o destino da transformação de Gravação e tornar dinâmicos os destinos de objetos de dados personalizados.

As seguintes propriedades podem ser configuradas:

### Especificar por

Para especificar colunas de destino e metadados para a transformação de Gravação, selecione uma das seguintes opções:

- Valor. A transformação de Gravação usa o objeto de dados associado para especificar metadados e colunas de destino.
- Parâmetro. A transformação de Gravação usa um parâmetro para especificar metadados e colunas de destino.

O padrão é a opção Valor.

### Objeto de Dados

Se você tiver criado a transformação de Gravação a partir de um objeto de dados existente, o campo exibirá o nome do objeto. Clique em **Procurar** para alterar o objeto de dados a ser associado à transformação de Gravação.

### Parâmetro

Escolha ou crie um parâmetro a ser associado à transformação de Gravação.

### Em tempo de execução, obter colunas de objetos da fonte de dados

Ao ativar essa opção, o Serviço de Integração de Dados busca alterações de metadados e definições de dados de tabelas de destino na transformação de Gravação.

Para visualizar como o Serviço de Integração de Dados busca as alterações de metadados e a definição de dados, visualize o mapeamento com parâmetros resolvidos.

## Propriedades de Portas

Na guia Portas, é possível configurar as seguintes propriedades:

### Colunas definidas por

Selecione uma das seguintes opções para definir colunas de transformação de Gravação:

- Objeto de dados associado. Use nomes de colunas, metadados e outras propriedades do objeto de dados na guia Objeto de Dados.
- Fluxo de mapeamento. O mapeamento busca nomes de colunas, metadados e outras propriedades dos objetos em upstream no mapeamento.

A opção padrão é Objeto de dados associado.

### Propriedades de recursos de colunas

Disponível para destinos de arquivo simples e personalizados. O recurso para cada coluna é o objeto de dados a partir do qual a coluna busca seu nome, seus metadados e outras propriedades. Selecione o nome do recurso para alterar suas propriedades.

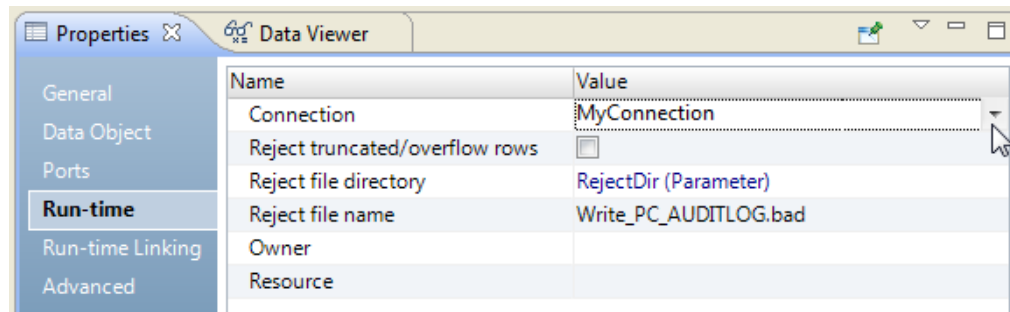
## Propriedades de Tempo de Execução

É possível configurar as seguintes propriedades da transação de Gravação na guia Tempo de execução:

### Conexão

Disponível para destinos relacionais. Conexão usada pela transformação. Clique no lado direito do campo para alterar a conexão.

A seguinte imagem mostra a localização do botão suspenso a ser clicado:



### Rejeitar linhas truncadas/de estouro

Disponível para destinos relacionais e personalizados.

A Developer tool permite converter dados transmitindo-os de porta em porta. Às vezes, uma conversão provoca um estouro de dados numéricos ou o truncamento de cadeias em colunas que contêm caracteres. Por exemplo, transmitir dados de uma porta Decimal (28, 2) para uma porta Decimal (19, 2) causa um estouro numérico. Da mesma forma, se você transmitir dados de uma porta String(28) para uma porta String(10), o Serviço de Integração de Dados truncará as cadeias em 10 caracteres.

Quando uma conversão causa um estouro, o Serviço de Integração de Dados, por padrão, ignora a linha. O Serviço de Integração de Dados não grava a linha no arquivo rejeitado. Para cadeias, o Serviço de Integração de Dados trunca a cadeia e a transmite para a próxima transformação.

Selecione essa opção para incluir todos os dados truncados e de estouro entre a última transformação e o destino no arquivo rejeitado da sessão. O Serviço de Integração de Dados envia todas as linhas truncadas e qualquer linha de estouro para o arquivo rejeitado da sessão ou para os logs de erros de linha, dependendo de como a sessão está configurada.

### Diretório do arquivo rejeitado

Diretório no qual o arquivo rejeitado existe. O padrão é o parâmetro do sistema RejectDir.

### Nome do arquivo rejeitado

Nome do arquivo rejeitado. O padrão é <output\_file\_name>.bad.

Se várias partições gravarem no destino de arquivo simples, cada uma delas gravará em um arquivo rejeitado separado denominado <output\_file\_name><partition\_number>.bad.

## Propriedades de Vinculação em Tempo de Execução

Crie e configure um link em tempo de execução na guia **Vinculação em Tempo de Execução**. Um link em tempo de execução é um link de grupo para grupo entre transformações que usa um parâmetro, uma política de link ou ambos para determinar quais portas devem ser vinculadas em tempo de execução. Links em tempo de execução aparecem como linhas grossas no editor de mapeamentos.

Crie e configure links em tempo de execução para uma transformação de Gravação nos seguintes casos:

- O objeto de dados de destino na transformação de Gravação usa um parâmetro.

- Portas da transformação em upstream podem mudar em tempo de execução.

**Nota:** Não crie um link em tempo de execução para uma transformação de Gravação ao definir as colunas de destino com base no fluxo de mapeamento.

É possível realizar as seguintes tarefas na guia **Vinculação em Tempo de Execução**:

#### **Criar um link em tempo de execução.**

Na área **Links**, clique no botão **Novo** e selecione a transformação a partir da qual você deseja vincular as portas à transformação de Gravação em tempo de execução na caixa de diálogo Novo Link.

#### **Configurar as propriedades do link em tempo de execução.**

Na área **Propriedades do Link**, configure as seguintes propriedades de link em tempo de execução:

##### **Parâmetro**

Selecione essa opção quando os nomes de portas puderem mudar entre execuções de mapeamento e você conhecer os valores desses nomes. Use um parâmetro do tipo Conjunto de Links de Entrada para conectar portas por valores de nome entre execuções de mapeamento. A sintaxe do parâmetro de mapeamento Conjunto de Links de Entrada consiste em pares de portas separados por ponto-e-vírgula: Afield1->Bfield2, Afield3->Bfield4 .

##### **Diretiva de Link**

Selecione essa opção quando quiser vincular automaticamente as portas que possuem nomes correspondentes. Por exemplo, quando ambos os objetos de mapeamento contêm uma porta denominada SALARY, o Serviço de Integração de Dados os vincula. É possível ignorar prefixos e sufixos nos nomes de portas.

Em tempo de execução, o Serviço de Integração de Dados estabelece e resolve links entre as portas na seguinte ordem:

- Links que você cria manualmente no editor de mapeamento.
- Links com base no parâmetro que você configurou para um link em tempo de execução.
- Links com base na diretiva de link que você configurou para um link em tempo de execução.

Para obter mais informações sobre links em tempo de execução, consulte o *Guia de Mapeamentos do Informatica Developer*.

## Propriedades Avançadas

Configure propriedades avançadas para determinar como o Serviço de Integração de Dados processa dados para a transformação de Gravação.

Configure as seguintes propriedades na guia **Avançado**:

##### **Nível de rastreamento**

Controla a quantidade de detalhes no arquivo de log de mapeamento.

##### **Tipo de carregamento de destino**

Tipo de carregamento de destino. Selecione Normal ou Em Massa. É possível definir o tipo de carregamento de destino para recursos relacionais ou objetos de dados personalizados.

Se você selecionar Normal, o Serviço de Integração de Dados carregará os destinos normalmente. Escolha Em Massa quando carregar no DB2, no Sybase, no Oracle ou no Microsoft SQL Server. Se você especificar Em Massa para outros tipos de bancos de dados, o Serviço de Integração de Dados reverterá para um carregamento normal. O carregamento em massa pode aumentar o desempenho do mapeamento, mas limita a capacidade de recuperação, pois o registro no banco de dados não ocorre.

Ao gravar em um destino Oracle com carregamento em massa, você pode otimizar o desempenho desativando restrições no banco de dados Oracle.

Escolha o modo Normal se o mapeamento contiver uma transformação Atualizar Estratégia. Se você selecionar Normal e o nome de destino do Microsoft SQL Server incluir espaços, configure o seguinte SQL de ambiente no objeto de conexão:

```
SET QUOTED_IDENTIFIER ON
```

### **Substituição de atualização**

Substitui a instrução UPDATE padrão no destino.

### **Excluir**

Exclui todas as linhas sinalizadas para exclusão.

Por padrão, essa opção fica ativada.

### **Inserir**

Insere todas as linhas sinalizadas para inserção.

Por padrão, essa opção fica ativada.

### **Estratégia de Esquema de Destino**

Tipo de estratégia de esquema de destino para a tabela de destino relacional ou Hive.

Você pode selecionar uma das seguintes estratégias de esquema de destino:

- **RETAIN** - Retenha o esquema de destino existente. O Serviço de Integração de Dados retém o esquema de destino existente.
- **CREATE** - Crie ou substitua a tabela em tempo de execução. O Serviço de Integração de Dados descarta a tabela de destino em tempo de execução e a substitui por uma tabela baseada em uma tabela de destino que você identifica.
- **Atribuir Parâmetro**. Você pode atribuir um parâmetro para representar o valor da estratégia do esquema de destino e, em seguida, alterar o parâmetro no tempo de execução.

### **Consulta DDL para criação ou substituição**

Cria ou substitui a tabela de destino no tempo de execução com base em uma consulta DDL que você define. Aplicável quando você seleciona a opção de estratégia de esquema de destino **CREATE - Crie ou substitua a tabela no tempo de execução**.

### **Truncar tabela de destino**

Trunca o destino antes de carregar os dados.

Por padrão, essa opção fica ativada.

### **Truncar a partição de destino**

Trunca um destino Hive particionado interno ou externo antes de carregar dados. Você deve escolher **Truncar a tabela de destino** antes de escolher essa opção.

Por padrão, essa opção fica desativada.

### **Estratégia de atualização**

A estratégia de atualização das linhas existentes. Você pode selecionar uma das seguintes estratégias:

- **Atualização como atualização**. O Serviço de Integração de Dados atualiza todas as linhas sinalizadas para atualização.
- **Atualização como inserção**. O Serviço de Integração de Dados insere todas as linhas sinalizadas para atualização. Você também deve selecionar a opção de destino **Inserir**.

- Atualizar e inserir. O Serviço de Integração de Dados atualizará as linhas sinalizadas para atualização se elas existirem no destino e, em seguida, inserirá as linhas restantes marcadas para inserção. Você também deve selecionar a opção de destino **Inserir**.

#### **PreSQL**

O comando SQL que o Serviço de Integração de Dados executa no banco de dados de destino antes de ler a origem.

A Developer tool não valida o SQL.

#### **PostSQL**

Comando SQL que o Serviço de Integração de Dados executa no banco de dados de destino depois de gravar no destino.

A Developer tool não valida o SQL.

#### **Manter ordem das linhas**

Mantenha a ordem das linhas dos dados de entrada para o destino. Selecione essa opção se o Serviço de Integração de Dados não realizar qualquer otimização que possa alterar a ordem de linhas.

Quando o Serviço de Integração de Dados realiza otimizações, ele pode perder a ordem das linhas anteriormente estabelecida no mapeamento. Você pode estabelecer a ordem das linhas em um mapeamento com uma origem de arquivo simples classificada, uma origem relacional classificada ou uma transformação do Classificador. Quando você configura um destino para manter a ordem das linhas, o Serviço de Integração de Dados não realiza otimizações para esse destino.

#### **Restrições**

Instruções SQL para restrições de integridade referencial em nível de tabela. Aplicável somente a destinos relacionais.

## Criar uma Transformação de Gravação

Ao criar uma transformação de Gravação, você escolhe um dos seguintes métodos com base no recurso a partir do qual a transformação é criada:

#### **Crie a transformação de um objeto de dados no repositório do Modelo.**

Use esse método se quiser basear os metadados da transformação de Gravação em um objeto de dados.

Para criar uma transformação de Gravação no editor de mapeamentos, use o assistente de **Nova Transformação de Gravação** ou arraste um objeto de dados da exibição **Object Explorer** e selecione Gravação como o tipo de transformação.

#### **Crie a transformação a partir do fluxo de objetos de mapeamento.**

Use esse método para basear as portas da transformação de Gravação no fluxo de metadados proveniente de transformações de mapeamento em upstream.

#### **Crie a transformação usando um parâmetro.**

Use esse método se quiser que a transformação de Gravação herde portas do objeto que o parâmetro representa.

#### **Crie a transformação usando outra transformação no mapeamento.**

Use esse método se quiser que a transformação de Gravação tenha as mesmas portas que a transformação de origem.

## Criando uma Transformação de Gravação a partir de um Objeto de Dados

Você pode criar e configurar uma transformação de Gravação no editor de mapeamento.

Convém usar esse método se você quiser configurar propriedades específicas para a transformação. Por exemplo, é possível criar a transformação e depois configurá-la para obter metadados de coluna da fonte de dados em tempo de execução.

1. Clique com o botão direito do mouse no editor de mapeamento e selecione **Adicionar Transformação**.  
A caixa de diálogo **Adicionar Transformação** é aberta.
2. Selecione **Gravar** e clique em **OK**.  
O assistente de **Nova Transformação de Gravação** é aberto.
3. Selecione **Objeto de Dados Físicos** ou **Objeto de Dados Lógicos** e clique em **Procurar**.  
A caixa de diálogo **Selecionar Objeto de Dados** é aberta.
4. Selecione uma fonte de dados e clique em **OK**.
5. Para definir colunas de transformação de Gravação pelo objeto de dados associado, selecione **Objeto de dados associado**.  
As portas da transformação de Gravação são iguais às colunas do objeto de dados associado.
6. Para atualizar dinamicamente colunas de objetos de destino em tempo de execução com alterações do arquivo de destino, selecione **Em tempo de execução, obter colunas de objetos de dados da fonte de dados**.  
O Serviço de Integração de Dados atualiza os metadados de colunas da transformação de Gravação quando o mapeamento é executado.
7. Clique em **Concluir**.

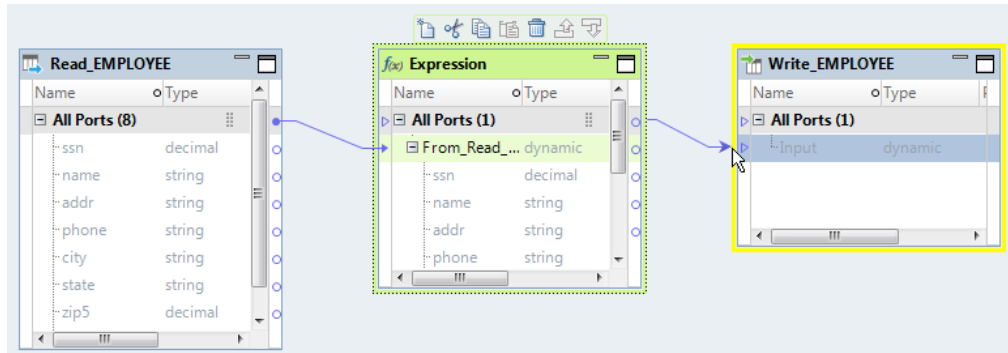
## Criando uma Transformação de Gravação de um Fluxo de Mapeamento

Você pode criar e configurar uma transformação de Gravação no editor de mapeamento.

Convém usar esse método se você quiser configurar propriedades específicas para a transformação. Por exemplo, é possível criar a transformação e depois configurá-la para definir colunas com base no fluxo de mapeamento e obter metadados de colunas da fonte de dados em tempo de execução.

1. Clique com o botão direito do mouse no editor de mapeamento e selecione **Adicionar Transformação**.  
A caixa de diálogo **Adicionar Transformação** é aberta.
2. Selecione **Gravar** e clique em **OK**.  
O assistente de **Nova Transformação de Gravação** é aberto.
3. Selecione **Objeto de Dados Físicos** ou **Objeto de Dados Lógicos** e clique em **Procurar**.  
A caixa de diálogo **Selecionar Objeto de Dados** é aberta.
4. Selecione uma fonte de dados e clique em **OK**.
5. Selecione **Fluxo de mapeamento** e depois clique em **Concluir**.
6. Arraste portas da porta Todas as Portas do upstream até a porta de **Entrada** da transformação de Gravação.

A seguinte imagem mostra como conectar portas em upstream à porta de **Entrada** da transformação de Gravação:



A transformação de Gravação recebe definições de colunas de objetos de mapeamento em upstream.

7. Para atualizar dinamicamente colunas de objetos de destino em tempo de execução com alterações do arquivo de destino, selecione **Em tempo de execução, obter colunas de objetos de dados da fonte de dados**.

O Serviço de Integração de Dados atualiza os metadados de colunas da transformação de Gravação quando o mapeamento é executado.

## Criando uma Transformação de Gravação de um Parâmetro

Um parâmetro é um valor constante que você pode alterar entre execuções de mapeamentos.

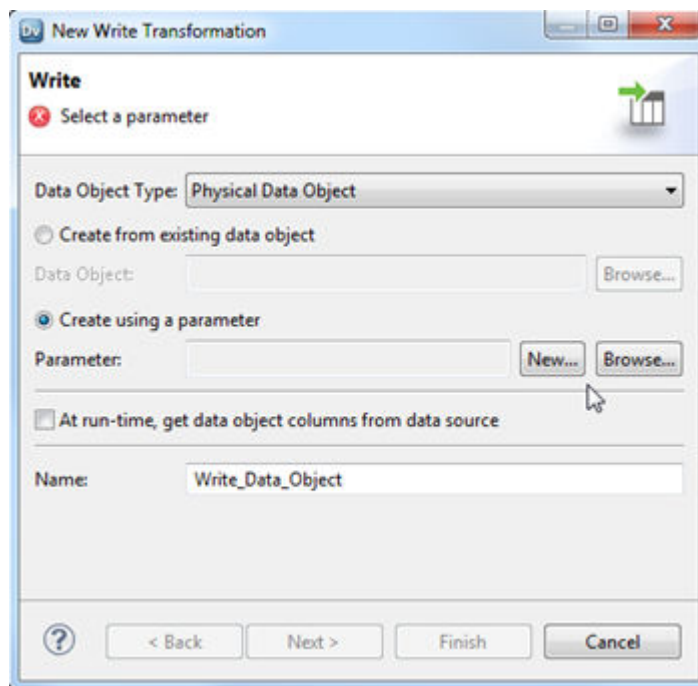
Em um mapeamento dinâmico, é possível usar parâmetros para alterar origens e destinos. Também é possível usar parâmetros para regras de entrada, regras de seleção, links em tempo de execução e propriedades de transformação. Quando você altera o valor do parâmetro, o Serviço de Integração de Dados cria ou recria o destino de acordo com o valor especificado no parâmetro.

1. Clique com o botão direito do mouse no editor e selecione **Adicionar Transformação**.
2. Selecione Gravar na lista de transformações e depois clique em **OK**.

**Sugestão:** Digite a primeira letra da transformação para filtrar a lista.

A caixa de diálogo Nova Transformação de Gravação é aberta.





3. Selecione **Criar usando um parâmetro**.
4. Navegue para selecionar o parâmetro ou clique em **Novo** para criar um novo parâmetro e, em seguida, navegue para selecionar um objeto de dados a partir do qual criar o parâmetro.
5. Opcionalmente, selecione a opção **Em tempo de execução, obter colunas de objetos da fonte de dados** para atualizar as colunas de transformação em tempo de execução.
6. Opcionalmente, clique em **Avançar** para escolher como definir as colunas da transformação.
7. Clique em **Concluir**.

## Criando uma Transformação de Gravação de uma Transformação Existente

É possível criar uma transformação de Gravação com as mesmas portas em uma transformação existente no mapeamento.

Você pode criar um destino para arquivos complexos, arquivos simples ou recursos relacionais. Se uma das portas na transformação existente contiver uma porta complexa, você deverá criar um destino de arquivo complexo e vincular portas por nome ou vincular portas em tempo de execução com base na diretiva de link. A Developer tool pode criar um destino de arquivo complexo Avro, Parquet, ORC ou JSON.

1. Abra um mapeamento no editor.
2. Clique com o botão direito do mouse no editor de mapeamento e selecione **Criar Destino**.  
A janela **Criar Destino** é aberta.
3. Escolha um arquivo complexo, um arquivo simples ou um tipo de objeto de dados relacionais.
4. Escolha um tipo de link.

Você pode escolher um dos seguintes tipos de link:

### Vincular portas por nome

Portas na transformação de Gravação correspondem às portas na origem e têm os mesmos nomes.

### Vincular porta dinâmica com base no fluxo de mapeamento

A transformação de Gravação contém portas dinâmicas com base em objetos upstream no fluxo de mapeamento.

### Vincular portas em tempo de execução com base na diretiva de link

Portas são criadas no destino em tempo de execução com base na diretiva de link que você configura na guia Vinculação em Tempo de Execução da transformação de Gravação.

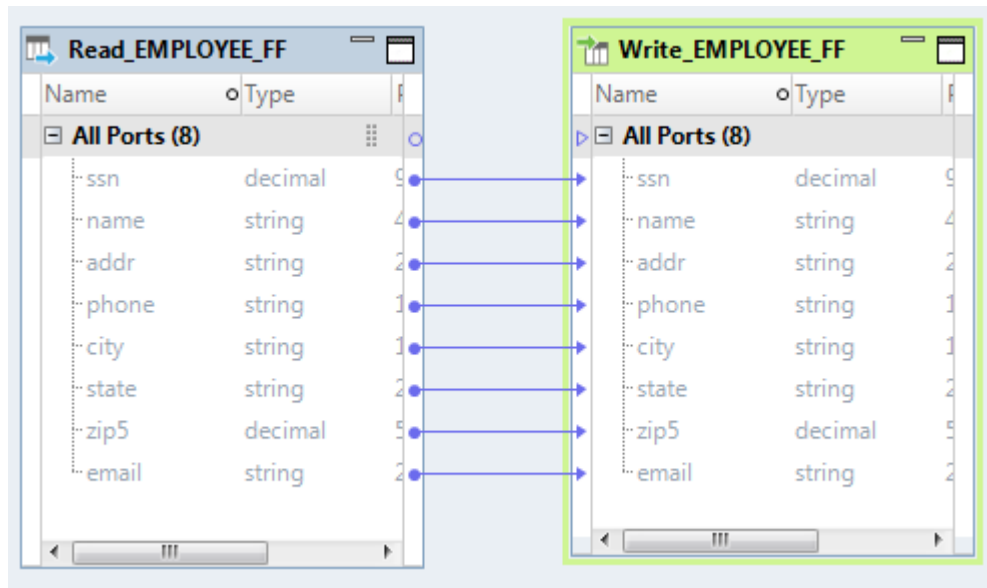
Para obter mais informações sobre portas dinâmicas e configuração de link de tempo de execução, consulte o *Guia de Mapeamento do Informatica Developer*.

5. Nomeie o novo objeto de dados.
6. Opcionalmente, clique em **Procurar** para selecionar uma localização para o objeto de dados.
7. Se você optou por criar um destino de arquivo complexo, selecione o formato de arquivo complexo como Avro ou Parquet na caixa suspensa **Formato do Recurso**.
8. Clique em **Concluir**.

A Developer tool realiza as seguintes tarefas:

- Adiciona uma transformação de Gravação ao mapeamento.

A imagem a seguir mostra uma transformação de Gravação criada a partir de uma transformação de Leitura:



- Vincula portas.
- Cria um objeto de dados físicos.  
Você pode configurar as propriedades do objeto de dados físicos. Por exemplo, você deve especificar uma conexão HDFS para o objeto de dados de arquivo complexo.

## APÊNDICE A

# Delimitadores de Transformação

Este apêndice inclui os seguintes tópicos:

- [Delimitadores de Transformação - Visão Geral, 735](#)

## Delimitadores de Transformação - Visão Geral

Os delimitadores de transformação especificam divisões entre as strings de dados.

A seguinte tabela lista os delimitadores usados pelas transformações para analisar e gravar strings de dados:

Nome do Delimitador	Símbolo do Delimitador
arroba	@
vírgula	,
traço	-
aspas duplas	"
barra	/
ponto	.
cerquilha	#
barra vertical	
ponto e vírgula	;
aspas simples	'
espaço	[Barra de espaços]
tabulação	[Tecla Tab]
sublinhado	_

# ÍNDICE

## A

- ação SOAP
  - substituindo na transformação de Consumidor de Serviço da Web [684](#)
- Analizador
  - descrição [233](#)
- análise de correspondência de campos
  - definição de análise de campos [461](#)
  - fluxo de processo [486](#)
- análise de correspondência de identidade
  - armazenamento persistente de dados de índice [469](#)
  - códigos de status de persistência [478](#), [479](#)
  - conjunto de dados mestre [469](#)
  - definição de análise de identidade [461](#)
  - descrições de status de persistência [478](#), [479](#)
  - fluxo de processo [499](#)
  - propriedades de saída para a análise de dados de índice [508](#)
  - regras e diretrizes para dados de índice persistentes [469](#)
- análise de correspondência de origem dupla [461](#)
- análise de correspondência de origem única [461](#)
- análise de mensagens SOAP
  - descrição [697](#)
  - elemento de união [705](#)
  - elementos QName [703](#)
  - saída desnormalizada [700](#)
  - saída dinamizada [701](#)
  - saída normalizada [699](#)
  - tipos derivados [702](#)
- anyType
  - mapear portas [716](#)
- armazenamento persistente de dados de índice de identidade [469](#)
- arquivo de entrada de exemplo
  - transformação do Processador de Dados [235](#)
- arquivo de origem de amostra
  - definindo na transformação do Processador de Dados [251](#)
- arquivo WSDL
  - elemento de associação [673](#)
  - elemento de operação [673](#)
  - elemento de porta [673](#)
  - elemento de serviço [673](#)
- arquivos de cache
  - diretório [72](#)
  - visão geral [71](#)
- associação
  - elemento de arquivo WSDL [673](#)
- atribuindo portas
  - problemas de qualidade [161](#)
- atributos anyAttribute
  - Transformação de Consumidor de Serviço da Web [678](#), [681](#)
- atributos de porta
  - propagando [63](#)
- Atualizar ou Inserir (propriedade)
  - descrição [453](#)
- autenticação por cookies
  - Transformação de Consumidor de Serviço da Web [677](#)

- autenticação por cookies ()
- transformação de Consumidor de Serviço da Web REST [590](#)

## B

- Biblioteca
  - criando na transformação do Processador de Dados [252](#)
- bloqueando transformações
  - descrição [41](#)

## C

- cabeçalho HTTP
  - adicionando à transformação de Consumidor de Serviço da Web [676](#)
  - adicionando à transformação de Consumidor de Serviço da Web REST [589](#)
- cache
  - transformação de Pesquisa [415](#)
- cache de pesquisa
  - visão geral [415](#)
- cache de pesquisa compartilhado
  - diretrizes [443](#)
  - particionamento, diretrizes [443](#)
- cache de pesquisa dinâmica
  - casos de uso [446](#)
  - descrição [445](#)
  - substituição SQL de pesquisa [451](#)
  - usando origens de arquivo simples [445](#)
- cache de pesquisa estático
  - visão geral [440](#)
- cache de pesquisa persistente
  - visão geral [441](#)
- caches
  - arquivos de cache [70](#)
  - cache de pesquisa dinâmica [445](#)
  - cache de pesquisa estático [440](#)
  - dados [71](#)
  - diretório para arquivos [72](#)
  - índice [71](#)
  - otimização do tamanho [75](#)
  - particionamento [75](#)
  - tamanho [72](#)
  - tamanho aumentado pelo Serviço de Integração de Dados [74](#)
  - tipos [71](#)
  - transformação de Pesquisa [438](#)
  - transformações [70](#)
  - visão geral [70](#)
- caches de dados
  - transformações [71](#)
- caches de índice
  - transformações [71](#)
- caches de pesquisa
  - definição [438](#)

- caches de pesquisa ()
  - dinâmico [445](#)
  - estático [440](#)
  - persistente [441](#)
  - visão geral [438](#)
- Cadastros de Pessoas Físicas
  - máscara de código de área [223](#)
  - maskamento de dados repetíveis [223](#)
- caixa de diálogo Editar Grupo de Saída
  - transformação de Normalizador [534](#)
- cálculos
  - usando variáveis com [47](#)
- campos de ocorrência múltipla
  - transformação de Normalizador [521](#)
- caracteres da string de origem
  - transformação de Mascaramento de Dados [217](#)
- caracteres de substituição da string de resultados
  - transformação de Mascaramento de Dados [218](#)
- chave de criptografia de armazenamento
  - transformação de Mascaramento de Dados [226](#)
- chaves
  - criando com a transformação Gerador de Sequência [608](#)
  - hierarquia da mensagem SOAP [709](#)
- chaves compostas
  - Transformação de Consumidor de Serviço da Web [678](#)
  - transformação de Consumidor de Serviço da Web REST [591](#)
- Chaves de classificação
  - configurando [620](#)
- chaves de composição
  - criando com a transformação Gerador de Sequência [608](#)
- chaves externas
  - criando com a transformação Gerador de Sequência [608](#)
- chaves geradas
  - grupos de saída de serviços da Web [699](#)
- chaves primárias
  - criando com a transformação Gerador de Sequência [608](#)
- ciclo
  - Propriedade da Transformação de Gerador de Sequência [610](#)
- Ciclo (propriedade)
  - Propriedade da Transformação de Gerador de Sequência [610](#)
- Classe EDataType
  - expressões Java [354](#)
- Classe JExpression
  - expressões Java [356](#), [357](#)
- Classe JExprParaMetadata
  - expressões Java [354](#)
- classificação
  - grupos de dados [560](#)
  - valores de string [558](#)
- Classificador
  - tamanho variável [624](#)
- classpath
  - propriedade de mapeamento [318](#)
- clusters
  - Transformação de Exceção de Registro Duplicado [266](#)
- código de resposta
  - transformação de Consumidor de Serviço da Web REST [590](#)
- código de resposta HTTP
  - transformação de Consumidor de Serviço da Web REST [590](#)
- código estático
  - Transformações Java [321](#)
- código Java
  - em transformações Java [318](#)
  - localizando erros [330](#)
- coluna dependente
  - maskamento de dados [213](#)
- coluna Localização
  - transformação de serviço da Web [708](#)
- compactação SOAP
  - Transformação de Consumidor de Serviço da Web [687](#)
- compilando
  - Transformações Java [329](#)
- compilando código Java
  - guia Código Completo [325](#)
- componente de inicialização
  - transformação do Processador de Dados [237](#)
- condição de associação
  - visão geral [368](#)
- condição de pesquisa
  - descrição [414](#)
  - especificar por parâmetro [425](#)
  - transformação de Mascaramento de Dados [212](#)
  - usando portas geradas [422](#)
- condição do filtro de grupo
  - transformação de Roteador [601](#)
- condições
  - transformação de Associador [368](#)
  - transformação de Roteador [601](#)
- conexão
  - serviços da Web [684](#)
  - serviços da Web REST [595](#)
- conexão HTTP
  - serviços da Web REST [595](#)
- conexões de serviços da Web
  - visão geral [684](#)
- configurações da porta de saída
  - descrição [291](#)
- configurações de codificação
  - transformação do Processador de Dados [239](#)
- configurações de controle de saída
  - transformação do Processador de Dados [241](#)
- conjuntos de resultados
  - posicionamento de parâmetros de saída [653](#)
  - procedimentos armazenados de amostra [651](#)
  - transformação SQL [651](#)
- considerações
  - transformações Java [314](#)
- consolidação automática
  - Transformação de Exceção de Registro Duplicado [267](#)
- consolidação manual
  - Transformação de Exceção de Registro Duplicado [267](#)
- constantes
  - substituindo valores nulos por [51](#)
- consulta de pesquisa
  - consulta padrão [409](#)
  - descrição [409](#)
  - diretrizes de substituição [411](#)
  - ORDER BY [409](#)
  - palavras reservadas [411](#)
  - substituindo [409](#), [411](#)
- consulta SQL
  - Transformação SQL [638](#)
- consultas
  - substituindo a pesquisa [409](#)
  - Transformação Pesquisa [409](#)
- contagem de erros
  - incrementando para transformações Java [341](#)
- Contagem Máx. de Linhas de Saída
  - Transformação SQL [639](#)
- Contagem Máxima de Linhas de Saída
  - Transformação SQL [641](#)
- Continuar com Erro de SQL
  - Transformação SQL [639](#), [643](#)
- criando
  - Objeto de Dados de Sequência [612](#)
  - transformação de Gerador de Sequência [615](#)

- criando ()
  - transformações Java [314, 328](#)
- criptografar o armazenamento
  - transformação de Mascaramento de Dados [226](#)
- critérios de seleção
  - seletor de portas [283, 423](#)

## D

- dados
  - armazenamento temporário [46](#)
- Dados de Cluster
  - grupo de saída [270](#)
- dados dinamizados
  - mensagens SOAP [713](#)
- Data Integration Service)
  - modo de reinicialização [345](#)
- defineJExpression
  - Método de API de transformação Java [355](#)
- definição de janelas
  - propriedades [286](#)
- dependências
  - caminho do link [63](#)
  - implícitas [64](#)
- desempenho
  - usando variáveis para melhorar [46](#)
- destinos
  - relacional [665](#)
- dicionário
  - mascaramento de dados de substituição [210](#)
  - mascaramento repetível de expressão [206](#)
- dígito de início
  - CPF [224](#)
- diretório de arquivo de cache
  - transformação de Associação [151](#)
  - transformação de Consolidação [188](#)
  - Transformação de Exceção de Registro Duplicado [271](#)
  - transformação de Gerador de Chaves [390](#)
- diretório de cache
  - transformação de Mascaramento de Dados [226](#)
- diretrizes de codificação
  - transformação do Processador de Dados [241](#)

## E

- Editor de Expressão
  - descrição [44](#)
  - testando expressões [281](#)
  - validando expressões [44](#)
- efeitos colaterais
  - Transformação de Consumidor de Serviço da Web [689](#)
  - transformação SQL [644](#)
- elemento de união
  - analisando mensagens SOAP [705](#)
- elementos
  - união [718](#)
- elementos anyType
  - analisando [701](#)
  - Transformação de Consumidor de Serviço da Web [678, 681](#)
- elementos de escolha
  - exibindo na transformação de Consumidor de Serviço da Web [679](#)
  - exibindo na transformação de Consumidor de Serviço da Web REST [594](#)
- elementos de lista
  - analisando mensagens SOAP [704](#)
  - descrição [718](#)

- elementos de opção
  - analisando mensagens SOAP [704](#)
  - descrição [717](#)
  - exibindo na transformação de Consumidor de Serviço da Web [682](#)
- elementos de tipo derivado
  - Transformação de Consumidor de Serviço da Web [678, 681](#)
- elementos de união
  - descrição [718](#)
- elementos QName
  - analisando mensagens SOAP [703](#)
- embaçando
  - valores de data [219](#)
  - valores numéricos [219](#)
- Encaminhando Linhas Rejeitadas
  - configurando [667](#)
  - opção [667](#)
- entrada desnormalizada
  - portas de serviço da Web [714](#)
- erros
  - aumentando o limite em transformações Java [341](#)
  - manipulando [53](#)
- erros de código de não usuário
  - em transformações Java [331](#)
- erros de código do usuário
  - Transformações Java [330](#)
- erros de compilação
  - Identificando a origem da transformação Java [330](#)
- escopo
  - mestre e detalhado [366](#)
  - seletor de portas [283, 423](#)
- Escopo de Transformação
  - transformações Java [315](#)
- esquemas
  - Transformação de Hierárquica para Relacional [307](#)
  - Transformação de Relacional para Hierárquica [579](#)
  - transformação do Processador de Dados [238](#)
- estratégia de correspondência
  - Transformação de Exceção de Registro Duplicado [273](#)
- evento de aviso
  - transformação do Processador de Dados [246](#)
- evento de erro fatal
  - transformação do Processador de Dados [246](#)
- evento de falha
  - transformação do Processador de Dados [246](#)
- evento de falha opcional
  - transformação do Processador de Dados [246](#)
- evento de notificação
  - transformação do Processador de Dados [246](#)
- eventos
  - transformação do Processador de Dados [246](#)
- Excel
  - configurando portas de transformação [67](#)
  - copiando para a Developer tool [67](#)
  - copiando portas [66](#)
  - editando uma transformação [68](#)
  - copiando regras e diretrizes [69](#)
- exemplo
  - expressão dinâmica [290](#)
  - chaves de partição e ordem [288](#)
- Exemplo de Origem
  - criando na transformação do Processador de Dados [253](#)
- Exibição Ajuda de Scripts
  - transformação do Processador de Dados [234](#)
- exibição Componente
  - transformação do Processador de Dados [234](#)
- Exibição de Atribuição de Problema
  - Transformação de Exceção de Registro Inválido [160](#)

- Exibição de configuração
  - Exemplo de transformação de Exceção de Registro Inválido [164](#)
  - Transformação de Exceção de Registro Duplicado [267](#)
  - Transformação de Exceção de Registro Inválido [158](#)
- Exibição de Configurações
  - transformação do Processador de Dados [239](#)
- Exibição de Eventos do Processador de Dados
  - exibindo log de eventos [249](#)
  - transformação do Processador de Dados [247](#)
- Exibição de eventos, Processador de Dados
  - exibindo log de eventos [249](#)
- Exibição de referência
  - Transformação de Hierárquica para Relacional [307](#)
  - Transformação de Relacional para Hierárquica [579](#)
- Exibição de Referência
  - transformação do Processador de Dados [238](#)
- Exibição de tempo de execução
  - Transformação SQL [648](#)
- exibição do Normalizador
  - descrição [523](#)
- exibição Eventos
  - transformação do Processador de Dados [247](#)
- Exibição Origem Hexadecimal do Processador de Dados
  - descrição [234](#)
- exibição Origem Hexadecimal, Processador de Dados
  - descrição [234](#)
- exibição Portas
  - saída de transformação SQL [632](#)
- Exibição Repositório
  - transformação do Processador de Dados [234](#)
- expressão dinâmica
  - criando [292](#)
- expressões
  - adicionando a uma porta [44](#)
  - adicionando comentários [44](#)
  - em transformações [42](#)
  - inserindo [44](#)
  - simplificando [46](#)
  - Transformações Java [348](#)
  - validando [44](#)
- expressões dinâmicas
  - exemplo [290](#)
  - configurações da porta de saída [291](#)
  - visão geral [290](#)
- expressões Java
  - chamando com interface simples [351](#)
  - Classe EDataType [354](#)
  - Classe JExpression [356](#), [357](#)
  - Classe JExprParaMetadata [354](#)
  - configurando [349](#)
  - configurando funções [350](#)
  - criando [350](#)
  - criando na caixa de diálogo Definir Função [350](#)
  - exemplo de interface avançada [356](#)
  - exemplo de interface simples [352](#)
  - gerando [349](#)
  - gerar código Java [350](#)
  - interface avançada [353](#)
  - interface simples [351](#)
  - invocando [342](#)
  - invocando com interface avançada [353](#)
  - método API invokeJExpression [351](#)
  - regras e diretrizes [351](#), [353](#)
  - regras e diretrizes para invocar [342](#)
  - Transformações Java [348](#)

## F

- falha em mapeamentos
  - Transformações Java [339](#)
- falhas SOAP genéricas
  - Transformação de Consumidor de Serviço da Web [687](#)
- filtragem das linhas de origem
  - transformação de Pesquisa [412](#)
- filtro da origem de pesquisa
  - limitação de pesquisas [412](#)
- formato de máscara
  - formatos especiais de máscara [220](#)
  - maskamento de valores de string [216](#)
- função ABORT
  - usando [51](#)
- função ERROR
  - usando [51](#)
- função IIF
  - substituindo chaves ausentes com a transformação Gerador de Sequência [609](#)

## G

- GCID
  - descrição [521](#)
  - exemplo de transformação de Normalizador [532](#)
- gerar código Java
  - expressões Java [350](#)
- gerar linha de saída
  - Transformações Java [339](#)
- gerar tabela de registros inválidos
  - Transformação de Exceção de Registro Inválido [160](#)
- Grupo de Saída Padrão
  - Transformação de Exceção de Registro Duplicado [270](#)
- grupo de sequência
  - exibindo na transformação de Consumidor de Serviço da Web REST [594](#)
- grupo definido pelo usuário
  - transformação de Roteador [601](#)
- grupo todos
  - exibindo na transformação de Consumidor de Serviço da Web [679](#), [682](#)
  - exibindo na transformação de Consumidor de Serviço da Web REST [594](#)
- grupos
  - adicionando à transformação de Roteador [604](#)
  - definido pelo usuário [601](#)
  - transformação de Roteador [601](#)
- grupos de saída
  - Transformação de Exceção de Registro Inválido [153](#)
  - transformação de Normalizador [532](#)
- grupos de substituição
  - analisando mensagens SOAP [704](#)
  - serviços da Web [717](#)
  - Transformação de Consumidor de Serviço da Web [678](#), [681](#)
- grupos na análise de correspondência [462](#)
- grupos padrão
  - transformação de Roteador [601](#)
- guia Agrupar por
  - descrição [139](#)
  - parâmetros de lista de portas [140](#), [561](#)
- guia Auxiliares
  - transformações Java [323](#)
- Guia Auxiliares
  - Transformações Java [321](#)
- Guia Classificação
  - Transformação do Classificador [620](#)

- guia Código Completo
  - erros de compilação Java [330](#)
  - transformações Java [325](#)
- guia Funções
  - transformações Java [324](#)
- guia Importações
  - transformações Java [323](#)
- Guia Importações
  - Transformações Java [321](#)
- guia Na Entrada
  - transformações Java [323](#)
- guia No Final
  - transformações Java [324](#)
- guia Objeto de Dados
  - descrição [570](#)
  - descrições dos campos [417](#)
  - parametrizar usando um novo objeto de dados [420](#)
- GZip
  - compactando mensagens SOAP [687](#)

## H

- hierarquia de entrada
  - Transformação do Normalizador [523](#)
- hierarquia SOAP
  - relacionamento com portas de entrada [708](#)

## I

- incrementando
  - configuração do intervalo da sequência [611](#)
- Incremento De (propriedade)
  - Propriedade da Transformação de Gerador de Sequência [610](#)
- informações do dicionário
  - transformação de Mascaramento de Dados [212](#)
- inicializando
  - variáveis [49](#)
- Inserir ou Atualizar (propriedade)
  - descrição [452](#)
- interface avançada
  - chamando expressões Java [353](#)
  - Classe EDataType [354](#)
  - Classe JExpression [356](#), [357](#)
  - Classe JExprParaMetadata [354](#)
  - exemplo [356](#)
  - expressões Java [353](#)
- interface simples
  - exemplo [352](#)
  - expressões Java [351](#)
  - Métodos API de transformação Java [351](#)
- intervalo
  - mascaramento de valores numéricos [218](#)
- intervalo de confirmação de armazenamento
  - transformação de Mascaramento de Dados [226](#)
- invoke
  - Método de API de transformação Java [359](#)
- invokeJExpression
  - método API [351](#)

## J

- JDK
  - transformação Java [310](#)
- JRE
  - transformação Java [310](#)

## L

- limite inferior
  - configurando [158](#), [267](#)
- limite superior
  - configurando [158](#), [267](#)
- linhas
  - sinalizando para atualização [666](#)
- linhas de entrada
  - obtendo o tipo de linha [340](#)
- link
  - um para muitos [61](#)
  - um para um [61](#)
- localizando
  - erros no código Java [330](#)
- log
  - definição [247](#)
- log de eventos
  - exibindo [249](#)
- log de eventos de tempo de design
  - descrição e localização [248](#)
- log de eventos de tempo de execução
  - transformação do Processador de Dados [248](#)
- log de eventos, tempo de design
  - descrição e localização [248](#)
- log do usuário
  - transformação do Processador de Dados [249](#)
- log, eventos de tempo de design
  - descrição e localização [248](#)
- logs
  - Transformações Java [344](#)

## M

- manter ordem das linhas
  - transformação de Gerador de Sequência [611](#)
- Mapeador
  - descrição [233](#)
- mapeamento de entrada
  - Transformação de Consumidor de Serviço da Web [678](#)
  - transformação de Consumidor de Serviço da Web REST [591](#)
- mapeamento de saída
  - Transformação de Consumidor de Serviço da Web [681](#)
  - transformação de Consumidor de Serviço da Web REST [593](#)
- mapeamentos
  - sinalizando linhas para atualização [666](#)
  - usando transformações de Roteador [605](#)
- mapeamentos dinâmicos
  - seletor de portas [284](#), [424](#)
  - Condição de pesquisa [422](#)
  - regras de seleção [284](#), [424](#)
  - transformação de Agregador [135](#)
  - transformação de Classificação [558](#)
  - transformação de Estratégia de Atualização [666](#)
  - transformação de Filtro [298](#)
  - transformação de Roteador [600](#)
  - Transformação do Classificador [619](#)
  - transformações de Associador [365](#)
  - transformações de Pesquisa [416](#)
- Maplet
  - referência [238](#)
- mascaramento aleatório
  - mascaramento de valores de data [205](#)
  - mascaramento de valores de string [204](#)
  - valores numéricos [204](#)
- mascaramento de chave
  - descrição [208](#)



- mascamamento de chave ([\)](#)
  - mascamamento de valores numéricos [209](#)
  - mascamando valores de data e hora [209](#)
  - valores numéricos [208](#)
- mascamamento de chaves
  - valores de string de mascaramento [208](#)
- mascamamento de expressão
  - descrição [206](#)
  - exemplo de mascaramento repetido [207](#)
  - mascamamento repetível [206](#)
  - regras e diretrizes [208](#)
- mascamamento de formato especial
  - CPF [223](#)
  - Endereços de e-mail [220](#)
  - Endereços IP [222](#)
  - números de cartão de crédito [220](#)
  - Números de Seguridade Social [222](#)
  - números de telefone [222](#)
  - números SIN repetíveis [223](#)
  - URLs [223](#)
- mascamamento de substituição
  - descrição [209](#)
  - propriedades de mascaramento [211](#)
- mascamamento dependente
  - descrição [213](#)
- mensagem SOAP
  - chaves [709](#)
- mensagens de solicitação de serviço da Web simultâneas
  - habilitando na transformação de Consumidor de Serviço da Web [684](#)
- mensagens SOAP
  - analisando elementos anyType [701](#)
  - analisando elementos de lista [704](#)
  - analisando elementos de opção [704](#)
  - analisando grupos de substituição [704](#)
  - dinamizando dados [713](#)
  - elementos de lista de mapas [718](#)
  - mapear elementos de opção [717](#)
  - mapear nós de ocorrência múltipla [699](#)
  - mapear portas [711](#)
  - mapear portas para elementos de união [718](#)
  - mapear várias portas de entrada [713](#)
  - visão geral [673](#)
- Método API de transformação Java
  - getResultDataType [358](#)
- Método defineJExpression
  - Transformações Java [338](#)
- Método failSession
  - Transformações Java [339](#)
- Método generateRow
  - Transformações Java [339](#)
- Método GetBytes
  - Transformações Java [357](#)
- Método GetDouble
  - Transformações Java [358](#)
- Método getInRowType
  - Transformações Java [340](#)
- Método getInt
  - Transformações Java [358](#)
- Método getLong
  - Transformações Java [358](#)
- Método getMetadata
  - transformações de Java [341](#)
- Método getResultDataType
  - Transformações Java [358](#)
- Método getResultMetadata
  - Transformações Java [358](#)

## N

## O

- Objeto de Dados de Sequência ()
  - propriedades [612](#)
- objeto Regra de Validação
  - criando na transformação do Processador de Dados [253](#)
- objeto XMap
  - criando na transformação do Processador de Dados [252](#)
- objetos de dados físicos
  - sincronização [569](#)
- objetos de dados relacionais
  - criando uma transformação de leitura dinâmica [573](#)
- objetos de dados relacionais dinâmicos
  - criando transformações de leitura [573](#)
- Opção de parâmetro
  - transformação de Associador [371](#)
- operação
  - elemento de arquivo WSDL [673](#)
- ORDER BY
  - consulta de pesquisa [409](#)
  - substituir [409](#)
- origem de pesquisa
  - parametrizando [420](#)
- otimização com seleção antecipada
  - Transformação de Consumidor de Serviço da Web [689](#)
  - transformação SQL [644](#)
- otimização de envio
  - habilitando na transformação SQL [645](#)
  - Transformação de Consumidor de Serviço da Web [689](#)
  - transformação SQL [644](#)
- otimização de origem
  - restrições [572](#)

## P

- Pacotes Java
  - importando [321](#)
- palavras reservadas
  - consulta de pesquisa [411](#)
- parâmetro de expressão
  - condição de pesquisa [425](#)
  - para uma condição de associação [371](#)
- parâmetros
  - condição de filtro [299](#)
  - transformação de Roteador [604](#)
- parâmetros de lista de portas
  - guia Agrupar por [140](#), [561](#)
- parâmetros de mapeamento
  - na substituição SQL da pesquisa [409](#)
- pesquisa
  - sem cache [440](#)
- pesquisa de substituição
  - diretrizes [411](#)
- pesquisas conectadas
  - descrição [406](#)
  - visão geral [407](#)
- pesquisas não conectadas
  - descrição [406](#)
  - visão geral [408](#)
- pontuações de controle
  - Transformação de Correspondência [466](#)
- pontuações de correspondência nula
  - Transformação de Correspondência [464](#)
- pontuações de vínculo
  - Transformação de Correspondência [466](#)
- porta de classificação
  - transformação de Classificação [560](#)
- porta de entrada de arquivo
  - transformação do Processador de Dados [235](#)

- porta de entrada de buffer
  - transformação do Processador de Dados [235](#)
- porta de filtro
  - Transformação de Consumidor de Serviço da Web [689](#)
- porta de parâmetro de serviço
  - transformação do Processador de Dados [235](#)
- porta de saída de arquivo
  - transformação do Processador de Dados [237](#)
- porta de saída de buffer
  - transformação do Processador de Dados [237](#)
- Porta de Saída Padrão
  - descrição [158](#)
- Porta NEXTVAL
  - Gerador de Sequência [608](#)
- porta SQLException
  - Transformação SQL [634](#)
- portas
  - atributos propagados por transformação [64](#)
  - configurar [60](#)
  - copiando do Excel [66](#)
  - criando [59](#)
  - entrada desnormalizada de serviço da Web [714](#)
  - mapear para mensagens SOAP [711](#)
  - ordem de avaliação [48](#)
  - portas variáveis [46](#)
  - regras e diretrizes de vinculação [62](#)
  - Transformação de Exceção de Registro Duplicado [269](#)
  - Transformação de Exceção de Registro Inválido [157](#)
  - Transformação de Gerador de Sequência [607](#)
  - transformação de Roteador [604](#)
  - transformações Java [314](#)
  - vinculação [60](#)
  - vinculando automaticamente [61](#)
  - vinculando manualmente [61](#)
  - vinculando por nome [61](#)
  - vinculando por posição [62](#)
  - visão geral de valores padrão [49](#)
- portas de entrada
  - transformações Java [314](#)
  - Transformações Java [315](#)
  - valores padrão [49](#)
- portas de entrada SQL
  - Transformação SQL [632](#)
- portas de parâmetro de serviço
  - transformação do Processador de Dados [236](#)
- portas de passagem
  - transformação de Expressão [280](#)
  - Transformação SQL [633](#)
  - valores padrão [49](#)
- Portas de problema de qualidade
  - em branco versus NULL [153](#)
- portas de problemas de qualidade
  - descrição [157](#)
- portas de saída
  - transformação Java [314](#)
  - Transformações Java [315](#)
  - tratamento de erros [49](#)
  - valores padrão [49](#)
- portas de transformação
  - visão geral [59](#)
- portas dinâmicas
  - condição de associação [371](#)
  - condição de pesquisa [417](#)
- portas variáveis
  - transformação de Expressão [280](#)
  - visão geral [46](#)
- problemas de qualidade
  - atribuindo portas para [161](#)

- procedimentos armazenados
  - exemplo [653](#)
  - exemplos de conjunto de resultados [651](#)
  - parâmetros [650](#)
  - porta de valor de retorno [649](#)
  - portas de entrada e saída [649](#)
  - registrando em variáveis [48](#)
  - Transformação SQL [649](#)
  - valor de retorno [650](#)
- processamento de alta precisão
  - transformações Java [315](#)
- processamento de nanossegundo
  - transformações Java [315](#)
- propriedade de classificação
  - transformação de Consolidação [188](#)
  - Transformação de Exceção de Registro Duplicado [271](#)
- propriedades
  - transformação de Gravação
    - propriedades [724](#)
  - transformação de Leitura
    - propriedades [566](#)
- propriedades avançadas
  - transformação de Associação [151](#)
  - transformação de Consolidação [188](#)
  - Transformação de Consumidor de Serviço da Web [684](#)
  - transformação de Consumidor de Serviço da Web REST [595](#)
  - Transformação de Exceção de Registro Duplicado [271](#)
  - transformação de Gerador de Chaves [390](#)
  - transformações Java [315](#)
- Propriedades avançadas
  - Transformação de Exceção de Registro Inválido [161](#)
- Propriedades Avançadas
  - Transformação SQL [635](#)
- propriedades de definição de janelas
  - ordem [286](#), [288](#)
  - partição [286](#), [288](#)
  - quadro [286](#), [287](#)
  - regras e diretrizes [289](#)

## Q

- quaisquer elementos
  - Transformação de Consumidor de Serviço da Web [678](#), [681](#)

## R

- redefinir
  - Transformação de Gerador de Sequência [611](#)
- registros
  - transformação de Normalizador [522](#)
- registros rejeitados
  - Transformação de Exceção de Registro Inválido [158](#)
- regras
  - valores padrão [55](#)
- regras de mascaramento
  - caracteres da string de origem [217](#)
  - caracteres de substituição da string de resultados [218](#)
  - embaçando [219](#)
  - formato de máscara [216](#)
  - formatos especiais de máscara [220](#)
  - intervalo [218](#)
- regras de seleção
  - mapeamentos dinâmicos [284](#), [424](#)
  - seletores de portas [283](#), [423](#)
  - transformação de Associador [366](#)

- regras e diretrizes
  - propriedades de definição de janelas [289](#)
- rejeitar arquivo
  - estratégias de atualização [667](#)
- restrições
  - otimização de origem [572](#)
- Reutilizável
  - Transformação de Gerador de Sequência [612](#)

## S

- saída de erros HTTP
  - habilitando na transformação de Consumidor de Serviço da Web [684](#)
- saída de falha genérica
  - habilitando na transformação de Consumidor de Serviço da Web [684](#)
- saída desnormalizada
  - análise de mensagens SOAP [700](#)
- saída dinamicada
  - análise de mensagens SOAP [701](#)
- saída distinta
  - Transformação do Classificador [620](#)
- saída exclusiva
  - transformação de Mascaramento de Dados [212](#)
- Script
  - criando na transformação do Processador de Dados [251](#)
- segurança da camada de transporte
  - Transformação de Consumidor de Serviço da Web [674](#)
  - Transformação de Consumidor de Serviços da Web REST [674](#)
- seletor de portas
  - regras de seleção [283](#), [284](#), [423](#), [424](#)
  - criando [285](#), [367](#), [426](#)
  - em expressões dinâmicas [290](#)
  - regras de seleção [283](#), [284](#), [423](#), [424](#)
  - transformação de Associador [365](#), [366](#)
- seletores de portas
  - em transformações de Associador [370](#)
- Sem estado
  - transformações Java [315](#)
- serviço
  - elemento de arquivo WSDL [673](#)
- serviço da Web
  - grupos de substituição [717](#)
  - mapear portas para anyTypes [716](#)
  - tipos derivados [716](#)
- serviço de Data Transformation
  - importando para o repositório do Modelo [255](#)
- Serviço do Data Transformation
  - importando para o repositório do Modelo [255](#)
  - importar vários serviços [255](#)
- sincronização
  - objetos de dados físicos [569](#)
  - objetos de dados personalizados [569](#)
- sincronizar com o editor
  - Transformação do Processador de Dados [254](#)
- solução de problemas
  - transformações Java [329](#)
- Streamer
  - descrição [233](#)
- strings
  - classificação [558](#)
- subregistros
  - transformação de Normalizador [522](#)
- substituição de strings
  - Transformação SQL [639](#)

- substituição SQL de pesquisa
  - caches dinâmicos [451](#)
- substituir ação SOAP
  - Transformação de Consumidor de Serviço da Web [684](#)

## T

- tabela de armazenamento
  - maskamento de dados de substituição [211](#)
  - maskamento de expressão [206](#)
- tabela de armazenamento compartilhado
  - transformação de Maskamento de Dados [226](#)
- Tabela de problemas
  - gerando [160](#)
- tabela de registros exclusivos
  - criando [278](#)
- tamanho do arquivo de cache
  - transformação de Associação [151](#)
  - transformação de Consolidação [188](#)
  - Transformação de Exceção de Registro Duplicado [271](#)
  - transformação de Gerador de Chaves [390](#)
- tamanho do cache
  - automático [72](#)
  - transformação de Maskamento de Dados [226](#)
- tamanho do cache automático
  - descrição [72](#)
- tamanho variável
  - na transformação de Classificador [624](#)
- tarefas Humanas
  - exceções de registro inválido [156](#)
- técnicas de maskamento
  - maskamento de dados [204](#)
- tipo de condição
  - avancado para transformação de Associador [369](#)
  - simples para a transformação de Associador [369](#)
- tipo de condição avançada
  - transformação de Associador [369](#)
- tipo de condição simples
  - transformação de Associador [369](#)
- tipos de dados
  - Transformações Java [311](#)
- tipos de dados primitivos Java
  - transformações Java [311](#)
- tipos de eventos
  - transformação do Processador de Dados [246](#)
- tipos derivados
  - analisando mensagens SOAP [702](#)
  - serviços da Web [716](#)
- transformação Agregador
  - usando variáveis [46](#)
- transformação Classificação
  - usando variáveis [46](#)
- transformação de agregador
  - tamanho do cache [70](#)
- transformação de Agregador
  - agrupar por portas [138](#)
  - caches [141](#)
  - classificando dados [142](#)
  - combinação de Estratégia de Atualização [668](#)
  - criando não reutilizável [144](#)
  - criando reutilizável [144](#)
  - desenvolvendo [135](#)
  - dicas [145](#)
  - entrada classificada [141](#)
  - expressões agregadas [136](#)
  - expressões não agregadas [141](#)
  - funções de agregação [137](#)

- transformação de Agregador ()
  - funções de agregação aninhadas [138](#)
  - mapeamentos dinâmicos [135](#)
  - portas [135](#)
  - propriedades avançadas [143](#)
  - solução de problemas [145](#)
  - visão geral [134](#)
- Transformação de agregador
  - ambiente não nativo [146](#)
  - Mecanismo Blaze [146](#)
  - Mecanismo Databricks Spark [147](#)
  - Mecanismo Spark [147](#)
- transformação de Analisador
  - visão geral [544](#)
- Transformação de Analisador
  - Ambiente não nativo [555](#)
- transformação de Associação
  - níveis de rastreamento [151](#)
  - propriedades avançadas [151](#)
  - visão geral [149](#)
- transformação de Associador
  - armazenando linhas mestras em cache [382](#)
  - associação externa completa [374](#)
  - associação externa detalhada [374](#)
  - associação externa mestra [373](#)
  - associação normal [372](#)
  - associando dados da mesma origem [379](#)
  - bloqueando pipelines de origem [381](#)
  - classificada [382](#)
  - condições [368](#)
  - configurando a ordem de classificação [375](#)
  - desempenho [382](#)
  - entrada classificada [375](#)
  - mapeamentos dinâmicos [365](#)
  - não classificado [382](#)
  - parâmetro de expressão [371](#)
  - portas [364](#)
  - portas dinâmicas [371](#)
  - propriedades avançadas [362](#)
  - regras de seleção [366](#)
  - regras e diretrizes [383](#)
  - seletores de portas [365](#)
  - tipo de condição [369](#)
  - tipo de condição avançada [369](#)
  - tipo de condição simples [369](#)
  - tipos de associação [372](#)
  - usando seletores de portas [370](#)
  - visão geral [361](#)
- Transformação de Associador
  - ambiente não nativo [384](#)
  - caches [363](#)
  - Mecanismo Blaze [384](#)
  - Mecanismo Databricks Spark [385](#)
  - Mecanismo Spark [384](#)
  - tamanho do cache [70](#)
- transformação de Classificação
  - definindo grupos para [560](#)
  - mapeamentos dinâmicos [558](#)
  - opções [558](#)
  - porta de classificação [560](#)
  - porta RANKINDEX [559](#)
  - portas [559](#)
  - propriedades avançadas [562](#)
  - visão geral [557](#)
- Transformação de Classificação
  - Ambiente não nativo [563](#)
  - caches [562](#)
  - Mecanismo Blaze [563](#)

- Transformação de Classificação ()
  - Mecanismo Databricks Spark [564](#)
  - Mecanismo Spark [564](#)
  - tamanho do cache [70](#)
- transformação de Classificador
  - algoritmo classificador [172](#)
  - ambiente não nativo [179](#)
  - visão geral [171](#), [618](#)
- Transformação de Classificador
  - Ambiente não nativo [627](#)
  - Mecanismo Blaze [627](#)
  - Mecanismo Databricks Spark [629](#)
  - Mecanismo Spark [628](#)
  - tamanho do cache [70](#)
- transformação de Comparação
  - visão geral [180](#)
- Transformação de Comparação
  - ambiente não nativo [185](#)
- transformação de Consolidação
  - níveis de rastreamento [188](#)
  - propriedades avançadas [188](#)
  - registros de classificação [188](#)
  - visão geral [186](#)
- Transformação de Consolidação
  - ambiente não nativo [202](#)
- Transformação de Consumidor de Serviço da Web
  - adicionando cabeçalhos HTTP [676](#)
  - autenticação por cookies [677](#)
  - compactação SOAP [687](#)
  - criando [691](#)
  - exibindo chaves [679](#), [682](#)
  - falhas SOAP genéricas [687](#)
  - habilitando a otimização de envio [690](#)
  - habilitando a saída de erros HTTP [684](#)
  - habilitando a saída de falha genérica [684](#)
  - mapeamento de entrada [678](#)
  - mapeamento de saída [681](#)
  - mensagens de solicitação de serviço da Web simultâneas [684](#)
  - mensagens SOAP [673](#)
  - nome WS-Security dinâmico [677](#)
  - nós de saída de mapeamento [681](#)
  - operações [674](#)
  - otimização com seleção antecipada [689](#)
  - otimização de envio [689](#)
  - otimização de filtros [689](#)
  - portas de entrada de mapeamento [678](#)
  - propriedades avançadas [684](#)
  - segurança [674](#)
  - segurança da camada de transporte [674](#)
  - tratamento de erros [687](#)
  - URL de ponto de extremidade [677](#)
  - URL de serviço da Web dinâmico [677](#)
  - visão geral [672](#)
- transformação de Consumidor de Serviço da Web REST
  - configuração [583](#)
  - configuração de mensagens [583](#)
  - configurando a URL base [595](#)
  - entrada classificada [595](#)
  - entrada de mapeamento [581](#)
  - identificação do recurso [584](#)
  - mapeamento de entrada [591](#)
  - mapeamento de saída [593](#)
  - mapeamento elementos para portas [588](#)
  - método Delete [587](#)
  - método Get [585](#)
  - método Post [586](#)
  - método Put [586](#)
  - Métodos HTTP [585](#)
- transformação de Consumidor de Serviço da Web REST ()
  - não reutilizável [596](#)
  - nível de rastreamento [595](#)
  - personalizando a exibição de mapeamentos de saída [594](#)
  - porta RequestInput [588](#)
  - portas [588](#)
  - portas de argumentos [589](#)
  - Portas de cabeçalho HTTP [589](#)
  - portas de código de resposta [590](#)
  - portas de entrada [588](#)
  - portas de entrada de mapeamento [592](#)
  - portas de passagem [588](#)
  - portas de saída [588](#)
  - portas de saída de mapeamento [594](#)
  - portas de saída XML [590](#)
  - Portas de URL [589](#)
  - portas para cookies [590](#)
  - processo [583](#)
  - propriedade de conexão [595](#)
  - propriedades avançadas [595](#)
  - regras de mapeamento de entrada [591](#)
  - regras de saída de mapeamento [594](#)
  - reutilizável [596](#)
  - saída de mapeamento [581](#)
  - suporte a servidores proxy [581](#)
  - tipo de mídia da internet [595](#)
  - validação de esquema XML [595](#)
  - visão geral [581](#)
- Transformação de Consumidor de Serviços da Web REST
  - segurança [674](#)
  - segurança da camada de transporte [674](#)
- transformação de Conversor de Maiúsculas/Minúsculas
  - visão geral [168](#)
- Transformação de Conversor de maiúsculas/minúsculas
  - ambiente não nativo [170](#)
- transformação de Correspondência
  - casos de uso [459](#)
  - definição de análise de campos [461](#)
  - definição de análise de identidade [461](#)
- Transformação de correspondência
  - exemplo de estratégia de correspondência [273](#)
- Transformação de Correspondência
  - Ambiente não nativo [484](#)
  - análise de origem dupla [461](#)
  - análise de origem única [461](#)
  - códigos de status de persistência [478](#), [479](#)
  - conceitos na análise de correspondência [460](#)
  - configurando uma operação de análise de correspondência [483](#)
  - criando um mapplet para análise de correspondência [482](#)
  - dados agrupados [462](#)
  - descrições de status de persistência [478](#), [479](#)
  - exibindo dados de análise de cluster [471](#)
  - exibindo dados de desempenho [471](#)
  - fatores de desempenho [470](#)
  - fluxo de processo de análise de campos [486](#)
  - fluxo do processo de análise de identidade [499](#)
  - formatos de saída [506](#)
  - Mecanismo Blaze [484](#)
  - Mecanismo Spark [484](#)
  - opções da exibição Saída de Correspondência [490](#)
  - opções de pontuação de cluster [465](#)
  - pontuações de controle [466](#)
  - pontuações de correspondência nula [464](#)
  - pontuações de vínculo [466](#)
  - portas de entrada predefinidas [477](#)
  - portas de saída predefinidas [477](#)
  - propriedades da exibição Saída de Correspondência [491](#)

- Transformação de Correspondência ()
  - propriedades de Correspondência na exibição Saída de Correspondência [508](#)
  - propriedades de Saída na exibição Saída de Correspondência [508](#)
  - tabelas de índice de identidade [469](#)
- transformação de Decisão
  - visão geral [258](#)
- Transformação de decisão
  - estratégia de exemplo de problema de qualidade [156](#)
- Transformação de Decisão
  - ambiente não nativo [264](#)
  - propriedades avançadas [264](#)
- transformação de Estratégia de Atualização
  - combinação de Agregador [668](#)
  - criando [666](#)
  - encaminhando linhas rejeitadas [667](#)
  - etapas para configurar [666](#)
  - expressões [667](#)
  - mapeamentos dinâmicos [666](#)
  - propriedades avançadas [667](#)
  - visão geral [665](#)
- Transformação de estratégia de atualização
  - Ambiente não nativo [669](#)
  - Mecanismo Blaze [669](#)
  - Mecanismo Spark [670](#)
- Transformação de Exceção
  - Exibição de Atribuição de Problema [160](#)
- Transformação de Exceção de Registro Duplicado
  - clusters [266](#)
  - configurando [278](#)
  - consolidação automática [267](#)
  - exemplo [272](#)
  - exemplo de definições de configuração [274](#)
  - exemplo de mapeamento [272](#)
  - exemplo de saída [275](#)
  - exemplo de saída de cluster [276](#)
  - Exibição de configuração [267](#)
  - gerar tabela de registros duplicados [268](#)
  - Grupo de Saída Padrão [270](#)
  - grupos de saída [270](#)
  - níveis de rastreamento [271](#)
  - portas [269](#)
  - propriedades avançadas [271](#)
  - registros de classificação [271](#)
  - visão geral [265](#)
- Transformação de Exceção de Registro Inválido
  - exemplo [163](#)
  - exemplo de mapplet [163](#)
  - exemplo de saída [165](#)
  - exemplo de saída de Problemas [166](#)
  - Exemplo de saída de registros inválidos [165](#)
  - Exibição de configuração [158](#)
  - fluxo de processo [154](#)
  - grupos de porta [157](#)
  - grupos de saída [153](#)
  - mapeamento [154](#)
  - portas de entrada [157](#)
  - portas de problema de qualidade em branco [153](#)
  - portas de saída [158](#)
  - problemas de qualidade [156](#)
  - propriedades avançadas [161](#)
  - saída de exemplo de registros válidos [167](#)
  - visão geral [152](#)
- Transformação de Exceção de Registros Inválidos
  - configurando [162](#)
- transformação de Expressão
  - configurações da porta de saída [291](#)
  - expressão dinâmica [290](#)

- transformação de Expressão ()
  - propriedades avançadas [295](#)
  - testando [281](#)
  - tipos de porta [280](#)
  - visão geral [279](#)
- Transformação de Expressão
  - ambiente não nativo [295](#)
  - Mecanismo Blaze [295](#)
  - Mecanismo Databricks Spark [296](#)
  - Mecanismo Spark [296](#)
- transformação de Filtro
  - condição de filtro [299](#)
  - dicas de desempenho [301](#)
  - linhas com valores nulos [301](#)
  - mapeamentos dinâmicos [298](#)
  - parâmetro de condição de filtro [299](#)
  - propriedades avançadas [301](#)
  - visão geral [297](#)
- Transformação de Filtro
  - ambiente não nativo [301](#)
  - Mecanismo Blaze [302](#)
- transformação de Gerador de Chaves
  - níveis de rastreamento [390](#)
  - propriedades avançadas [390](#)
  - visão geral [386](#)
- Transformação de Gerador de Chaves
  - Ambiente não nativo [391](#)
- transformação de Gerador de Sequência
  - criando [615](#)
  - manter ordem das linhas [611](#)
  - Propriedade Incrementar em [611](#)
- Transformação de Gerador de Sequência
  - Ambiente não nativo [617](#)
  - ciclo [610](#)
  - Ciclo [611](#)
  - criando chaves compostas [608](#)
  - intervalo de valores [611](#)
  - Mecanismo Blaze [617](#)
  - Mecanismo Spark [617](#)
  - Porta NEXTVAL [608](#)
  - portas [607](#)
  - propriedades [610](#), [612](#)
  - redefinir [611](#)
  - usando a função IIF para substituir chaves ausentes [609](#)
  - valor inicial [610](#)
  - visão geral [607](#)
- transformação de Gravação
  - visão geral [724](#)
- Transformação de Hierárquica para Relacional
  - configurar portas [307](#)
  - criando [308](#)
  - descrição [303](#)
  - exemplo [304](#)
- transformação de Leitura
  - visão geral [565](#)
- transformação de mascaramento de dados
  - mascarando os endereços IP [222](#)
- transformação de Mascaramento de Dados
  - arquivo de valor padrão [224](#)
  - caracteres da string de origem [217](#)
  - CPF repetível [223](#)
  - descrição [203](#)
  - dicionário para mascaramento de substituição [210](#)
  - diretório de cache [226](#)
  - diretrizes para mascaramento de expressão [208](#)
  - embaçando [219](#)
  - formato de máscara [216](#)
  - formatos especiais de máscara [220](#)

## transformação de Mascaramento de Dados ()

- intervalo [218](#)
  - intervalo de confirmação de armazenamento [226](#)
  - mascamamento aleatório [204](#)
  - mascamamento de cartão de crédito [220](#)
  - mascamamento de dados dependente [213](#)
  - mascamamento de expressão [206](#)
  - mascamamento de expressão do nome do dicionário [206](#)
  - mascamamento de substituição [209](#)
  - mascamamento de valores de data [219](#)
  - mascamamento repetível de expressão [206](#)
  - mascarando CPFs [223](#)
  - mascarando números de seguridade social [222](#), [225](#)
  - mascarando números de telefone [222](#)
  - mascarando o endereço de e-mail [220](#)
  - mascarando URLs [223](#)
  - números SIN repetíveis [223](#)
  - propriedades de mascaramento de substituição [211](#)
  - propriedades de tempo de execução [226](#)
  - propriedades do mascaramento de substituição [212](#)
  - Saída exclusiva [226](#)
  - tabela de armazenamento [206](#)
  - tabela de armazenamento compartilhado [226](#)
  - tabelas de armazenamento [211](#)
  - tamanho do cache [226](#)
  - técnicas de mascaramento [204](#)
- ## Transformação de Mascaramento de dados
- ambiente não nativo [230](#)
- ## Transformação de média ponderada
- Ambiente não nativo [722](#)
- ## transformação de Mesclagem
- visão geral [542](#)
- ## Transformação de Mesclagem
- Ambiente não nativo [543](#)
- ## transformação de Normalizador
- arrastando portas de outros objetos [537](#)
  - campos de ocorrência múltipla [521](#)
  - criando [536](#)
  - editando um grupo de saída [534](#)
  - exemplo de caso de uso [537](#)
  - GCID [521](#)
  - grupos de saída [531](#), [532](#)
  - mesclando campos [525](#)
  - portas de entrada [524](#)
  - registros de ocorrência múltipla [522](#)
  - subregistros [522](#)
- ## Transformação de Normalizador
- Ambiente não nativo [541](#)
  - exemplo de grupos de entrada e saída [539](#)
  - exemplo de mapeamento [538](#)
  - geração de chave [535](#)
- ## transformação de Padronizador
- visão geral [657](#)
- ## Transformação de Padronizador
- Ambiente não nativo [660](#)
- ## transformação de Pesquisa
- armazenando em cache [415](#)
  - cache dinâmico [445](#)
  - cache persistente [441](#)
  - caches [438](#)
  - condição de pesquisa [412](#)
  - conectada [407](#)
  - conflitos de nomes de portas [419](#)
  - criando não reutilizável [431](#)
  - criando pesquisas não conectadas [432](#)
  - criando reutilizável [430](#)
  - desenvolvendo [409](#)
  - mapeamentos dinâmicos [416](#)

## transformação de Pesquisa ()

- não conectada [408](#)
  - objeto de dados parametrizados relacionais [420](#)
  - parâmetros e variáveis de mapeamento [409](#)
  - pesquisa sem cache [440](#)
  - portas dinâmicas [417](#)
  - propriedades avançadas [429](#)
  - propriedades de consulta [416](#)
  - propriedades de tempo de execução [428](#)
  - regras e diretrizes da condição de pesquisa [414](#)
  - substituição de consulta de pesquisa [409](#)
  - substituindo a pesquisa [411](#)
  - tamanho do cache, reduzindo [409](#)
  - visão geral [405](#)
- ## Transformação de pesquisa
- Ambiente não nativo [435](#)
  - conectada [406](#)
  - Mecanismo Databricks Spark [437](#)
  - não conectada [406](#)
- ## Transformação de Pesquisa
- Mecanismo Blaze [435](#)
  - Mecanismo Spark [435](#)
  - tamanho do cache [70](#)
- ## Transformação de Processador de dados
- ambiente não nativo [257](#)
- ## transformação de Processamento de Dados
- componente de inicialização [237](#)
- ## Transformação de Relacional para Hierárquica
- criando [579](#)
  - criando portas [580](#)
  - descrição [575](#)
  - desenvolvimento [579](#)
  - exemplo [576](#)
  - portas [579](#)
- ## transformação de Roteador
- condição do filtro de grupo [601](#)
  - conectando em mapeamentos [605](#)
  - exemplo [601](#)
  - grupos [601](#)
  - mapeamentos dinâmicos [600](#)
  - parâmetros [604](#)
  - portas [604](#)
  - portas dinâmicas [603](#)
  - propriedades avançadas [605](#)
  - visão geral [599](#)
- ## Transformação de Roteador
- Ambiente não nativo [606](#)
- ## transformação de Rotulador
- visão geral [392](#)
- ## Transformação de Rotulador
- Ambiente não nativo [404](#)
- ## transformação de União
- visão geral [661](#)
- ## Transformação de União
- Ambiente não nativo [664](#)
  - Mecanismo Databricks Spark [664](#)
- ## Transformação de Validador de Endereço
- ambiente não nativo [132](#)
- ## transformação do Classificador
- Guia Classificação [620](#)
- ## Transformação do classificador
- modelo do classificador [172](#)
- ## Transformação do Classificador
- caches [624](#)
  - mapeamentos dinâmicos [619](#)
- ## Transformação do Consumidor de Serviços da Web REST
- criando [596](#)



- Transformação do Normalizador
  - hierarquia de entrada [523](#)
  - propriedades avançadas [535](#)
  - visão geral [520](#)
- transformação do Processador de Dados
  - configurações de codificação [239](#)
  - configurações [239](#)
  - configurações de controle de saída [241](#)
  - configurações de processamento [243](#)
  - configurações de XML [244](#)
  - descrição [233](#), [234](#)
  - exibições [234](#)
  - exportando como um serviço [255](#)
  - logs de usuário [249](#)
  - portas de entrada [235](#)
  - portas de parâmetro de serviço [236](#)
  - portas de saída [237](#)
  - testando no Visualizador de Dados [254](#)
- Transformação do Processador de Dados
  - Configurações de XMap [244](#)
  - criando [250](#)
  - portas [235](#)
- transformação Expressão
  - usando variáveis [46](#)
- Transformação Hierárquica para Relacional
  - criando portas [308](#)
  - desenvolvimento [308](#)
  - portas [306](#)
  - testando no Visualizador de Dados [309](#)
- transformação Java
  - métodos API [337](#)
  - visão geral [310](#)
- Transformação Java
  - Ambiente não nativo [334](#)
  - Mecanismo Blaze [334](#)
  - Mecanismo Spark [335](#)
- Transformação Normalizador
  - exemplo de definição [538](#)
  - exemplo de saída de mapeamento [540](#)
  - mesclando campos [526](#)
- Transformação Pesquisa
  - consulta padrão [409](#)
- transformação SQL
  - otimização com seleção antecipada [644](#)
  - otimização de envio [644](#)
  - propriedades da otimização de envio [645](#)
- Transformação SQL
  - parâmetros INOUT [650](#)
  - cardinalidade de linha de entrada para linha de saída [639](#)
  - chamando um procedimento armazenado [650](#)
  - conexão de tempo de execução [654](#)
  - conjuntos de resultados [651](#)
  - continuando com erro de SQL [643](#)
  - criando a partir de um procedimento armazenado [656](#)
  - criar como vazio [655](#)
  - definindo a conexão de banco de dados [648](#)
  - definindo a consulta [638](#)
  - definindo as portas de saída [632](#)
  - descrição das portas de entrada [632](#)
  - exemplo [645](#)
  - Exibição Propriedades Avançadas [635](#)
  - instrução de consulta [648](#)
  - número de linhas afetadas [635](#)
  - número de saída de linhas [641](#)
  - parâmetros de procedimento armazenado [650](#)
  - porta de valor de retorno [649](#)
  - porta SQLException [634](#)
  - portas [631](#)
- Transformação SQL ()
  - portas de passagem [633](#)
  - procedimentos armazenados [649](#)
  - restringindo linhas de saída [641](#)
  - substituição de strings de consulta [639](#)
  - vinculação de parâmetros [638](#)
  - visão geral [630](#)
- transformações
  - configurando portas no Excel [67](#)
  - copiando metadados [67](#)
  - ambiente Hadoop [36](#)
  - arquivos de cache [71](#)
  - ativo [32](#)
  - caches [70](#)
  - conectadas [33](#)
  - criando [58](#)
  - desenvolvendo [41](#)
  - editando no Excel [68](#)
  - editando reutilizável [57](#)
  - expressões [42](#)
  - Gerador de Sequência [607](#)
  - Java [310](#)
  - manipulação de erros [53](#)
  - não conectadas [33](#)
  - não reutilizável [58](#)
  - otimização do tamanho do cache [75](#)
  - particionamento do cache [75](#)
  - reutilizável [57](#)
  - tamanho do cache [72](#), [74](#)
  - validação de expressão [44](#)
  - vários grupos [41](#)
  - visão geral [32](#)
- transformações ativas
  - Classificação [557](#)
  - descrição [32](#)
  - Java [310](#), [311](#)
- transformações conectadas
  - Classificação [557](#)
  - Gerador de Sequência [607](#)
  - Java [310](#)
- transformações de Java
  - armazenamento de metadados [346](#)
  - Método getMetadata [341](#)
  - método storeMetadata [346](#)
  - recuperação de metadados [341](#)
- transformações de leitura
  - criando a partir de objetos de dados relacionais [573](#)
- transformações de serviços da Web
  - coluna Localização [708](#)
- transformações Java
  - classpath no nível de mapeamento [318](#)
  - código Java [318](#)
  - criando [314](#), [328](#)
  - criando portas [314](#)
  - erros de compilação [329](#)
  - Escopo de Transformação [315](#)
  - guia Auxiliares [323](#)
  - guia Código Completo [325](#)
  - guia Função [324](#)
  - guia Importações [323](#)
  - guia Na Entrada [323](#)
  - guia No Final [324](#)
  - obtendo o tipo de linha de entrada [340](#)
  - processamento de alta precisão [315](#)
  - processamento de nanossegundo [315](#)
  - propriedades avançadas [315](#)
  - Sem estado [315](#)
  - solução de problemas [329](#)



## Transformações Java

- ativa [311](#)
- compilando [329](#)
- configurando valores nulos em [345](#)
- conversão de tipo de dados [311](#)
- criando portas [315](#)
- criando trechos de código Java [320](#)
- erros de código de não usuário [331](#)
- erros de código do usuário [330](#)
- falha em mapeamentos [339](#)
- Guia Auxiliares [321](#)
- Guia Importações [321](#)
- Identificando a origem de erros de compilação [330](#)
- logs [344](#)
- Método defineJExpression [338](#)
- Método failSession [339](#)
- Método generateRow [339](#)
- Método getInRowType [340](#)
- Método incrementErrorCount [341](#)
- Método invokeJExpression [342](#)
- Método isNull [343](#)
- Método LogError [344](#)
- Método LogInfo [344](#)
- Método resetNotification [345](#)
- Método setNull [345](#)
- métodos API [337](#)
- passiva [311](#)
- portas de entrada [315](#)
- portas de saída [315](#)
- redefinindo variáveis em [345](#)
- tipos de dados primitivos Java [311](#)
- valores padrão para portas [315](#)
- verificando valores nulos em [343](#)
- transformações não conectadas
  - transformação de Pesquisa [408](#)
- transformações não reutilizáveis
  - descrição [58](#)
- transformações passivas
  - Gerador de Sequência [607](#)
  - Java [310](#), [311](#)
  - visão geral [33](#)
- transformações reutilizáveis
  - descrição [57](#)
  - editando [57](#)
- Transformador
  - descrição [233](#)
- tratar falha como erro
  - habilitando na transformação de Consumidor de Serviço da Web [684](#)
- trechos de código
  - criando para transformações Java [320](#)
- trechos de código Java
  - criando para transformações Java [320](#)

## U

- URL de ponto de extremidade
  - Transformação de Consumidor de Serviço da Web [677](#)
- URL dinâmico
  - Transformação de Consumidor de Serviço da Web [677](#)
- URL do ponto de extremidade
  - transformação de Consumidor de Serviço da Web REST [589](#)

## Usar mesclagem do Hive

- opção [667](#)

## V

- validando
  - valores padrão [53](#), [56](#)
- Valor Atual (propriedade)
  - Transformação de Gerador de Sequência [610](#)
- Valor Final (propriedade)
  - Transformação de Gerador de Sequência [610](#)
- valor inicial
  - Transformação de Gerador de Sequência [610](#)
- Valor Inicial (propriedade)
  - Transformação de Gerador de Sequência [610](#)
- valores de data
  - maskamento de dados aleatórios [205](#)
- Valores de Data e Hora
  - maskamento de dados [209](#)
- valores de porta
  - Transformações Java [315](#)
- valores de string
  - maskamento de dados chave [208](#)
  - maskamento de dados personalizado [204](#)
- valores não encontrados
  - substituindo com o Gerador de Sequência [609](#)
- valores nulos
  - configurando para transformação Java [345](#)
  - ignorando [52](#)
  - substituindo por uma constante [51](#)
  - verificando transformações Java [343](#)
- valores numéricos
  - maskamento aleatório [205](#)
  - maskamento de chave [209](#)
- valores padrão
  - definido pelo usuário [50](#)
  - inserção [53](#), [56](#)
  - maskamento de dados [224](#)
  - portas de entrada [49](#), [50](#)
  - portas de passagem [49](#), [50](#)
  - portas de saída [49](#), [50](#)
  - regras para [55](#)
  - validando [53](#), [56](#)
  - visão geral [49](#)
- variáveis
  - Inicializações [49](#)
  - ordem de avaliação da porta [48](#)
  - resultados de procedimento de armazenamento, captura [48](#)
  - Transformações Java [321](#)
  - visão geral [46](#)
- variáveis de instância
  - Transformações Java [321](#)
- variáveis de mapeamento
  - na substituição SQL da pesquisa [409](#)
- variáveis estáticas
  - Transformações Java [321](#)
- variáveis locais
  - visão geral [46](#)
- vários grupos
  - transformações [41](#)
- vinculação de parâmetros
  - Transformação SQL [638](#)