Informatica®
10.1

Guia de Perfil

Informatica Guia de Perfil
10.1
Junho 2016

© Copyright Informatica LLC 2014, 2018

Este software e documentação contêm informações de propriedade da Informatica LLC, são fornecidos sob um contrato de licença que contém restrições quanto a seu uso e divulgação, e são protegidos por leis de copyright. A engenharia reversa do software é proibida. 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) sem o consentimento prévio da Informatica LLC. Este Software pode estar protegido por patentes dos EUA e/ou internacionais e outras patentes pendentes.

O uso, duplicação ou divulgação do Software pelo Governo dos Estados Unidos estão sujeitos às restrições estipuladas no contrato de licença de software aplicável e como estabelecido em DFARS 227.7202-1(a) e 227.7702-3(a) (1995), DFARS 252.227-7013©(1)(ii) (OCT 1988), FAR 12.212(a) (1995), FAR 52.227-19 ou FAR 52.227-14 (ALT III), conforme aplicável.

As informações contidas neste produto ou documentação estão sujeitas a alteração sem aviso prévio. Informe-nos por escrito caso encontre quaisquer problemas neste produto ou documentação.

Informatica, Informatica Platform, Informatica Data Services, PowerCenter, PowerCenterRT, PowerCenter Connect, PowerCenter Data Analyzer, PowerExchange, PowerMart, Metadata Manager, Informatica Data Quality, Informatica Data Explorer, Informatica B2B Data Transformation, Informatica B2B Data Exchange Informatica On Demand, Informatica Identity Resolution, Informatica Application Information Lifecycle Management, Informatica Complex Event Processing, Ultra Messaging, Informatica Master Data Management e Live Data Map são marcas comerciais ou marcas registradas da Informatica LLC nos Estados Unidos e em jurisdições pelo mundo. Todos os outros nomes de outras companhias e produtos podem ser nomes ou marcas comerciais de seus respectivos proprietários.

Partes desta documentação e/ou software estão sujeitas a direitos autorais de terceiros, incluindo sem limitação: Copyright DataDirect Technologies. Todos os direitos reservados. Copyright © Sun Microsystems. Todos os direitos reservados. Copyright © RSA Security Inc. Todos os direitos reservados. Copyright © Ordinal Technology Corp. Todos os direitos reservados. Copyright © Aandacht c.v. Todos os direitos reservados. Copyright Genivia, Inc. Todos os direitos reservados. Copyright Isomorphic Software. Todos os direitos reservados. Copyright © Meta Integration Technology, Inc. Todos os direitos reservados. Copyright © Intalio. Todos os direitos reservados. Copyright © Oracle. Todos os direitos reservados. Copyright © Adobe Systems Incorporated. Todos os direitos reservados. Copyright © DataArt, Inc. Todos os direitos reservados. Copyright © ComponentSource. Todos os direitos reservados. Copyright © Microsoft Corporation. Todos os direitos reservados. Copyright © Rogue Wave Software, Inc. Todos os direitos reservados. Copyright © Teradata Corporation. Todos os direitos reservados. Copyright © Yahoo! Inc. Todos os direitos reservados. Copyright © Glyph & Cog, LLC. Todos os direitos reservados. Copyright © Thinkmap, Inc. Todos os direitos reservados. Copyright © Clearpace Software Limited. Todos os direitos reservados. Copyright © Information Builders, Inc. Todos os direitos reservados. Copyright © OSS Nokalva, Inc. Todos os direitos reservados. Copyright Edifecs, Inc. Todos os direitos reservados. Copyright Cleo Communications, Inc. Todos os direitos reservados. Copyright © International Organization for Standardization 1986. Todos os direitos reservados. Copyright © ej-technologies GmbH.Todos os direitos reservados. Copyright © Jaspersoft Corporation. Todos os direitos reservados. Copyright © International Business Machines Corporation. Todos os direitos reservados. Copyright © yWorks GmbH. Todos os direitos reservados. Copyright © Lucent Technologies.Todos os direitos reservados. Copyright © University of Toronto. Todos os direitos reservados. Copyright © Daniel Veillard.Todos os direitos reservados. Copyright © Unicode, Inc. Copyright IBM Corp. Todos os direitos reservados. Copyright © MicroQuill Software Publishing, Inc. Todos os direitos reservados. Copyright © PassMark Software Pty Ltd. Todos os direitos reservados. Copyright © LogiXML, Inc. Todos os direitos reservados. Copyright © 2003-2010 Lorenzi Davide, todos os direitos reservados. Copyright © Red Hat, Inc. Todos os direitos reservados. Copyright © The Board of Trustees of the Leland Stanford Junior University. Todos os direitos reservados. Copyright © EMC Corporation. Todos os direitos reservados. Copyright © Flexera Software.Todos os direitos reservados. Copyright © Jinfonet Software. Todos os direitos reservados. Copyright © Apple Inc. Todos os direitos reservados. Copyright © Telerik Inc. Todos os direitos reservados. Copyright © BEA Systems. Todos os direitos reservados. Copyright © PDFlib GmbH. Todos os direitos reservados. Copyright © Orientation in Objects GmbH. Todos os direitos reservados. Copyright © Tanuki Software, Ltd. Todos os direitos reservados. Copyright © Ricebridge. Todos os direitos reservados. Copyright © Sencha, Inc. Todos os direitos reservados. Copyright © Scalable Systems, Inc. Todos os direitos reservados. Copyright © jQWidgets. Todos os direitos reservados. Copyright © Tableau Software, Inc. Todos os direitos reservados. Copyright © MaxMind, Inc. Todos os direitos reservados. Copyright © TMate Software s.r.o. Todos os direitos reservados. Copyright © MapR Technologies Inc. Todos os direitos reservados. Copyright © Amazon Corporate LLC. Todos os direitos reservados. Copyright © Highsoft. Todos os direitos reservados. Copyright © Python Software Foundation. Todos os direitos reservados. Copyright © BeOpen.com. Todos os direitos reservados. Copyright © CNRI. Todos os direitos reservados.

Este produto inclui software desenvolvido pela Apache Software Foundation (http://www.apache.org/) e/ou outros softwares licenciados nas várias versões da Licença Apache (a "Licença"). Você pode obter uma cópia dessas Licenças em http://www.apache.org/licenses/. A menos que exigido pela legislação aplicável ou concordado por escrito, o software distribuído em conformidade com estas Licenças é fornecido "NO ESTADO EM QUE SE ENCONTRA", SEM GARANTIA OU CONDIÇÃO DE QUALQUER TIPO, seja expressa ou implícita. Consulte as Licenças para conhecer as limitações e as permissões que regulam o idioma específico de acordo com as Licenças.

Este produto inclui software desenvolvido pela Mozilla (http://www.mozilla.org/), direitos autorais de software de The JBoss Group, LLC; todos os direitos reservados; software copyright © 1999-2006 de Bruno Lowagie e Paulo Soares e outros produtos de software licenciados sob a Licença Pública GNU Lesser General Public License Agreement, que pode ser encontrada em http://www.gnu.org/licenses/lgpl.html. Os materiais são fornecidos gratuitamente pela Informatica, no estado em que se encontram, sem garantia de qualquer tipo, explícita nem implícita, incluindo, mas não limitando-se, as garantias implicadas de comerciabilidade e adequação a um determinado propósito.

O produto inclui software ACE(TM) e TAO(TM) com copyright de Douglas C. Schmidt e seu grupo de pesquisa na Washington University, University of California, Irvine e Vanderbilt University, Copyright (©) 1993-2006, todos os direitos reservados.

Este produto inclui o software desenvolvido pelo OpenSSL Project para ser usado no kit de ferramentas OpenSSL (copyright The OpenSSL Project. Todos os direitos reservados) e a redistribuição deste software está sujeita aos termos disponíveis em http://www.openssl.org e http://www.openssl.org/source/license.html.

Este produto inclui o software Curl com o Copyright 1996-2013, Daniel Stenberg, <daniel@haxx.se>. Todos os direitos reservados. Permissões e limitações relativas a este software estão sujeitas aos termos disponíveis em http://curl.haxx.se/docs/copyright.html. É permitido usar, copiar, modificar e distribuir este software com qualquer objetivo, com ou sem taxa, desde que a nota de direitos autorais acima e esta nota de permissão apareçam em todas as cópias.

O produto inclui software copyright 2001-2005 (©) MetaStuff, Ltd. Todos os direitos reservados. Permissões e limitações relativas a este software estão sujeitas aos termos disponíveis em http://www.dom4j.org/license.html.

O produto inclui o copyright de software © 2004-2007, The Dojo Foundation. Todos os direitos reservados. Permissões e limitações relativas a este software estão sujeitas aos termos disponíveis em http://dojotoolkit.org/license.

Este produto inclui o software ICU com o copyright International Business Machines Corporation e outros. Todos os direitos reservados. Permissões e limitações relativas a este software estão sujeitas aos termos disponíveis em http://source.icu-project.org/repos/icu/icu/trunk/license.html.

Este produto inclui o copyright de software © 1996-2006 Per Bothner. Todos os direitos reservados. O direito de usar tais materiais é estabelecido na licença que pode ser encontrada em http://www.gnu.org/software/kawa/Software-License.html.

Este produto inclui o software OSSP UUID com Copyright © 2002 Ralf S. Engelschall, Copyright © 2002 e OSSP Project Copyright © 2002 Cable & Wireless Deutschland. Permissões e limitações relativas a este software estão sujeitas aos termos disponíveis em http://www.opensource.org/licenses/mit-license.php.

Este produto inclui software desenvolvido pela Boost (http://www.boost.org/) ou sob a licença de software Boost. Permissões e limitações relativas a este software estão sujeitas aos termos disponíveis em http://www.boost.org/LICENSE_1_0.txt.

Este produto inclui software copyright © 1997-2007 University of Cambridge. Permissões e limitações relativas a este software estão sujeitas aos termos disponíveis em http://www.pcre.org/license.txt.

Este produto inclui o copyright de software © 2007 The Eclipse Foundation. Todos os direitos reservados. As permissões e as limitações relativas a este software estão sujeitas aos termos disponíveis em http://www.eclipse.org/org/documents/epl-v10.php e em http://www.eclipse.org/org/documents/edl-v10.php.

Este produto inclui softwares licenciados de acordo com os termos disponíveis em http://www.tcl.tk/software/tcltk/license.html, http://www.bosrup.com/web/overlib/?License, http://www.stlport.org/doc/ license.html, http:// asm.ow2.org/license.html, http://www.cryptix.org/LICENSE.TXT, http://hsqldb.org/web/hsqlLicense.html, http://httpunit.sourceforge.net/doc/ license.html, http://jung.sourceforge.net/license.txt , http://www.gzip.org/zlib/zlib_license.html, http://www.openldap.org/software/release/license.html, http://www.libssh2.org, http://slf4j.org/license.html, http://www.sente.ch/software/OpenSourceLicense.html, http://fusesource.com/downloads/license-agreements/fuse-message-broker-v-5-3- license-agreement; http://antlr.org/license.html; http://aopalliance.sourceforge.net/; http://www.bouncycastle.org/licence.html; http://www.jgraph.com/jgraphdownload.html; http://www.jcraft.com/jsch/LICENSE.txt; http://jotm.objectweb.org/bsd_license.html; . http://www.w3.org/Consortium/Legal/2002/copyright-software-20021231; http://www.slf4j.org/license.html; http://nanoxml.sourceforge.net/orig/copyright.html; http://www.json.org/license.html; http://forge.ow2.org/projects/javaservice/, http://www.postgresql.org/about/licence.html, http://www.sqlite.org/copyright.html, http://www.tcl.tk/software/tcltk/license.html, http://www.jaxen.org/faq.html, http://www.jdom.org/docs/faq.html, http://www.slf4j.org/license.html; http://www.iodbc.org/dataspace/iodbc/wiki/iODBC/License; http://www.keplerproject.org/md5/license.html; http://www.toedter.com/en/jcalendar/license.html; http://www.edankert.com/bounce/index.html; http://www.net-snmp.org/about/license.html; http://www.openmdx.org/#FAQ; http://www.php.net/license/3_01.txt; http://srp.stanford.edu/license.txt; http://www.schneier.com/blowfish.html; http://www.jmock.org/license.html; http://xsom.java.net; http://benalman.com/about/license/; https://github.com/CreateJS/EaselJS/blob/master/src/easeljs/display/Bitmap.js; http://www.h2database.com/html/license.html#summary; http://jsoncpp.sourceforge.net/LICENSE; http://jdbc.postgresql.org/license.html; http://protobuf.googlecode.com/svn/trunk/src/google/protobuf/descriptor.proto; https://github.com/rantav/hector/blob/master/LICENSE; http://web.mit.edu/Kerberos/krb5-current/doc/mitK5license.html; http://jibx.sourceforge.net/jibx-license.html; https://github.com/lyokato/libgeohash/blob/master/LICENSE; https://github.com/hjiang/jsonxx/blob/master/LICENSE; https://code.google.com/p/lz4/; https://github.com/jedisct1/libsodium/blob/master/LICENSE; http://one-jar.sourceforge.net/index.php?page=documents&file=license; https://github.com/EsotericSoftware/kryo/blob/master/license.txt; http://www.scala-lang.org/license.html; https://github.com/tinkerpop/blueprints/blob/master/LICENSE.txt; http://gee.cs.oswego.edu/dl/classes/EDU/oswego/cs/dl/util/concurrent/intro.html; https://aws.amazon.com/asl/; https://github.com/twbs/bootstrap/blob/master/LICENSE; https://sourceforge.net/p/xmlunit/code/HEAD/tree/trunk/LICENSE.txt.

Este produto inclui software licenciado de acordo com a Academic Free License (http://www.opensource.org/licenses/afl-3.0.php), a Common Development and Distribution License (http://www.opensource.org/licenses/cddl1.php), a Common Public License (http://www.opensource.org/licenses/cpl1.0.php), a Sun Binary Code License Agreement Supplemental License Terms, a BSD License (http://www.opensource.org/licenses/bsd-license.php), a nova BSD License (http://opensource.org/licenses/BSD-3-Clause), a MIT License (http://www.opensource.org/licenses/mit-license.php), a Artistic License (http://www.opensource.org/licenses/artistic-license-1.0) e a Initial Developer's Public License Version 1.0 (http://www.firebirdsql.org/en/initial-developer-s-public-license-version-1-0/).

Este produto inclui copyright do software © 2003-2006 Joe Walnes, 2006-2007 XStream Committers. Todos os direitos reservados. Permissões e limitações relativas a este software estão sujeitas aos termos disponíveis em http://xstream.codehaus.org/license.html. Este produto inclui software desenvolvido pelo Indiana University Extreme! Lab. Para obter mais informações, visite http://www.extreme.indiana.edu/.

Este produto inclui software Copyright © 2013 Frank Balluffi e Markus Moeller. Todos os direitos reservados. As permissões e limitações relativas a este software estão sujeitas aos termos da licença MIT.

Consulte as patentes em https://www.informatica.com/legal/patents.html.

ISENÇÃO DE RESPONSABILIDADE: a Informatica LLC fornece esta documentação no estado em que se encontra, sem garantia de qualquer tipo, expressa ou implícita, incluindo, mas não limitando-se, as garantias implícitas de não infração, comercialização ou uso para um determinado propósito. A Informatica LLC não garante que este software ou documentação não contenha erros. As informações fornecidas neste software ou documentação podem incluir imprecisões técnicas ou erros tipográficos. As informações deste software e documentação estão sujeitas a alterações a qualquer momento sem aviso prévio.

AVISOS

Este produto da Informatica (o "Software") traz determinados drivers (os "drivers da DataDirect") da DataDirect Technologies, uma empresa em funcionamento da Progress Software Corporation ("DataDirect"), que estão sujeitos aos seguintes termos e condições:

1. OS DRIVERS DA DATADIRECT SÃO FORNECIDOS NO ESTADO EM QUE SE ENCONTRAM, SEM GARANTIA DE QUALQUER TIPO, EXPRESSA OU IMPLÍCITA, INCLUINDO, MAS NÃO LIMITANDO-SE, AS GARANTIAS IMPLÍCITAS DE COMERCIALIZAÇÃO, ADEQUAÇÃO A UMA FINALIDADE ESPECÍFICA E NÃO INFRAÇÃO.
2. EM NENHUM CASO, A DATADIRECT OU SEUS FORNECEDORES TERCEIRIZADOS SERÃO RESPONSÁVEIS, EM RELAÇÃO AO CLIENTE FINAL, POR QUAISQUER DANOS DIRETOS, INDIRETOS, INCIDENTAIS, ESPECIAIS, CONSEQUENCIAIS OU DEMAIS QUE POSSAM ADVIR DO USO DE DRIVERS ODBC, SENDO OU NÃO ANTERIORMENTE INFORMADOS DAS POSSIBILIDADES DE TAIS DANOS. ESTAS LIMITAÇÕES SE APLICAM A TODAS AS CAUSAS DE AÇÃO, INCLUINDO, SEM LIMITAÇÕES, QUEBRA DE CONTRATO, QUEBRA DE GARANTIA, NEGLIGÊNCIA, RESPONSABILIDADE RIGOROSA, DETURPAÇÃO E OUTROS ATOS ILÍCITOS.

Data da Publicação: 2018-06-08

Conteúdo

Prefácio. 9

Recursos da Informatica. 9

Rede da Informatica. 9

Base de Dados de Conhecimento da Informatica. 9

Documentação da Informatica. 10

Matrizes de Disponibilidade de Produto Informatica. 10

Informatica Velocity. 10

Informatica Marketplace. 10

Suporte global a clientes Informatica. 10

Parte I: Introdução a Perfis. 11

Capítulo 1: Introdução a Perfis. 12

Visão Geral de Introdução a Perfis. 12

Processo de Criação de Perfil. 13

Ferramentas de Criação de Perfil. 13

Componentes de Perfil. 13

Capítulo 2: Conceitos de Perfil de Coluna. 15

Visão Geral de Conceitos de Perfil de ColunaVisão Geral de Perfis. 15

Opções de Perfil de Coluna. 16

Gerenciamento de Perfis com Controle de Versão e Bloqueios de Perfis de Repositório. 16

Scorecards. 17

Capítulo 3: Conceitos de Curadoria. 18

Visão Geral dos Conceitos de Curadoria. 18

Tarefas de Curadoria. 18

Parte II: Criação de perfil com o Informatica Analyst. 20

Capítulo 4: Perfis da Coluna no Informatica Analyst. 21

Visão Geral dos Perfis de Coluna no Informatica Analyst. 21

Processo de Criação de Perfil de Coluna. 22

Opções de Perfil. 22

Opções de Amostragem. 23

Opções de Busca Detalhada. 23

Opções de ambientes de tempo de execução. 23

Visão geral de perfis de sistema operacional no Informatica Analyst. 24

Selecionando um perfil do sistema operacional. 24

Visão Geral de Bloqueios de Ativos de Repositório e do Desenvolvimento Baseado em Equipe. . . . 24

Criando um perfil de coluna. 25
Editar um Perfil de Coluna. 26
Executar um Perfil. 27
Sincronizando um Objeto de Dados de Arquivo Simples. 28
Sincronizando um Objeto de Dados Relacional. 29

Capítulo 5: Regras no Informatica Analyst. 30

Regras na Visão Geral do Informatica Analyst. 30
Regras em um perfil de coluna. 30
Regras Predefinidas. 31
Processo de Regras Predefinidas. 31
Aplicar uma Regra Predefinida. 31
Regras de Expressão. 32
Processo de Regras de Expressão. 32
Criando uma Regra de Expressão. 33

Capítulo 6: Resultados do Perfil da Coluna no Informatica Analyst. 34

Resultados de perfil de coluna na visão geral do Informatica Analyst. 34
Exibição Resumida. 35
Propriedades da exibição resumida. 36
Filtros padrão na exibição resumida. 36
Exibição Detalhada. 37
Painéis da exibição detalhada. 38
Estatísticas. 39
Visualização de Dados. 40
Tipos de dados. 40
Exceções. 41
Padrões. 42
Valores. 43
Tipos de execução de perfil. 46
Execução de perfil mais recente. 46
Execução de perfil histórico. 46
Execução de perfil consolidada. 46
Selecionando uma execução de perfil. 47
Visão geral dos resultados de vários perfis de comparação. 48
Comparando resultados de vários perfis. 48
Exibição Resumida dos Resultados de Perfil de Comparação 49
Exibição detalhada dos resultados do perfil de comparação 51
Busca Detalhada de Perfil da Coluna. 52
Fazer Busca Detalhada na Linha de Dados. 52
Aplicar Filtros aos Dados de Busca Detalhada. 52
Curadoria na ferramenta Analyst. 53
Aprovando tipos de dados e domínios de dados. 53

Rejeitando tipos de dados e domínios de dados. . . . 53
Arquivos de exportação de perfil de coluna no Informatica Analyst. . . . 54
Resultados de Exportação de Perfil em um Arquivo CSV. . . . 54
Resultados de Exportação de Perfil no Microsoft Excel. . . . 54
Exportar Resultados de Perfil do Informatica Analyst. . . . 55

Capítulo 7: Scorecards no Informatica Analyst. . . . 56

Scorecards na Visão Geral do Informatica Analyst. . . . 56
Processo de Scorecard no Informatica Analyst. . . . 57
Criando um scorecard no Informatica Analyst. . . . 58
Adicionando colunas a um scorecard existente. . . . 59
Executar um Scorecard. . . . 60
Exibir um Scorecard. . . . 61
Editar um Scorecard. . . . 61
Métricas. . . . 62
Pesos de Métrica. . . . 62
Valor da Qualidade de Dados. . . . 62
Definir Limites. . . . 63
Grupos de Métrica. . . . 63
Criar um Grupo de Métrica. . . . 63
Mover Pontuações para um Grupo de Métrica. . . . 64
Editar um Grupo de Métrica. . . . 64
Excluir um Grupo de Métrica. . . . 64
Buscas Detalhadas em Colunas. . . . 65
Gráficos de Tendência. . . . 65
Gráfico de Tendência de Pontuação. . . . 66
Gráfico de Tendência de Custo. . . . 66
Exibir Gráficos de Tendência. . . . 67
Exportando Gráficos de Tendência. . . . 68
Arquivos de Exportação de Scorecard no Informatica Analyst. . . . 69
Resultados da Exportação do Scorecard no Microsoft Excel. . . . 69
Exportando Resultados do Scorecard do Informatica Analyst. . . . 69
Notificações de Scorecard. . . . 70
Modelo de Mensagem de Notificação de E-mail. . . . 70
Configurar as Notificações de Scorecard. . . . 71
Definir as Configurações Globais para Notificações de Scorecard. . . . 72
Linhagem de Scorecard. . . . 72
Exibir a Linhagem de Scorecard o Informatica Analyst. . . . 73

Parte III: Criação de perfil com o Informatica Developer. . . . 74

Capítulo 8: Perfis de Objeto de Dados. . . . 75

Perfis de Coluna no Informatica Developer. . . . 75

Opções de Filtragem. 75
Opções de Amostragem. 76
Criando um Único Perfil de Objeto de DadosCriando um Perfil. 76
Criando Vários Perfis de Objeto de Dados. 77
Sincronizando um Objeto de Dados de Arquivo Simples. 78
Sincronizar um Objeto de Dados Relacionais. 78

Capítulo 9: Perfis de coluna em fontes de dados semiestruturadas. 79

Visão geral de perfis de coluna em fontes de dados semiestruturadas. 79
Perfil de coluna em uma fonte de dados JSON ou XML. 80
Criando um objeto de dados em uma fonte de dados JSON ou XML. 80
Perfil de coluna em uma fonte de dados semiestruturada no HDFS. 81
Perfil de coluna em uma fonte de dados XML ou JSON no HDFS. 81
Perfil de coluna em uma fonte de dados Avro ou Parquet no HDFS. 81
Criando uma conexão HDFS. 81
Criando um objeto de dados de arquivo complexo em um arquivo JSON ou XML no HDFS. 82
Criando um objeto de dados em uma fonte de dados Avro ou Parquet. 83
Executando um perfil de coluna em fontes de dados semiestruturadas. 84

Capítulo 10: Regras no Informatica Developer. 86

Visão Geral de Regras no Informatica DeveloperDiretrizes para Regras. 86
Criar uma Regra no Informatica Developer. 87
Aplicando uma Regra no Informatica DeveloperAplicando uma Regra. 87

Capítulo 11: Coluna de Resultados de Perfil no Informatica Developer. 88

Resultados do Perfil de Colunas no Informatica DeveloperResultados do Perfil de Colunas. 88
Propriedades de Valor de Coluna. 89
Propriedades de Padrão de Coluna. 90
Propriedades de Estatísticas de Coluna. 90
Propriedades de tipo de dados da coluna. 91
Curadoria no Informatica DeveloperCuradoria no Informatica Developer. 92
Aprovando tipos de dadosAprovando tipos de dados no Informatica Developer. 92
Rejeitando tipos de dadosRejeitando tipos de dados no Informatica Developer. 92
Exportar Resultados de Perfil do Informatica Developer. 93

Capítulo 12: Scorecards no Informatica Developer. 94

Visão Geral dos Scorecards no Informatica Developer. 94
Criar um Scorecard. 94
Exportar um Arquivo de Recursos para a Linhagem de Scorecard. 95
Exibir a Linhagem de Scorecard do Informatica Developer. 95

Capítulo 13: Criação de Perfil de Mapplet e Mapeamento. 97

Visão Geral da Criação de Perfil de Mapplet e MapeamentoPerfis de Mapplet e Mapeamento. 97

Executar um Perfil em um Mapplet ou Objeto de Mapeamento. 97
Comparar Perfis de Objetos de Mapplet ou de Mapeamento. 98
Gerar um Mapeamento de Perfil. 98

Índice. 100

Prefácio

O *Guia de Perfil* da Informatica foi escrito para usuários do Informatica Analyst e do Informatica Developer. Ele contém informações sobre como você pode usar perfis para analisar o conteúdo e a estrutura dos dados.

Use perfis de coluna para determinar as características das colunas, como frequência de valor, porcentagens e padrões, em uma fonte de dados.

Recursos da Informatica

Rede da Informatica

A Rede da Informatica hospeda o Suporte Global a Clientes da Informatica, a Base de Dados de Conhecimento da Informatica e outros recursos de produtos. Para acessar a Rede da Informatica, visite https://network.informatica.com.

Como membro, você pode:

- Acessar todos os seus recursos Informatica em um só lugar.
- Pesquisar a Base de Dados de Conhecimento em busca de recursos de produtos, incluindo documentações, perguntas frequentes e práticas recomendadas.
- Visualizar informações sobre disponibilidade de produtos.
- Revisar seus casos de suporte.
- Encontrar a sua Rede de Grupo de Usuários da Informatica local e colaborar com seus colegas.

Como membro, você pode:

- Acessar todos os seus recursos Informatica em um só lugar.
- Pesquisar a Base de Dados de Conhecimento em busca de recursos de produtos, incluindo documentações, perguntas frequentes e práticas recomendadas.
- Visualizar informações sobre disponibilidade de produtos.
- 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 pesquisar a Rede da Informatica em busca de recursos de produtos, como documentações, artigos de instruções, práticas recomendadas e PAMs.

Para acessar a Base de Dados de Conhecimento, visite https://kb.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.

Documentação da Informatica

Para obter a documentação mais recente do seu produto, navegue pela Base de Dados de Conhecimento da Informatica
em .https://kb.informatica.com/_layouts/ProductDocumentation/Page/ProductDocumentSearch.aspx

Em caso de dúvidas, comentários ou ideias sobre esta documentação, entre em contato com a equipe de Documentação da Informatica pelo e-mail infa_documentation@informatica.com.

Matrizes de Disponibilidade de Produto Informatica

As Matrizes de Disponibilidade de Produto (PAMs) indicam as versões dos sistemas operacionais, os bancos de dados e outros tipos de fontes e destinos de dados com os quais uma versão de produto é compatível. Se você for membro da Rede da Informatica, poderá acessar PAMs 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. Desenvolvido com base na experiência no mundo real de centenas de projetos de gerenciamento de dados, o Informatica Velocity representa o conhecimento coletivo de nossos consultores, que trabalharam com organizações de todo o mundo para planejar, desenvolver, implantar e manter soluções de gerenciamento de dados bem-sucedidas.

Se você for membro da Rede da Informatica, poderá acessar 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.

Informatica Marketplace

O Informatica Marketplace é um fórum onde você pode encontrar soluções que aumentam, ampliam ou aprimoram suas implementações da Informatica. Aproveitando qualquer uma das centenas de soluções fornecidas por desenvolvedores e parceiros da Informatica, você pode melhorar sua produtividade e agilizar o tempo de implementação nos seus projetos. Você pode acessar o Informatica Marketplace através do link https://marketplace.informatica.com.

Suporte global a clientes Informatica

Você pode entrar em contato com um Centro de Suporte Global por telefone ou via Suporte Online na 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:
http://www.informatica.com/us/services-and-training/support-services/global-support-centers.

Se você for membro da Rede da Informatica, poderá usar o Suporte Online em http://network.informatica.com.

Parte I: Introdução a Perfis

Esta parte contém os seguintes capítulos:

- Introdução a Perfis, 12
- Conceitos de Perfil de Coluna, 15
- Conceitos de Curadoria, 18

CAPÍTULO 1

Introdução a Perfis

Este capítulo inclui os seguintes tópicos:

- Visão Geral de Introdução a Perfis, 12
- Processo de Criação de Perfil, 13
- Ferramentas de Criação de Perfil, 13
- Componentes de Perfil, 13

Visão Geral de Introdução a Perfis

Crie e execute um perfil para localizar o conteúdo, a qualidade e a estrutura das fontes de dados de um aplicativo, esquema ou empresa. O conteúdo da fonte de dados inclui frequências de valor e tipos de dados. A estrutura de fonte de dados inclui chaves e dependências funcionais.

Como parte do processo de descoberta, você pode criar e executar perfis. Um perfil é um objeto de repositório que localiza e analisa todas as irregularidades nas fontes de dados da empresa e problemas ocultos em dados que colocam os projetos em risco. A execução de um perfil em qualquer fonte de dados da empresa oferece uma boa compreensão das vantagens e desvantagens de seus dados e metadados.

Você pode usar o Informatica Analyst e o Informatica Developer para analisar os dados e metadados de origem. Analistas e desenvolvedores podem usar essas ferramentas para colaborar, identificar problemas de qualidade dos dados e analisar relacionamentos de dados. Com base na função do seu cargo, você pode usar os recursos da ferramenta Analyst ou da Developer tool. O grau de criação de perfil que você pode executar difere com base na ferramenta usada.

Você pode executar as seguintes tarefas na ferramenta Developer e na ferramenta Analyst:

- Execute a criação de perfil de coluna. O processo inclui descobrir o número de valores exclusivos, valores nulos e padrões de dados em uma coluna.
- Adicionar regras a perfis de coluna.
- Faça a curadoria dos tipos de dados inferidos nos resultados do perfil.
- Use scorecards para monitorar a qualidade de dados.
- Gere um mapeamento a partir de um perfil.

Processo de Criação de Perfil

Quando você inicia um projeto de integração de dados, a criação de perfil geralmente é a primeira etapa. Você pode criar perfis para analisar o conteúdo, a qualidade e a estrutura de fontes de dados. Como parte do processo de criação de perfil, você descobre os metadados das fontes de dados.

Você usa diferentes perfis para diferentes tipos de análise de dados, como um perfil de coluna. Você revela e documenta problemas de qualidade de dados. Conclua as seguintes tarefas para executar a criação de perfil:

1. Localize e analise o conteúdo dos dados das fontes de dados. Inclui tipos de dados, a frequência de valor, a frequência de padrão e as estatísticas de dados, como os valores mínimo e máximo.
2. Consulte os resultados da criação de perfil.
3. Crie dados de referência.
4. Faça uma busca detalhada nos resultados do perfil.
5. Documente problemas de dados.
6. Crie e execute regras.
7. Crie scorecards para monitorar a qualidade dos dados.

Ferramentas de Criação de Perfil

Você usa várias ferramentas da Informatica para gerenciar o processo de criação de perfil.

Você pode usar as seguintes ferramentas para gerenciar o processo de criação de perfil:

Informatica Administrator

Gerencie usuários, grupos, privilégios e funções. Você pode administrar o Serviço Analyst e gerenciar permissões para projetos e objetos no Informatica Analyst. Você pode controlar as permissões de acesso no Informatica Developer usando esta ferramenta.

Informatica Developer

Crie e execute perfis nessa ferramenta para localizar e analisar os metadados de uma ou mais fontes de dados. Você cria perfis usando um assistente.

Informatica Analyst

Você pode executar um perfil de coluna na ferramenta Analyst. Depois de executar um perfil, você pode fazer uma busca detalhada nos dados de linhas em uma fonte de dados.

Componentes de Perfil

Um perfil tem vários componentes que você pode usar para analisar o conteúdo e a estrutura das fontes de dados de maneira eficaz.

Um perfil tem os seguintes componentes:

Filtro

Cria um subconjunto da fonte de dados original que atende a critérios específicos. Você pode então executar um perfil na amostra de dados.

Regra

A lógica comercial que define condições aplicadas aos dados quando você executa um perfil. Adicione uma regra ao perfil para validar os dados.

Marca

Os metadados que definem um objeto no repositório do Modelo com base no uso comercial. Crie marcas para agrupar objetos de acordo com seu uso comercial.

Comentário

Descrição sobre o perfil. Use comentários para compartilhar informações sobre perfis com outros usuários das ferramentas Analyst e Developer.

Scorecard

Uma representação gráfica dos valores válidos para uma coluna ou a saída de uma regra nos resultados do perfil. Use scorecards para medir o progresso da qualidade dos dados.

CAPÍTULO 2

Conceitos de Perfil de Coluna

Este capítulo inclui os seguintes tópicos:

- Visão Geral de Conceitos de Perfil de ColunaVisão Geral de Perfis, 15
- Opções de Perfil de Coluna, 16
- Gerenciamento de Perfis com Controle de Versão e Bloqueios de Perfis de Repositório, 16
- Scorecards, 17

Visão Geral de Conceitos de Perfil de ColunaVisão Geral de Perfis

Um perfil de coluna determina as características das colunas em uma fonte de dados, como a frequência de valor, porcentagens e padrões.

A criação de perfil de coluna revela os seguintes fatos sobre dados:

- O número de valores nulos, distintos e não distintos em cada coluna, expresso como um número e uma porcentagem.
- Os padrões de dados em cada coluna e as frequências com que esses valores ocorrerem.
- As estatísticas sobre os valores de coluna, como o número máximo e mínimo de tamanhos de valores e o primeiro e o último valores em cada coluna.
- Tipos de dados documentados, tipos de dados inferidos e possíveis conflitos entre os tipos de dados documentados e inferidos.
- Exceções de frequência de valor e padrão.

Você pode configurar as seguintes opções ao criar ou editar um perfil:

- Opções de perfil de coluna. É possível selecionar as colunas nas quais você deseja executar um perfil e escolher uma opção de amostragem e uma opção de busca detalhada.
- Adicionar, editar ou excluir filtros e regras.

Nos resultados do perfil, você pode adicionar comentários e marcas a um perfil e às colunas em um perfil. Termos comerciais podem ser atribuídos a colunas.

O repositório do Modelo usa bloqueios de perfis de repositório para bloquear perfis e, assim, impedir que os usuários sobrescrevam trabalhos. O sistema de controle de versão salva várias versões de um perfil e atribui um número de versão a cada uma. Você pode fazer o check-out de um perfil e então submetê-lo a check-in depois de fazer alterações. É possível desfazer a ação de check-out de um perfil antes de submetê-lo novamente a check-in.

Crie scorecards para analisar periodicamente a qualidade de dados. Crie scorecards antes e depois de aplicar regras para criação de perfis de forma que você possa exibir uma representação gráfica dos valores válidos para as colunas.

Use o Serviço de Agendador para agendar execuções de perfil e de scorecard em um horário específico ou em intervalos. O Serviço de Agendador gerencia agendamentos para perfis, scorecards, mapeamentos implantados e fluxos de trabalho implantados. Você pode criar, gerenciar e executar agendamentos no Informatica Administrator.

Opções de Perfil de Coluna

Ao criar um perfil, você pode usar o assistente de perfil para definir filtros, regras, opções de busca detalhada, opções de amostragem e conexão. Essas opções determinam como o perfil lê as linhas a partir dos dados de origem.

Você pode definir as seguintes opções em um perfil de coluna, em um perfil de descoberta do domínio de dados ou em um perfil de descoberta empresarial:

- Filtros. Você pode criar e aplicar filtros a um perfil.
- Regras. Você pode adicionar regras ao criar um perfil. Você pode reutilizar as regras criadas na ferramenta Analyst ou na Developer tool.
- Opções de busca detalhada. Você pode optar por ler dados atuais na fonte de dados ou ler dados de perfil que são preparados no warehouse de criação de perfil.
- Opções de amostragem. Você pode escolher uma das opções de amostragem para determinar o número de linhas nas quais executar um perfil.
- Conexão. É possível escolher Nativo, Hive ou Hadoop como ambiente de tempo de execução.

Gerenciamento de Perfis com Controle de Versão e Bloqueios de Perfis de Repositório

O repositório do Modelo bloqueia perfis para impedir que os usuários sobrescrevam trabalhos. Quando você começa a editar um perfil, o perfil é bloqueado para impedir que outros usuários salvem alterações nele. O bloqueio é liberado quando você salva o perfil. O gerenciamento de perfis com controle de versão cria versões de um perfil, e é possível visualizar o histórico de versões.

O repositório do Modelo bloqueia um perfil quando você o edita na Developer tool ou na ferramenta Analyst. Se a ferramenta parar inesperadamente, o bloqueio será mantido para que, quando você se conectar novamente ao repositório do Modelo, possa visualizar os perfis que bloqueou. Você pode continuar a editar os perfis ou pode desbloqueá-los.

Quando o repositório do Modelo está integrado a um sistema de controle de versão, é possível gerenciar versões de um perfil. Por exemplo, você pode fazer o check-out e o check-in de perfis, desfazer check-outs, visualizar versões de histórico específicas do perfil e visualizar os perfis que estão em check-out. Para obter informações sobre bloqueios de ativos de repositório e o gerenciamento de ativos com controle de versão na ferramenta Analyst, consulte o *Guia da Ferramenta Analyst*. Para obter informações sobre bloqueios de objetos de repositório e o gerenciamento de objetos com controle de versão na Developer tool, consulte o *Guia da Developer Tool*.

Scorecards

Um scorecard é a representação gráfica dos valores válidos para uma coluna ou saída de uma regra nos resultados do perfil. Use scorecards para medir o progresso da qualidade de dados. Você pode criar um scorecard a partir de um perfil e monitorar o progresso de qualidade de dados ao longo do tempo.

Um scorecard tem vários componentes, como métricas, grupos de métrica e limites. Depois de executar um perfil, você pode adicionar colunas de origem como métricas para um scorecard e configurar os valores válidos para as métricas. Scorecards ajudam a organização a medir o valor da qualidade de dados, rastreando o custo de dados inválidos em níveis de métricas e scorecard. Para medir o custo de dados inválidos para cada métrica, atribua uma unidade de custo à métrica e defina um custo fixo ou variável. Quando você executar o scorecard, os resultados de scorecard incluirão o custo de dados inválidos para cada métrica e o valor de custo total para todas as métricas.

Use um grupo de métrica para categorizar as métricas relacionadas em um scorecard para um conjunto. Um limite identifica a faixa, em porcentagem, de dados inválidos que é aceitável para colunas em um registro. Os limites para uma faixa de dados podem ser definidos como bom, aceitável ou inaceitável.

Ao executar um scorecard, configure se você deseja fazer uma busca detalhada das métricas de pontuação nos dados ativos ou nos dados preparados. Depois que executar um scorecard e exibir as pontuações, faça uma busca detalhada em cada métrica para identificar registros de dados válidos e registros que não são válidos. Também pode exibir a linhagem do scorecard para cada métrica ou grupo de métrica em um scorecard. Para rastrear a qualidade de dados eficientemente, você pode usar gráficos de tendências de pontuação e gráficos de tendências de custo. Esses gráficos monitoram como as pontuações e o custo de dados inválidos mudaram no decorrer de um período de tempo.

O depósito de criação de perfil armazena as estatísticas de scorecard e informações de configuração. Você pode configurar um aplicativo de terceiros para obter resultados de scorecard e executar relatórios. Também pode exibir os resultados do scorecard em um aplicativo da Web, portal ou relatório, como um relatório de business intelligence.

CAPÍTULO 3

Conceitos de Curadoria

Este capítulo inclui os seguintes tópicos:

- Visão Geral dos Conceitos de Curadoria, 18
- Tarefas de Curadoria, 18

Visão Geral dos Conceitos de Curadoria

Curadoria é o processo de validação e gerenciamento dos metadados descobertos em uma fonte de dados de forma que os metadados estejam prontos para uso e preparação de relatórios.

Você pode fazer curadoria dos seguintes resultados de perfil inferidos:

- Tipos de dados
- Domínios de dados
- Chaves primárias
- Chaves externas

Você faz curadoria dos resultados de perfil inferidos para tornar precisos os metadados sobre colunas, domínios de dados e relacionamentos de objeto de dados nos bancos de dados e esquemas. Em seguida, você pode localizar os metadados mais relevantes ao usar a pesquisa de descoberta para pesquisar informações em vários repositórios. Você pode também encontrar os metadados mais relevantes ao exibir o diagrama de relacionamento de chave externa nos resultados da descoberta empresarial.

Você pode fazer curadoria das inferências de metadados específicas que um perfil gera como parte da execução do perfil. Por exemplo, você pode aprovar ou rejeitar os tipos de dados inferidos nos resultados de perfil de coluna e nos resultados da descoberta de domínio de dados. Você também pode aprovar ou rejeitar as chaves primárias e as chaves externas inferidas nos resultados da descoberta empresarial.

Tarefas de Curadoria

Você pode fazer curadoria dos resultados do perfil após a execução do perfil. Você também pode reverter uma decisão de curadoria que tomou quando você executou o perfil anteriormente.

Você pode executar as seguintes tarefas de curadoria na ferramenta Analyst:

- Aprovar ou rejeitar os tipos de dados inferidos para várias colunas e domínios de dados.

- Restaurar tipos de dados aprovados ou rejeitados para o status inferido.
- Restaure domínios de dados aprovados ou rejeitados para o status inferido.
- Exibir ou ocultar as linhas de resultados rejeitadas.
- Excluir colunas de execuções de perfil com base nas preferências específicas de metadados, como tipos de dados e domínios de dados aprovados.

Você pode executar as seguintes tarefas de curadoria na Developer tool:

- Aprovar ou rejeitar os tipos de dados inferidos para várias colunas.
- Restaurar tipos de dados aprovados ou rejeitados para o status inferido.
- Restaure domínios de dados aprovados ou rejeitados para o status inferido.
- Exibir ou ocultar as linhas de resultados rejeitadas.
- Aprovar ou rejeitar os objetos de dados nos resultados da detecção de chave primária.
- Aprovar ou rejeitar os resultados da descoberta empresarial, incluindo os resultados da detecção de chave externa.
- Excluir colunas de execuções de perfil com base nas preferências específicas de metadados, como tipos de dados e domínios de dados aprovados.

Parte II: Criação de perfil com o Informatica Analyst

Esta parte contém os seguintes capítulos:

- Perfis da Coluna no Informatica Analyst, 21
- Regras no Informatica Analyst, 30
- Resultados do Perfil da Coluna no Informatica Analyst, 34
- Scorecards no Informatica Analyst, 56

CAPÍTULO 4

Perfis da Coluna no Informatica Analyst

Este capítulo inclui os seguintes tópicos:

- Visão Geral dos Perfis de Coluna no Informatica Analyst, 21
- Processo de Criação de Perfil de Coluna, 22
- Opções de Perfil, 22
- Visão geral de perfis de sistema operacional no Informatica Analyst, 24
- Visão Geral de Bloqueios de Ativos de Repositório e do Desenvolvimento Baseado em Equipe, 24
- Criando um perfil de coluna, 25
- Editar um Perfil de Coluna, 26
- Executar um Perfil, 27
- Sincronizando um Objeto de Dados de Arquivo Simples, 28
- Sincronizando um Objeto de Dados Relacional, 29

Visão Geral dos Perfis de Coluna no Informatica Analyst

Ao criar um perfil, selecione as colunas no objeto de dados no qual você deseja executar um perfil. Você pode configurar as opções de amostragem e de busca detalhada para obter uma criação de perfil mais rápida. É possível escolher um ambiente de tempo de execução. Ao criar um perfil, você pode adicionar regras e filtros a ele. Depois de executar o perfil, examine as estatísticas de criação de perfil para entender os dados.

Você pode criar um perfil de tabelas amplas e arquivos simples que tenham no máximo 1000 colunas. Ao criar ou executar um perfil, você pode optar por selecionar todas as colunas ou cada coluna de um perfil. Você pode selecionar todas as colunas para fazer uma busca detalhada e exibir frequências de valores para essas colunas.

Você pode criar perfis de coluna com os seguintes métodos no Informatica Analyst:

- Clique com o botão direito do mouse no objeto de dados no espaço de trabalho **Biblioteca** para criar um perfil.
- Usar opções padrão para criar um perfil de coluna padrão.

- Personalizar as configurações do perfil para criar um perfil personalizado.

Nota: Você pode exibir e executar o perfil em fontes de dados Avro, JSON, Parquet e XML. Você pode criar e editar um perfil de coluna em fontes de dados Avro, JSON, Parquet e XML no Informatica Developer.

Processo de Criação de Perfil de Coluna

Como parte do processo de criação de perfil de colunas, você pode optar por incluir todas as colunas de origem para criação de perfil ou por selecionar colunas específicas. Você também pode aceitar as opções de perfil padrão ou configurar as opções de amostragem, as opções de busca detalhada e o ambiente de tempo de execução.

As seguintes etapas descrevem o processo de criação de perfil de coluna:

1. Escolha um nome, uma descrição e uma localização para o perfil de coluna.
2. Selecione um objeto de dados importado ou uma origem externa na qual você deseja executar o perfil.
3. Opcionalmente, visualize os dados de origem.
4. Selecione as colunas nas quais você deseja executar o perfil.
5. Determine se você deseja criar o perfil com as opções padrão ou alterá-las. As opções que podem ser configuradas incluem opções de amostragem, opções de busca detalhada e o ambiente de tempo de execução.
6. Opcionalmente, adicione regras e filtros ao criar o perfil.
7. Execute o perfil.

Nota: Leve em consideração as regras e diretrizes a seguir para nomes de coluna e criação de perfil multilíngue e dados Unicode:

- É possível criar perfis de dados multilíngues de diferentes origens e exibir resultados do perfil com base em definições de localidade no navegador. A ferramenta Analyst altera os tipos de dados de Data e Hora, Numérico e Decimal com base na localidade do navegador.
- Classifique em dados multilíngues. É possível classificar em dados multilíngues. A ferramenta Analyst exibe a ordem de classificação com base no local do navegador.
- Para criar o perfil de dados Unicode em um banco de dados DB2, defina a variável de ambiente de banco de dados DB2CODEPAGE no banco de dados e reinicie o Serviço de Integração de Dados.

Opções de Perfil

As opções de perfil incluem opções de amostragem de dados e de busca detalhada de dados. Você pode configurar essas opções ao criar ou editar um perfil de coluna para um objeto de dados.

Use o espaço de trabalho **Descoberta** para configurar as opções de perfil. Você pode escolher criar um perfil com as opções padrão para as opções de colunas, amostragem e busca detalhada. Use a opção de busca detalhada para escolher entre dados ativos e dados de fase.

Opções de Amostragem

As opções de amostragem determinam o número de linhas escolhido pela ferramenta Analyst para executar um perfil. É possível configurar as opções de amostragem quando você define ou executa um perfil.

A tabela a seguir descreve as opções de amostragem para um perfil:

Opção	Descrição
Todas as linhas	Selecione todas as linhas no objeto de dados.
Fazer amostragem das primeiras <number> linhas	O número de linhas que você deseja executar em relação ao perfil. A ferramenta Analyst escolhe entre as primeiras linhas na origem.
Amostragem aleatória de <number> linhas	O número de linhas de uma amostra aleatória na qual executar o perfil. A amostragem aleatória força a ferramenta Analyst a realizar a busca detalhada nos dados de fase. Observe que essa opção pode afetar o desempenho da busca detalhada.
Amostra aleatória	O tamanho de amostra aleatória com base no número de linhas no objeto de dados. A amostragem aleatória força a ferramenta Analyst a realizar a busca detalhada nos dados de fase. Observe que essa opção pode afetar o desempenho da busca detalhada.
Exclui tipos de dados e domínios de dados aprovados da inferência de tipo de dados e de domínio de dados nas execuções de perfil subsequentes	Exclui o tipo de dados ou o domínio de dados aprovado da inferência de tipo de dados e domínio dados a partir da próxima execução de perfil.

Opções de Busca Detalhada

Você pode configurar as opções de busca detalhada ao definir ou editar um perfil.

A tabela a seguir descreve as opções de busca detalhada para um perfil:

Opções	Descrição
Ativo	Faz uma busca detalhada nos dados ativos para fazer a leitura dos dados atuais na fonte de dados.
Em fase	Faz uma busca detalhada nos dados de fase para fazer a leitura de dados de perfil que estão em fase no depósito de criação de perfil.
Selecionar colunas	Identifica as colunas para busca detalhada que você não selecionou para criação de perfil.

Opções de ambientes de tempo de execução

Você pode escolher a opção Nativo, Hive ou Hadoop como ambiente de tempo de execução para um perfil de coluna. O Informatica Analyst define o ambiente de tempo de execução na definição de perfil depois que

você escolhe um ambiente de tempo de execução. O ambiente de tempo de execução não afeta os resultados do perfil.

A seguinte tabela descreve as opções de ambiente de tempo de execução para um perfil de coluna:

Opção	Descrição
Nativo	A ferramenta Analyst envia os trabalhos de perfil ao Módulo do Serviço de Criação de Perfil. Em seguida, o Módulo do Serviço de Criação de Perfil decompõe os trabalhos de perfil em um conjunto de mapeamentos. O Serviço de Integração de Dados executa esses mapeamentos e grava os resultados do perfil no warehouse de perfil.
Hive	O Serviço de Integração de Dados envia a lógica de perfil por push para o mecanismo Hive no cluster Hadoop para executar perfis.
Hadoop	O Serviço de Integração de Dados envia a lógica de perfil por push para o mecanismo Blaze no cluster Hadoop para executar perfis.

Visão geral de perfis de sistema operacional no Informatica Analyst

Você pode escolher um perfil do sistema operacional na ferramenta Analyst. Depois que você escolhe um perfil do sistema operacional, o Serviço de Integração de Dados cria e executa os perfis de coluna, os perfis de descoberta empresarial e os scorecards com base na permissão do usuário do perfil do sistema operacional.

A ferramenta Analyst usa o perfil padrão para executar perfis e scorecards. Se você tiver apenas um perfil do sistema operacional, ele será selecionado por padrão. Se você tiver vários perfis de sistema operacional, poderá escolher um deles.

Selecionando um perfil do sistema operacional

Você pode selecionar um perfil do sistema operacional no Informatica Analyst. O Serviço de Integração de Dados usa as permissões do usuário do perfil do sistema operacional para executar os trabalhos de criação de perfil.

1. Na área de cabeçalho do Informatica Analyst, clique em **<Nome de usuário>** > **Configurações**.

   A caixa de diálogo **Configurações** é exibida.
2. Selecione um perfil do sistema operacional. Clique em **Salvar**.

Visão Geral de Bloqueios de Ativos de Repositório e do Desenvolvimento Baseado em Equipe

O repositório do Modelo bloqueia perfis para impedir que os usuários sobrescrevam o trabalho de outro usuário. Se o repositório do Modelo estiver integrado a um sistema de controle de versão, ele salvará várias versões de ativos e atribuirá um número de versão a uma versão. Você pode fazer o check-out e o check-in

de perfis e também desfazer check-outs. É possível visualizar uma versão específica de um perfil que você colocou em check-out.

Quando você começar a editar um perfil na ferramenta Analyst, o repositório do Modelo bloqueará esse perfil para que outros usuários não possam editá-lo. Ao salvar o perfil, você mantém o bloqueio. Quando você fecha o perfil, o repositório do Modelo o desbloqueia.

Com o gerenciamento de ativos com controle de versão, o repositório do Modelo impede que perfis sejam substituídos por outros membros da equipe. Quando você tentar editar um perfil colocado em check-out por outro usuário, receberá uma notificação indicando o usuário que fez check-out desse perfil. É possível abrir um perfil em check-out no modo somente leitura ou salvá-lo com um nome diferente.

Você pode selecionar uma versão do perfil na caixa de diálogo Propriedades do Perfil para visualizar a definição do perfil para essa versão. A opção Propriedades de Perfil pode ser acessada no menu Ações. Para obter mais informações sobre bloqueios de ativos de repositório e o gerenciamento de ativos com controle de versão, consulte o *Guia da Ferramenta Analyst*.

Criando um perfil de coluna

Você pode criar um perfil personalizado ou um perfil padrão. Ao criar um perfil personalizado, você pode configurar as colunas, as linhas de amostra e as opções de busca detalhada. Ao criar um perfil padrão, o perfil de coluna e a descoberta de domínio de dados são executados em todo o conjunto de dados com todos os domínios de dados.

1. No espaço de trabalho **Descoberta**, clique em **Perfil** ou selecione **Novo** > **Perfil** na área de cabeçalho.

   Nota: Você pode clicar com o botão direito do mouse no objeto de dados no espaço de trabalho **Biblioteca** e criar um perfil. Nesse perfil, o nome do perfil, o nome da localização e o objeto de dados são extraídos das propriedades do objeto de dados. Você pode criar um perfil padrão ou personalizar as configurações para criar um perfil personalizado.

   O assistente de **Novo Perfil** é exibido.
2. A opção **Origem única** é selecionada por padrão. Clique em **Avançar**.
3. Na tela **Especificar Propriedades Gerais**, insira um nome e uma descrição opcional para o perfil. No campo Localização, selecione o projeto ou a pasta na qual você deseja criar o perfil. Clique em **Avançar**.
4. Na tela **Selecionar Origem**, clique em **Escolher** para selecionar um objeto de dados ou clique em **Novo** para importar um objeto de dados. Clique em **Avançar**.
   - Na caixa de diálogo **Escolher Objeto de Dados**, selecione um objeto de dados. Clique em **OK**.
     O painel Propriedades exibe as propriedades do objeto de dados selecionado. O painel Visualização de Dados exibe as colunas no objeto de dados.
   - Na caixa de diálogo **Novo Objeto de Dados**, você pode escolher uma conexão, um esquema, uma tabela ou uma exibição no qual criar um perfil, selecionar uma localização e criar uma pasta para importar o objeto de dados. Clique em **OK**.
5. Na tela **Selecionar Origem**, selecione as colunas nas quais você deseja executar um perfil. Opcionalmente, selecione **Nome** para selecionar todas as colunas. Clique em **Avançar**.

   Todas as colunas são selecionadas por padrão. A ferramenta Analyst lista as propriedades de coluna, como o nome, tipo de dados, precisão, escala, valor nulo e participantes na chave primária de cada coluna.

6. Na tela **Especificar Configurações**, opte por executar um perfil de coluna, uma descoberta de domínio de dados ou um perfil de coluna com descoberta de domínio de dados. Por padrão, a opção de perfil de coluna está selecionada.
   - Escolha **Executar perfil de coluna** para executar um perfil de coluna.
   - Escolha **Executar descoberta de domínio de dados** para realizar a descoberta de domínio de dados. No painel **Domínio de dados**, selecione os domínios de dados que você deseja descobrir, selecione critérios de conformidade e selecione as colunas para a descoberta do domínio de dados na caixa de diálogo **Editar seleção de colunas para a descoberta do domínio de dados**.
   - Escolha **Executar perfil de coluna** e **Executar descoberta de domínio de dados** para executar o perfil de coluna com a descoberta de domínio de dados. Selecione as opções de domínio de dados no painel **Domínio de dados**.

     Nota: Por padrão, as colunas que você seleciona são para o perfil de coluna e a descoberta de domínio de dados. Clique em **Editar** para marcar ou desmarcar colunas para a descoberta de domínio de dados.
   - Escolha dados, colunas ou dados e colunas para executar a descoberta de domínio de dados.
   - Escolha uma opção de amostragem. Você pode escolher **Todas as linhas (análise completa)**, **Fazer amostragem das primeiras**, **Amostra aleatória** ou **Amostra aleatória (automática)** como opção de amostragem no painel **Executar perfil em**. Essa opção aplica-se a um perfil de coluna e uma descoberta de domínio de dados.
   - Escolha uma opção de busca detalhada. Você pode escolher a opção de busca detalhada **Ativa** ou **Preparado** ou pode escolher **Desativar** para desativar a busca detalhada no painel **Busca Detalhada**. Opcionalmente, clique em **Selecionar Colunas** para selecionar colunas para fazer uma busca detalhada. Você pode optar por omitir o tipo de dados e a inferência de domínio de dados das colunas com um tipo de dados ou domínio de dados aprovado.
   - Escolha **Nativo**, **Hive** ou **Hadoop** como ambiente de tempo de execução. Se você escolher Hive, clique em **Escolher** para selecionar uma conexão do Hive na caixa de diálogo **Selecionar uma Conexão do Hive**. Se você escolher Hadoop, clique em **Escolher** para selecionar uma conexão do Hadoop na caixa de diálogo **Selecionar uma Conexão do Hadoop**.
7. Clique em **Avançar**.

   A tela **Especificar Regras e Filtros** é aberta.
8. Na tela **Especificar Regras e Filtros**, é possível realizar as seguintes tarefas:
   - Criar, editar ou excluir uma regra. Você pode aplicar as regras existentes no perfil.
   - Criar, editar ou excluir um filtro.

     Nota: Ao criar um scorecard nesse perfil, você pode reutilizar os filtros criados para o perfil.
9. Clique em **Salvar e Concluir** para criar o perfil ou clique em **Salvar e Executar** para criar e executar o perfil.

Editar um Perfil de Coluna

Você pode fazer alterações em um perfil de coluna depois de executá-lo.

1. No espaço de trabalho **Biblioteca**, selecione o projeto que contém o perfil ou selecione o perfil no painel **Ativos**.
2. Clique no nome do perfil.

A exibição resumida é exibida no espaço de trabalho **Descoberta**.

3. Se o sistema de controle de versão estiver ativado, clique em **Ações** > **Fazer Check-Out** para fazer check-out do perfil.
4. Clique em **Ações** > **Editar Perfil**.

   O assistente de **Perfil** é exibido.
5. Com base nas alterações que você deseja fazer, escolha uma das seguintes opções de página:
   - **Especificar Propriedades Gerais**. Altere as propriedades básicas, como o nome, a descrição e a localização.
   - **Selecionar Origem**. Escolha outras fontes de dados e colunas correspondentes para executar o perfil.
   - **Especificar Configurações**. Escolha executar o perfil de coluna ou o perfil de coluna com a descoberta de domínio de dados. Selecione os domínios de dados que você deseja descobrir e modifique as opções de descoberta do domínio de dados, amostragem e busca detalhada.
   - **Especificar Regras e Filtros**. Crie, edite ou exclua regras e filtros.
6. Clique em **Salvar e Concluir** para concluir a edição do perfil ou clique em **Salvar e Executar** para editar e executar o perfil.
7. Se o sistema de controle de versão estiver ativado, será necessário realizar as seguintes tarefas:
   - Clique em **Salvar e Concluir** para concluir a edição do perfil.
   - Na exibição resumida, clique em **Fazer Check-In** para fazer o check-in do perfil.
   - Clique em **Ações** > **Executar Perfil** para executar o perfil.

Executar um Perfil

Execute um perfil para analisar uma fonte de dados quanto ao conteúdo e estrutura e selecione colunas e regras para fazer uma busca detalhada. É possível fazer uma busca detalhada em dados de fase ou ativos para colunas e regras. Você pode executar um perfil somente em uma coluna ou regra sem executar o perfil em todas as colunas de origem após a execução de perfil inicial.

1. No espaço de trabalho **Biblioteca**, selecione o projeto ou a pasta que contém o perfil no painel Projetos ou selecione o perfil no painel **Ativos**.
2. Clique em **Ações** > **Abrir**.

   A exibição resumida é exibida no espaço de trabalho **Descoberta**.
3. Clique em **Ações** > **Executar Perfil.**

   A ferramenta Analyst realiza uma execução de perfil e mostra os resultados do perfil na exibição resumida.
4. Na exibição resumida, clique em uma coluna para exibir os resultados de coluna.

   A exibição detalhada é exibida.

Sincronizando um Objeto de Dados de Arquivo Simples

Você pode sincronizar as alterações para uma fonte de dados de arquivo simples externa com seu objeto de dados na ferramenta Analyst. Use o assistente **Sincronizar Arquivo Simples** para sincronizar os objetos de dados.

1. Abra o espaço de trabalho **Biblioteca**.
2. Na seção **Projetos**, selecione um objeto de dados de arquivo simples de um projeto.

   A ferramenta Analyst exibe a visualização de dados para o arquivo simples na guia **Visualização de Dados**.
3. Clique na guia **Propriedades**.
4. No menu Ações, clique em **Sincronizar**.

   O assistente de **Sincronizar Arquivo Simples** é exibido.
5. Escolha procurar uma localização ou insira um caminho de rede para importar o arquivo simples.
   - Procure uma localização e clique em **Escolher Arquivo** para selecionar o arquivo simples em um diretório que sua máquina pode acessar.
   - Para inserir um caminho de rede, selecione **Inserir um Caminho de Rede** e configure o caminho e nome do arquivo.
6. Clique em **Avançar**.
7. Escolha entre importar um arquivo simples delimitado ou com largura fixa.
   - Para importar um arquivo simples delimitado, aceite a opção **Delimitado**.
   - Para importar um arquivo simples de largura fixa, selecione a opção **Largura Fixa**.
8. Clique em **Avançar**.
9. Configure as opções de arquivo simples para o arquivo simples delimitado ou com largura fixa.
10. Clique em **Avançar**.
11. Opcionalmente, altere os atributos da coluna.
12. Clique em **Avançar**.
13. Aceite o nome padrão ou insira outro nome para o arquivo simples.
14. Opcionalmente, insira uma descrição.
15. Clique em **Concluir**.

    Uma mensagem de sincronização solicita que você confirme a ação.
16. Clique em **Sim** para sincronizar o arquivo simples.

    Uma mensagem é exibida informando que a sincronização foi concluída. Para exibir detalhes das alterações de metadados, clique em **Mostrar Detalhes**.
17. Clique em **OK**.

Sincronizando um Objeto de Dados Relacional

Você pode sincronizar as alterações de uma fonte de dados relacionais externa com seu objeto de dados de tabela. As alterações de fonte de dados externas incluem adicionar, alterar e remover colunas de origem e colunas de regra.

1. Abra o espaço de trabalho **Biblioteca**.
2. Na seção **Projetos**, selecione um objeto de dados de tabela de um projeto.

   A ferramenta Analyst exibe a visualização de dados para a tabela na guia **Visualização de Dados**.
3. Clique na guia **Propriedades**.
4. No menu Ações, clique em **Sincronizar**.

   Uma mensagem solicita que você confirme a ação.
5. Para concluir o processo de sincronização, clique em **Sim**.

   Uma mensagem de status de sincronização é exibida.
6. Uma mensagem é exibida informando que a sincronização foi concluída.

   Para exibir detalhes das alterações de metadados, clique em **Mostrar Detalhes**.
7. Clique em **OK**.

CAPÍTULO 5

Regras no Informatica Analyst

Este capítulo inclui os seguintes tópicos:

- Regras na Visão Geral do Informatica Analyst, 30
- Regras Predefinidas, 31
- Regras de Expressão, 32

Regras na Visão Geral do Informatica Analyst

Uma regra é uma lógica de negócios que define condições aplicadas a dados de origem quando você executa um perfil. Você pode adicionar uma regra ao perfil para validar dados.

Talvez você queira usar uma regra em diferentes circunstâncias. Você pode adicionar uma regra para limpar uma ou mais colunas de dados. Também pode adicionar uma regra de pesquisa que fornece informações que os dados de origem não fornecem. Pode, ainda, adicionar uma regra para validar uma regra de limpeza para um projeto de qualidade de dados ou integração de dados.

Ao adicionar uma regra a um perfil, você pode criar ou aplicar uma regra.

É possível criar ou aplicar os seguintes tipos de regra a um perfil:

- **Regras de expressão**. Use as funções de expressão e colunas para definir a lógica da regra. Crie regras de expressão na ferramenta Analyst. Um analista pode criar uma regra de expressão e promovê-la para uma regra reutilizável que outros analistas podem usar em vários perfis.
- **Regras predefinidas**. Inclui regras reutilizáveis que um desenvolvedor cria na Developer tool. As regras que um desenvolvedor cria na Developer tool como mapplets podem aparecer na ferramenta Analyst como regras reutilizáveis.

Regras em um perfil de coluna

Você pode criar regras ao criar um perfil de coluna.

Depois de adicionar uma regra a um perfil, você poderá executar o perfil novamente para a coluna de regra. A ferramenta Analyst exibe os resultados de perfil da coluna de regra na exibição resumida. Você pode exibir os resultados de coluna de uma regra na exibição detalhada. A saída de uma regra pode ser uma ou mais colunas virtuais. As colunas virtuais existem nos resultados de perfil. A ferramenta Analyst executa um perfil nas colunas virtuais. Por exemplo, use uma regra predefinida que divida uma coluna que contém o nome e o sobrenome em colunas virtuais FIRST_NAME e LAST_NAME. A ferramenta Analyst executa o perfil nas colunas FIRST_NAME e LAST_NAME.

Nota: Se você excluir um objeto de regra que outros tipos de objeto referenciam, a ferramenta Analyst exibe uma mensagem que lista os tipos de objeto. Determine o impacto de excluir a regra antes de excluí-la.

Regras Predefinidas

Regras predefinidas são regras criadas na Developer tool ou fornecidas com as ferramentas Developer e Analyst. Aplique regras predefinidas para a ferramenta Analyst criar perfis para modificar ou validar dados de origem.

Regras predefinidas usam transformações para definir a lógica da regra. As regras predefinidas podem ser usadas com vários perfis. No repositório do Modelo, uma regra predefinida é um mapplet com um grupo de entrada, um grupo de saída e transformações que definem a lógica da regra.

Processo de Regras Predefinidas

Use o **Assistente de Nova Regra** para aplicar uma regra predefinida a um perfil.

Você pode executar as etapas a seguir para aplicar uma regra predefinida:

1. Abra um perfil.
2. Selecione uma regra predefinida.
3. Consulte os parâmetros da regra.
4. Selecione a coluna de entrada. Você pode selecionar várias colunas caso deseje aplicar a regra em mais de uma coluna.
5. Configure as opções de criação de perfil.

Aplicar uma Regra Predefinida

Quando você aplicar uma regra predefinida, selecione a regra e configure as colunas de entrada e saída para a regra. Aplique uma regra predefinida para usar uma regra promovida a regra reutilizável ou use uma regra criados por um desenvolvedor.

1. No espaço de trabalho **Biblioteca**, selecione o projeto que contém o perfil ou selecione o perfil no painel **Ativos**.
2. Clique em **Ações** > **Abrir** para abrir o perfil.

   A exibição resumida é exibida no espaço de trabalho **Descoberta**.
3. Clique em **Ações** > **Editar Perfil**.

   O **Assistente de Perfil** é exibido.
4. Clique em **Especificar Regras e Filtros**.
5. Na tela **Especificar Regras e Filtros**, clique em **Ações** > **Aplicar uma Regra Existente** no painel **Regras**.

   A caixa de diálogo **Aplicar Assistente de Regra** é exibida.
6. Selecione uma regra e clique em **Avançar**.
7. Clique em **Adicionar**.

   A caixa de diálogo **Escolher colunas para a porta de entrada** é exibida.
8. Selecione um campo e uma coluna de entrada. Clique em **OK**.

As colunas de entrada e saída são exibidas na caixa de diálogo **Aplicar Assistente de Regra**.

9. Na caixa de diálogo **Aplicar Assistente de Regra**, clique em **OK**.

   A regra é exibida na tela **Especificar Regras e Filtros**.

Regras de Expressão

As regras de expressão usam funções de expressão e colunas para definir a lógica da regra. Crie regras de expressão e adicione-as a um perfil na ferramenta Analyst.

Use regras de expressão para alterar ou validar valores para colunas em um perfil. Uma ou mais regras de expressão podem ser criadas para uso em um perfil. Funções de expressão são funções como SQL usadas para transformar dados de origem. A lógica da regra de expressão pode ser criada com os seguintes tipos de funções:

- Caractere
- Conversão
- Limpeza de Dados
- Data
- Codificação
- Financeiro
- Numérico
- Científico
- Especial
- Testar

Processo de Regras de Expressão

Use o assistente de **Perfil** para criar uma regra de expressão e adicioná-la a um perfil.

Use o editor de expressão para adicionar funções de expressão, configurar colunas como entrada para as funções, validar a expressão e configurar o tipo de retorno, precisão e escala.

Depois de criar e validar uma regra de expressão, você pode editar o valor de precisão da coluna da regra de saída. Por padrão, o valor de precisão da coluna da regra de saída está definido como 10. O valor de precisão é truncado quando a coluna da regra de saída excede o valor de precisão definido.

A saída de uma regra de expressão é uma coluna virtual que usa o nome da regra como o nome da coluna. A ferramenta Analyst executa um perfil de coluna na coluna virtual. Por exemplo, use uma regra de expressão para validar um CEP. A regra retornará 1 se o CEP for válido e 0 se o CEP for inválido. O Informatica Analyst executa um perfil de coluna nos valores de saída de 1 e 0 da regra.

Você pode executar as seguintes etapas para criar uma regra de expressão:

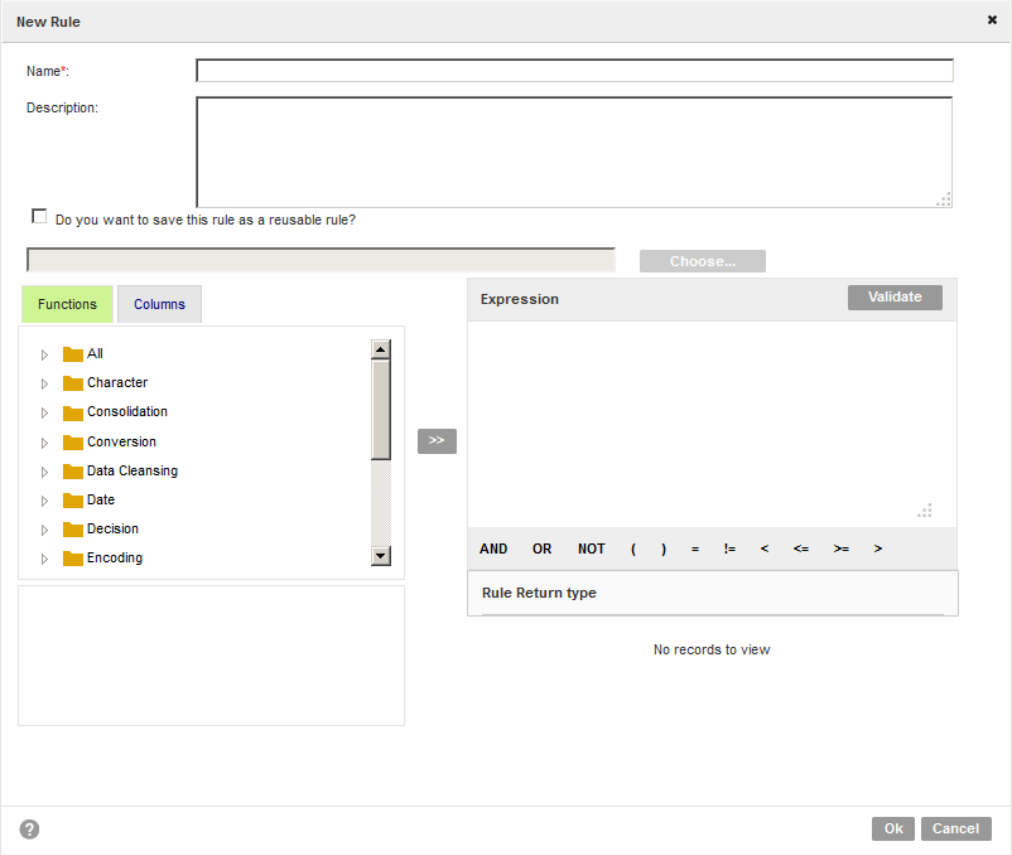
1. Abra um perfil.
2. Configure a lógica da regra usando funções de expressão e colunas como parâmetros.
3. Configure as opções de criação de perfil.

Criando uma Regra de Expressão

Use o assistente de **Perfil** para criar uma regra de expressão e adicioná-la a um perfil. Crie uma regra de expressão para validar valores para colunas em um perfil.

1. Abra um perfil.
2. Na exibição resumida, clique em **Ações** > **Editar Perfil** para abrir o assistente de **Perfil**.
3. Clique em **Especificar Regras e Filtros**.
4. No painel Regras, clique em **Ações** > **Adicionar uma Regra**.

   A caixa de diálogo **Nova Regra** é exibida.



5. Na caixa de diálogo **Nova Regra**, insira um nome e uma descrição opcional para a regra. Você pode criar uma regra no painel Funções ou no painel Colunas.
   - No painel Funções, selecione uma categoria de função e clique no botão de setas para a direita (>>). Na caixa de diálogo, especifique os parâmetros e clique em **OK**.
     A função juntamente com as colunas e valores são exibidas no painel Expressão.
   - No painel Colunas, selecione uma coluna e clique no botão de setas para a direita (>>). A coluna é exibida no painel Expressão. Adicione funções, expressões e valores para criar uma regra.
6. Para verificar a regra, clique em **Validar**.
7. Como alternativa, escolha promover a regra como regra reutilizável e configure a localização do projeto e da pasta. Se uma regra for promovida a regra reutilizável, você ou outros usuários poderão usar a regra em outro perfil como uma regra predefinida.
8. Clique em **OK**.

   A tela **Especificar Regras e Filtros** é exibida com a regra no painel Regras.

CAPÍTULO 6

Resultados do Perfil da Coluna no Informatica Analyst

Este capítulo inclui os seguintes tópicos:

- Resultados de perfil de coluna na visão geral do Informatica Analyst, 34
- Exibição Resumida, 35
- Exibição Detalhada, 37
- Estatísticas, 39
- Tipos de execução de perfil, 46
- Visão geral dos resultados de vários perfis de comparação, 48
- Busca Detalhada de Perfil da Coluna, 52
- Curadoria na ferramenta Analyst, 53
- Arquivos de exportação de perfil de coluna no Informatica Analyst, 54

Resultados de perfil de coluna na visão geral do Informatica Analyst

Exiba os resultados do perfil para compreender e analisar o conteúdo, a estrutura e a qualidade de dados. Você pode exibir todas as colunas e regras em um perfil na exibição resumida. Você pode exibir as propriedades de uma coluna ou regra em detalhes na exibição detalhada.

Você pode exibir os resultados do perfil no espaço de trabalho **Descoberta**. O cabeçalho da exibição exibe o tipo de perfil, o número de colunas no perfil, o número de regras no perfil, os dados de amostragem e a data e hora de criação.

Na exibição resumida, você pode exibir as propriedades de cada coluna como um valor, um gráfico de barras horizontais ou como uma porcentagem. É possível exibir propriedades de colunas, como valores nulos, distintos e não distintos, padrões, tipos de dados e domínios de dados. Você pode exibir os resultados do perfil na exibição resumida com base nos filtros padrão.

Na exibição detalhada, é possível exibir valores nulos, distintos e não distintos, tipos de dados inferidos, domínios de dados inferidos, padrões inferidos, valores e termos comerciais, além de visualizar os dados em painéis.

Você pode exibir os resultados do perfil da execução mais recentes, da execução de histórico e da execução consolidada. Você pode comparar os resultados do perfil de duas execuções de perfil e exibir os resultados

nas exibições resumida e detalhada. Você pode exibir as estatísticas de perfil e fazer a curadoria dos dados. As estatísticas de perfil incluem valores, padrões, tipos de dados, exceções e estatísticas de colunas e regras. É possível realizar descobertas de dados e buscas detalhadas nesses dados.

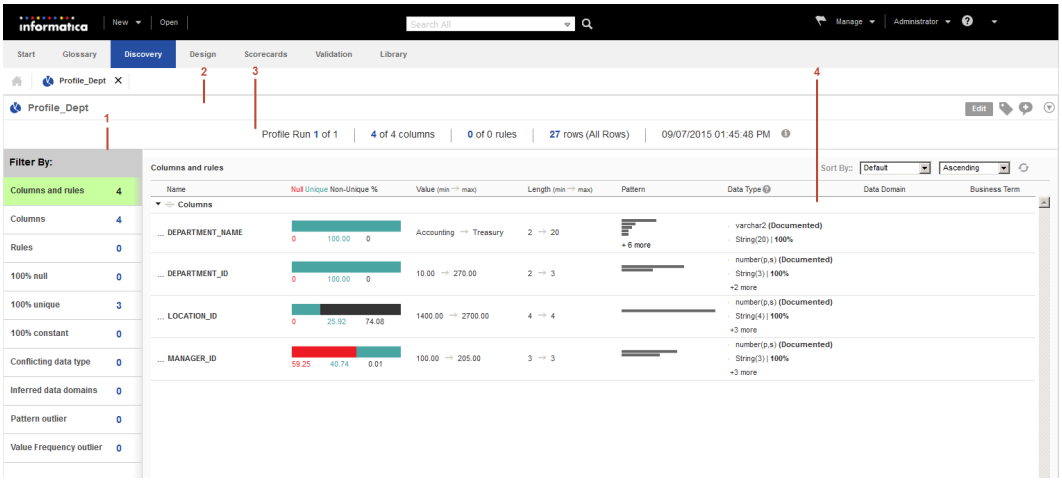
Nota: Você pode exibir e executar um perfil em fontes de dados Avro, JSON, Parquet e XML. Você pode exibir os resultados do perfil de execução mais recentes, execução de histórico e execução consolidada, e comparar os resultados do perfil de duas execuções de perfil.

Você pode exportar frequências de valor, frequências padrão, dados de busca detalhada, comentários, marcas e termos comerciais para um arquivo CSV. Você pode exportar as informações do resumo do perfil para um arquivo do Microsoft Excel de forma a exibir todos os dados em um arquivo para análise futura. Você pode exibir as informações de regra nos resultados do perfil. Os resultados do perfil que são exibidos dependem das opções de configuração do perfil e de amostragem.

Exibição Resumida

O resumo dos resultados do perfil aparecem em formato de grade na exibição resumida. Você pode usar os filtros padrão na exibição resumida para exibir estatísticas específicas. Por exemplo, quando você escolhe Regras, a exibição resumida mostra todas as regras no perfil.

A imagem a seguir apresenta uma exibição gráfica de amostra da exibição resumida:



1. Filtros padrão. Você pode visualizar os resultados do perfil na exibição resumida com base nos filtros padrão.
2. Cabeçalho de perfil. Você pode visualizar o nome do perfil no cabeçalho. Você pode usar o botão Editar para editar o perfil, usar os ícones de marcas e comentários para adicionar ou editar marcas e comentários e escolher as opções no menu Ações.
3. Cabeçalho da exibição resumida. Você pode visualizar informações específico do perfil no cabeçalho da exibição resumida. É possível visualizar o número de execuções de perfil, o número total de execuções de perfil, o número de colunas e regras, bem como o número de linhas no perfil.
4. Exibição resumida. Você pode visualizar as propriedades de todas as colunas e regras do perfil.

Na exibição resumida, é possível executar ou editar o perfil, detectar exceções de padrões ou frequências de valor, adicionar colunas a um scorecard, escolher uma execução de perfil, comparar duas execuções de perfil, exportar resultados de perfil ou resultados de descoberta de domínio de dados para uma planilha do Microsoft Excel, verificar os resultados de inferência de várias colunas, adicionar ou excluir comentários e marcas ou visualizar propriedades de perfil.

Propriedades da exibição resumida

A exibição resumida mostra as propriedades de todas as colunas e regras em um perfil. A exibição resumida inclui uma representação visual das propriedades. Para classificar os valores das propriedades, clique em cada propriedade do resumo.

A seguinte tabela descreve as propriedades do resumo dos resultados do perfil:

Propriedade	Descrição
Nome	Exibe o nome da coluna ou regra no perfil.
% de nulos, distintos e não distintos	Exibe os valores nulos, distintos e não distintos em porcentagens para uma saída de coluna ou regra. Você pode exibir os valores em um gráfico de barras horizontais.
Padrão	Exibe vários padrões na coluna como gráficos de barras horizontais. Você pode exibir os caracteres padrão e o número de padrões semelhantes em uma coluna, como uma porcentagem, ao posicionar o mouse sobre o gráfico de barras.
Valor	Exibe os valores mínimo e máximo na coluna ou regra de saída.
Comprimento	Exibe o tamanho mínimo e máximo dos valores na coluna ou regra de saída.
Tipo de dados	Exibe o tipo de dados documentado da coluna ou regra. Exibe os tipos de dados inferidos ao posicionar o mouse sobre o campo. A ferramenta Analyst pode inferir nos seguintes tipos de dados: - String - Varchar - Decimal - Número inteiro - Data Você também pode exibir a porcentagem de conformidade com base nos tipos de dados inferidos. **Nota:** A ferramenta Analyst não pode gerar o tipo de dados dos valores de uma coluna numérica que tenha uma precisão maior que 38. A ferramenta Analyst não pode gerar o tipo de dados dos valores de uma coluna de string que tenha uma precisão maior que 255. Se você tiver uma coluna de data na qual cria um perfil de coluna com um valor de ano anterior a 1800, o tipo de dados inferido poderá aparecer como uma string de tamanho fixo. Altere o valor padrão para o parâmetro de ano mínimo em InferDateTimeConfig.xml, conforme necessário.
Domínio de dados	Exibe os nomes dos domínios de dados associados à coluna, juntamente com a porcentagem de conformidade e o número de linhas em conformidade.
Termo comercial	Exibe o termo comercial atribuído a coluna.

Filtros padrão na exibição resumida

Você pode exibir os resultados do perfil na exibição resumida com base nos filtros padrão.

A exibição resumida exibe os resultados do perfil para todas as colunas de origem, colunas virtuais e colunas de regra por padrão. O painel Filtrar por exibe o número de colunas na qual você pode aplicar os filtros padrão.

Na exibição resumida, você pode exibir os resultados do perfil usando as seguintes opções de filtro padrão:

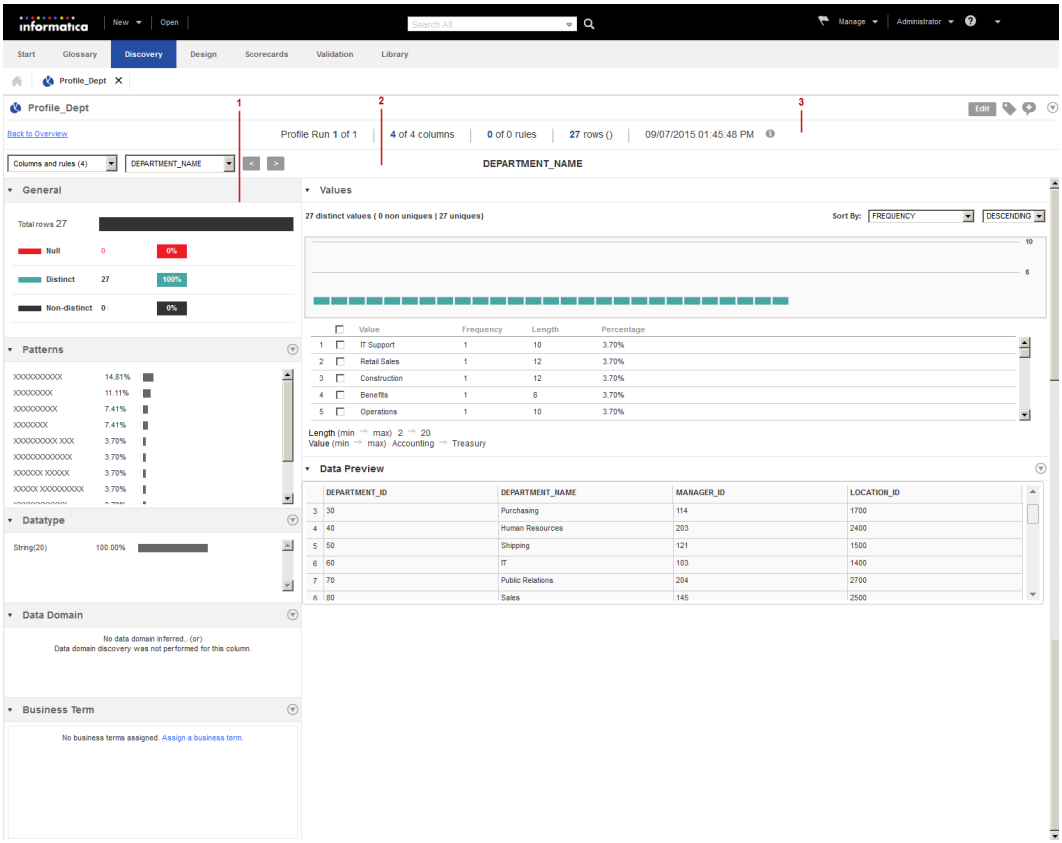
Opção de filtro padrão	Descrição
Colunas e regras	Exibe os resultados do perfil das colunas de origem e de regra. Você pode expandir e recolher as colunas de origem e de regra para exibir os resultados.
Colunas	Exibe os resultados do perfil das colunas de origem.
Regras	Exibe os resultados do perfil das colunas de regra.
100% nulo	Exibe os resultados do perfil das colunas que possuem valores 100% nulos.
100% distinto	Exibe os resultados de perfil para as colunas que possuem valores 100% distintos.
100% constante	Exibe os resultados do perfil de coluna que possui o mesmo valor para todos os registros. Por exemplo, o filtro 100% constante incluirá os resultados do perfil de uma coluna de país se contiver somente um valor "EUA".
Tipos de dados em conflito	Exibe os resultados do perfil das colunas nas quais o tipo de dados documentado e o tipo de dados inferido não correspondem. Por exemplo, o filtro exibe a coluna CustomerTier porque o tipo de dados documentado para a coluna é um número inteiro (2) e o tipo de dados inferido é uma cadeia.
Domínios de dados inferidos	Exibe os resultados do perfil das colunas nas quais o domínio de dados inferido é o mesmo que o domínio de dados configurado.
Exceção padrão	Exibe os resultados do perfil das colunas que possuem exceções padrão.
Exceção de frequência de valor	Exibe os resultados do perfil das colunas que possuem exceções de valor ou de frequência.

Exibição Detalhada

Resultados de colunas aparecem na exibição detalhada. É possível ver as propriedades de colunas em detalhes.

A exibição detalhada de uma coluna aparece depois que você clica nessa coluna na exibição resumida.

A seguinte imagem mostra uma amostra de visualização gráfica das propriedades de colunas na exibição detalhada:



1. Painéis. Você pode ver propriedades gerais, valores na coluna, uma visualização de dados, padrões inferidos, tipos de dados inferidos, domínios de dados inferidos e termos comerciais em painéis.
2. Cabeçalho de detalhes de coluna. Você pode visualizar resultados de coluna selecionando a coluna na lista suspensa ou usando os botões de navegação.
3. Cabeçalho da exibição resumida. Você pode visualizar informações específicas do perfil no cabeçalho da exibição resumida. Você pode ver a execução de perfil, o número de colunas, regras e linhas nessa execução de perfil, bem como a data e a hora do perfil.

Na exibição detalhada, você pode executar ou editar o perfil, adicionar a coluna a um scorecard, escolher uma execução de perfil, comparar duas execuções de perfil, exportar os resultados do perfil para uma planilha do Microsoft Excel, exportar frequências de valores, frequências de padrão, tipos de dados, dados de busca detalhada de valores selecionados ou dados de busca detalhada de padrões selecionados para um arquivo csv, além de adicionar ou excluir comentários e marcas na coluna e visualizar as propriedades do perfil.

Use o menu Ações em cada painel para realizar outras ações nas propriedades da coluna. Você pode recolher ou expandir os painéis.

Painéis da exibição detalhada

A exibição detalhada mostra as propriedades de coluna em painéis, como o número e a porcentagem de valores distintos, não distintos e nulos, os padrões, os tipos de dados inferidos, os domínios de dados inferidos, os valores, a visualização de dados e os termos comerciais vinculados.

Quando for possível clicar na coluna ou na regra, a exibição detalhada da coluna ou da regra será aberta.

A seguinte tabela descreve os painéis na exibição detalhada:

Painéis	Descrição
Geral	Exibe o número de linhas com valores nulos, distintos e não distintos em cores diferentes. Você pode exibir os valores em porcentagens. Você pode exibir o aumento e a diminuição dos valores gerais em cada perfil consecutivo executado como um minigráfico. Um minigráfico exibe a variação no número de valores nulos, valores distintos ou valores não distintos nas últimas cinco execuções de perfil consecutivas em um gráfico de linhas. Você pode exibir o número e a porcentagem de valores ao mover o ponteiro sobre o minigráfico para cada execução de perfil. Você pode adicionar marcas e comentários à coluna.
Padrões	Exibe os padrões dos valores de coluna. A frequência nas quais os padrões são mostrados em uma coluna é exibida como um gráfico de barras horizontais e em porcentagens. Você pode fazer uma busca detalhada em um padrão, adicionar um padrão a uma tabela de referência ou criar um domínio de dados com o padrão selecionado.
Tipo de dados	Exibe os tipos de dados inferidos da coluna. A frequência dos tipos de dados em uma coluna é exibida como um gráfico de barras horizontais e em porcentagens. Você pode fazer uma busca detalhada em um tipo de dados, além de aprovar, rejeitar ou redefinir o tipo de dados inferido selecionado. A opção **Mostrar Rejeitados** exibe os tipos de dados inferidos rejeitados.
Domínio de dados	Exibe os domínios de dados inferidos da coluna. Você pode fazer uma busca detalhada em um domínio de dados para linhas em conformidade, linhas fora de conformidade ou linhas com valores nulos. Você pode aprovar, rejeitar ou redefinir o valor do domínio de dados. A opção **Mostrar Rejeitados** exibe os domínios de dados rejeitados. Você pode verificar o valor do domínio de dados.
Termo comercial	Exibe o termo comercial atribuído à coluna. Você pode atribuir ou remover a atribuição de um termo comercial a uma coluna.
Valores	Exibe todos os valores na coluna em uma representação gráfica junto com a frequência e porcentagem. Você pode fazer uma busca detalhada em cada valor. Você pode adicionar o valor a uma tabela de referência, criar uma regra de frequência de valor e criar um domínio de dados.
Visualização de Dados	Exibe os dados de busca detalhada para o padrão, o tipo de dados, o domínio de dados ou o valor selecionado.

Estatísticas

Você pode exibir estatísticas, como valores, padrões, tipos de dados, domínio de dados e exceções das colunas e regras em um perfil.

Você pode exibir estatísticas de perfil na exibição resumida e exibir estatísticas de coluna nas exibições resumida e detalhada. Você pode exibir estatísticas da execução de perfil mais recente, execução de perfil histórico e execução de perfil consolidado. Você pode comparar os resultados de perfil de duas execuções de perfil e exibir as estatísticas do perfil e das colunas nas exibições resumida e detalhada.

Visualização de Dados

É possível exibir os dados de busca detalhada para o padrão, o tipo de dados, o domínio de dados ou o valor selecionado no painel Visualização de Dados.

Também é possível exibir o painel Visualização de Dados na exibição detalhada. Quando você clicar em uma coluna na exibição resumida, a exibição detalhada é exibida, e o painel Visualização de Dados é recolhido por padrão. Para visualizar os dados da coluna, você pode clicar em **Ações** > **Mostrar Visualização**.

A seguinte tabela descreve as opções do menu **Ações** no painel Visualização de Dados:

Opção	Descrição
Adicionar ao Filtro	Crie um filtro de busca detalhada para filtrar os dados de busca detalhada, de forma que você possa analisar irregularidades de dados nos subconjuntos de resultados do perfil.
Salvar Filtro	Salva o filtro de busca detalhada.
Mostrar Visualização	Exibe as linhas de origem.
Exportar Dados	Exporta os resultados da busca detalhada para um arquivo CSV ou um arquivo do Microsoft Excel.

Tipos de dados

Os tipos de dados incluem todos os tipos de dados inferidos de cada coluna nos resultados do perfil.

Você pode exibir os tipos de dados nas exibições resumida e detalhada. Na exibição resumida, você pode exibir o tipo de dados documentado e os tipos de dados inferidos. O filtro **Conflito no tipo de dados** exibe as colunas nas quais existe um conflito entre o tipo de dados documentado e o inferido. Na exibição detalhada, você pode exibir os tipos de dados inferidos da coluna. A frequência dos tipos de dados em uma coluna é exibida como um gráfico de barras horizontais e em porcentagens. Você pode fazer uma busca detalhada, aprovar, rejeitar ou redefinir o tipo de dados inferido selecionado. A opção Mostrar Rejeitados exibe os tipos de dados inferidos rejeitados.

A seguinte tabela descreve as propriedades para os tipos de dados:

Propriedade	Descrição
Tipo de dados	Exibe a lista de tipos de dados inferidos e documentados para a coluna no perfil.
Frequência	Exibe o número de vezes que um tipo de dados de uma coluna aparece, expresso como um número.
Porcentagem	Exibe a porcentagem na qual o tipo de dados de uma coluna é exibido.

Propriedade	Descrição
Busca detalhada	Faça uma busca detalhada para especificar linhas de origem com base em um tipo de dados de coluna. **Nota:** Você não poderá realizar uma ação de busca detalhada se selecionar vários tipos de dados inferidos.
Status	Indica o status do tipo de dados. Os status são inferido, aprovado ou rejeitado. **Inferido** Indica o tipo de dados da coluna inferido pela ferramenta Analyst. **Aprovado** Indica um tipo de dados aprovado da coluna. Ao aprovar um tipo de dados, você confirma esse tipo de dados no repositório do modelo. **Rejeitado** Indica um tipo de dados rejeitado da coluna.

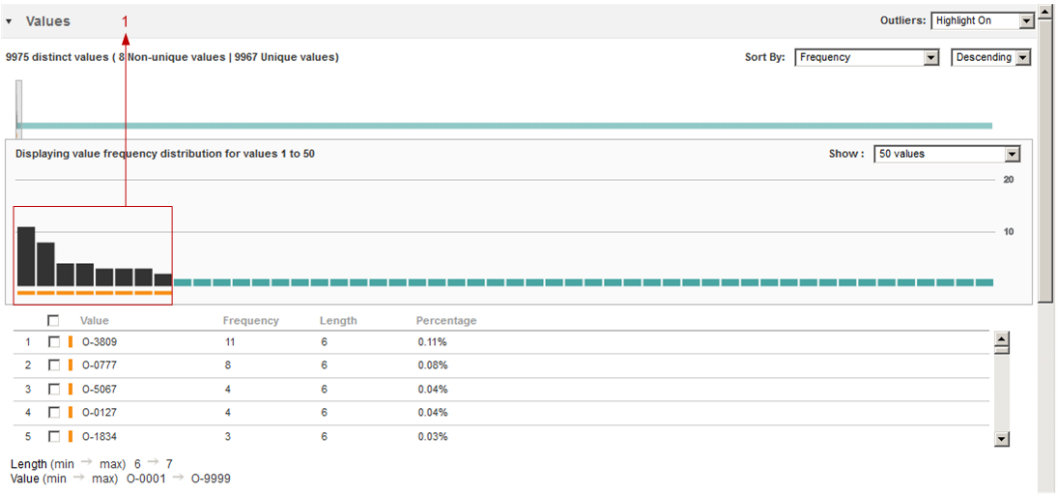
Exceções

Uma exceção é um padrão, um valor ou uma frequência de uma coluna nos resultados do perfil que não estão dentro de um intervalo de valores esperado.

O plug-in de criação de perfil no Serviço de Integração de Dados executa um algoritmo para identificar os valores que não estão dentro do intervalo da maioria dos valores na coluna. Qualquer padrão, valor ou frequência que não estiver dentro do intervalo esperado da maioria desses valores na coluna é uma exceção.

Por padrão, a ferramenta Analyst não determina exceções nos resultados do perfil. Na exibição resumida, você pode executar a exceção para exibir os resultados de exceção. O filtro de exceção de padrão mostra as exceções com base em padrões na coluna. O filtro de exceção de frequência de valor exibe as exceções com base nos valores ou frequências na coluna. A detecção de exceção ocorre no plano de fundo para que você possa realizar outras ações na exibição resumida.

Na exibição detalhada, você pode exibir os valores de exceção no painel Valores ao selecionar a opção **Realçar em** na lista. O valor de exceção aparece como uma barra vertical com sublinhado laranja. Para visualizar apenas o valor de exceção, você deve selecionar a opção **Filtrar** na lista.



1. Valores de exceção. Um valor de exceção aparece como uma barra vertical com sublinhado laranja.

Executando uma exceção

Execute uma exceção para identificar padrões, valores ou frequências em uma coluna que não está dentro de um intervalo de valores esperado.

1. Na exibição resumida, clique em **Ações** > **Detectar Exceção**.

   A exceção de Padrão e a exceção de Frequência de Valor no painel **Filtrar por** muda de N/A para o número de exceções detectadas.
2. No painel **Filtrar por**, clique em **Exceção padrão**.

   As colunas com exceções padrão são exibidas na exibição resumida.
3. No painel **Filtrar por**, clique em **Exceção de Frequência de Valor**.

   As colunas com exceções de valor ou frequência aparecem na exibição resumida.
4. Na exibição detalhada, selecione **Realçar** na lista suspensa de exceção.

   No painel Valores, as exceções são exibidas como barras verticais com sublinhados em laranja.
5. Clique em **Filtro** na lista suspensa Exceções para visualizar apenas os valores de exceções.

Padrões

Você pode exibir os padrões dos valores de coluna e a frequência nas quais os padrões aparecem nas exibições resumida e detalhada.

Na exibição resumida, você pode exibir vários padrões na coluna como gráficos de barras horizontais. Você pode exibir os caracteres padrão e o número de padrões semelhantes em uma coluna, como uma porcentagem, ao posicionar o mouse sobre o gráfico de barras. Na exibição detalhada, você pode exibir a frequência com a qual os padrões aparecem em uma coluna como um gráfico de barras horizontal e em porcentagens. Você pode fazer uma busca detalhada, adicionar o padrão a uma tabela de referência ou criar um domínio de dados com o padrão selecionado.

O depósito da criação de perfil armazena um máximo de 16.000 valores exclusivos da frequência mais alta incluindo valores NULL para resultados do perfil por padrão. Se houver pelo menos um valor NULL nos resultados do perfil, a ferramenta Analyst poderá exibir valores NULL como padrão.

Nota: A ferramenta Analyst não pode gerar o padrão para uma coluna numérica que tenha uma precisão maior que 38. A ferramenta Analyst não pode gerar o padrão para uma coluna de string que tenha uma precisão maior que 255.

A tabela a seguir descreve as propriedades para os padrões de coluna:

Propriedade	Descrição
Padrão	Exibe o padrão da coluna no perfil.
Frequência	Exibe o número de vezes que um padrão de uma coluna aparece, expresso como um número.
Porcentagem	Exibe a porcentagem na qual o padrão de uma coluna é exibido.

A tabela a seguir descreve os caracteres padrão e quais eles representam:

Caractere	Descrição
9	Representa qualquer caractere numérico. O Informatica Analyst exibe até três caracteres separadamente no formato "9". A ferramenta exibe mais de três caracteres como um valor entre parênteses. Por exemplo, o formato "9(8)" representa um valor numérico com oito dígitos.
X	Representa qualquer caractere alfabético. O Informatica Analyst exibe até três caracteres separadamente no formato "X". A ferramenta exibe mais de três caracteres como um valor entre parênteses. Por exemplo, o formato "X(6)" pode representar o valor "Boston". **Nota:** O caractere padrão X não diferencia maiúsculas de minúsculas e pode representar caracteres em letras maiúsculas ou minúsculas dos dados de origem.
p	Representa "(", o parêntese de abertura.
q	Representa ")", o parêntese de fechamento.

Nota: Padrões de coluna também podem incluir caracteres especiais. Por exemplo, ~, [,], =, -, ?, =, {, *, -, >, < e $.

Valores

Você pode exibir os valores das colunas e a frequência em que os valores são exibidos na coluna.

Exiba valores mínimos e máximos em uma coluna na exibição resumida. Na exibição detalhada, você pode exibir as propriedades de valor para uma coluna.

Valores na exibição resumida

Você pode exibir os valores mínimo e máximo de todas as colunas e regras da execução de perfil mais recente, execução de perfil histórico e execução de perfil consolidado na exibição resumida.

Exemplo

Um banco de dados de uma loja de varejo tem uma coluna denominada ID do Funcionário na tabela Funcionário preenchida com IDs de funcionários que variam de 100 a 250 e com nomes como Bob e Robert. Quando você executa um perfil de coluna na tabela Funcionário, a coluna Valor da ID do Funcionário na exibição resumida mostra 100 --> Robert

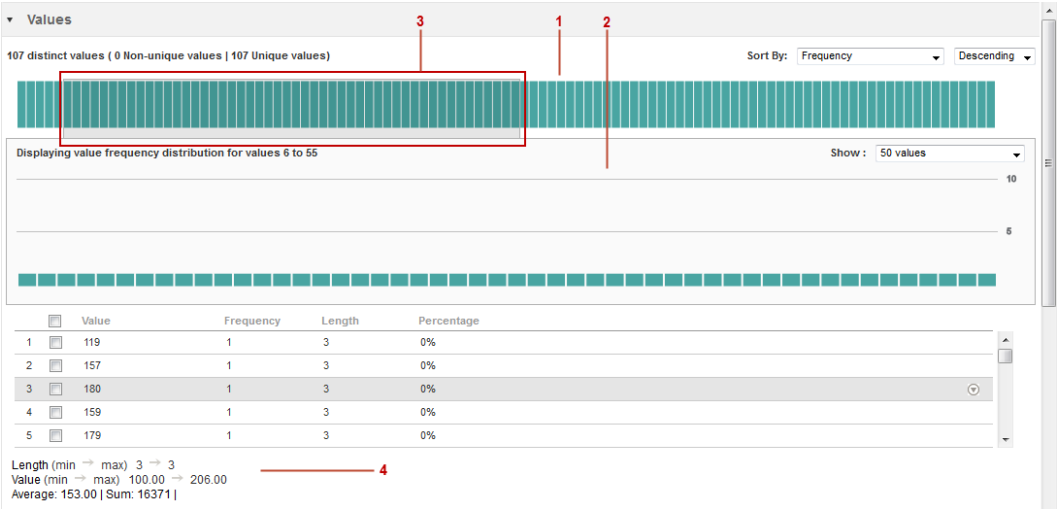
Valores na exibição detalhada

Os valores de coluna na exibição detalhada incluem valores para uma coluna e a frequência na qual os valores são exibidos na coluna.

O painel **Valores** exibe os valores de colunas em uma representação gráfica. Você pode exibir a frequência, o tamanho e a porcentagem de cada valor. Você pode classificar os valores com base no valor ou na frequência. Você pode fazer uma busca detalhada nos dados, adicionar os valores a uma tabela de referência, criar uma regra de frequência de valor ou criar um domínio de dados. Você pode exibir os valores nulos como uma barra vertical vermelha, a frequência de valores como uma barra vertical preto e os valores de exceção como barras verticais com destaque em laranja. Você pode realçar os valores de exceção, desativar exceções ou filtrar os resultados para exibir apenas valores de exceção na coluna.
O painel Valores contém o layout gráfico e as seções de valor.

O layout gráfico é dividido em dois painéis.

A seguinte imagem mostra o painel Valores na exibição detalhada:



1. Painel superior. Você pode exibir os valores como um gráfico de barras verticais. Você pode classificar os valores de frequência e valor. Você pode classificar os valores em ordem crescente ou decrescente. Você pode exibir os valores de exceção como barras verticais com realces em laranja.
2. Painel inferior. Você pode exibir os valores no controle deslizante do painel inferior, onde cada valor é representado por uma barra vertical. Você pode fazer uma busca detalhada no valor, adicionar esse valor a uma tabela de referência, criar uma regra de frequência de valor ou criar um domínio de dados no valor. É possível exibir 50, 75 ou 100 valores de uma só vez.
3. Controle deslizante. Você pode deslizar o controle deslizante sobre os valores no painel superior. O painel inferior mostra os valores no controle deslizante.
4. Propriedades do valor. A seção de propriedades do valor exibe os valores e as propriedades.

A seguinte tabela descreve os painéis no layout gráfico:

Painel	Descrição
Painel superior	Exibe todos os valores como um gráfico de barras verticais. É possível exibir um máximo de 16.000 valores no painel superior. Você pode usar o controle deslizante para visualizar um lote de valores.
Painel inferior	Exibe os valores do lote selecionados no painel superior. Por padrão, a ferramenta Analyst exibe 50 valores. Você pode optar por exibir 75 ou 100 valores por vez.

A seguinte tabela descreve as propriedades dos valores de coluna na seção de valor:

Propriedade	Descrição
Valor	Exibe uma lista de valores do lote selecionado no painel superior. **Nota:** A ferramenta Analyst exclui o CLOB, o BLOB, o Raw e os tipos de dados binários em valores de coluna.
Frequência	Exibe o número de vezes que um valor aparece na coluna, expresso como um número.
Comprimento	Exibe o tamanho do valor de coluna.
Porcentagem	Exibe a porcentagem que um valor aparece na coluna.

A seguinte tabela descreve as estatísticas da coluna selecionada:

Estatísticas	Descrição
Length (min - max)	Exibe o comprimento do menor valor e do maior valor da coluna.
Valor (mín - máx)	Exibe os valores mínimo e máximo na coluna.
Média	Exibe a média dos valores de coluna.
Soma	Exibe a soma de todos os valores na coluna.

Valores na exibição detalhada de comparação de resultados de perfil

O painel Valores na exibição detalhada de comparação de resultados de perfil exibe as propriedades de valor, como o número de valores diferentes, o valor mínimo, o valor máximo, o tamanho mínimo e máximo, a média, o desvio padrão e a soma dos valores.

A exibição detalhada de uma coluna de comparação de resultados de perfil exibe as propriedades de valor, o valor e a frequência do valor com um gráfico de barras horizontais.
A tabela a seguir descreve as propriedades dos valores de coluna na exibição detalhada quando você compara os resultados de duas execuções de perfil.:

Propriedade	Descrição
Nº de valores distintos	Exibe o número de valores distintos na coluna.
Valor mín.	Exibe o valor mínimo na coluna.
Valor máx.	Exibe o valor máximo na coluna.
Comprimento (mín. - máx.)	Exibe o tamanho do menor valor e maior valor da coluna.
Média	Exibe a média dos valores de coluna.

Propriedade	Descrição
Desvio padrão	Exibe o desvio padrão ou a variabilidade entre os valores de coluna de todos os valores de coluna.
Soma	Exibe a soma de todos os valores na coluna.

Tipos de execução de perfil

Você pode exibir os resultados do perfil da execução de perfil mais recente, execução de perfil histórico e execução de perfil consolidado. Você pode exibir os resultados de execução de perfil na exibição resumida.

Execução de perfil mais recente

Exiba os resultados do perfil da execução de perfil mais recente na exibição resumida.

Você pode visualizar os resultados de perfil da última execução de perfil na exibição resumida ao:

- Criar, salvar e executar um perfil.
- Abrir um perfil anteriormente executado no espaço de trabalho **Biblioteca**.
- Clicar no link **Voltar para Última Execução de Perfil** na exibição resumida ou detalhada da execução de perfil consolidada.
- Clicar no link **Voltar para Última Execução de Perfil** na exibição resumida ou detalhada de uma execução de perfil histórica.
- Selecionar a execução de perfil mais recente na caixa de diálogo **Selecionar Execução do Perfil** e clicar em **OK**.

Execução de perfil histórico

Exiba os resultados de perfil de uma execução de perfil anterior na exibição resumida.

O depósito de criação de perfil salva os resultados de perfil de todas as execuções de um perfil. Você pode escolher exibir os resultados de uma versão anterior da execução de perfil selecionando-a na caixa de diálogo Selecionar Execução de Perfil.

Execução de perfil consolidada

Exiba os resultados do perfil mais recentes de cada coluna no perfil na exibição resumida.

Na execução de perfil consolidado, você pode exibir os resultados mais recentes de cada coluna no perfil. Quando você escolhe a execução de perfil Consolidada na caixa de diálogo **Selecionar Execução de Perfil**, o warehouse de criação de perfil recupera os resultados de colunas mais recentes de todas as execuções do perfil. Você pode exibir os resultados na exibição resumida e o cabeçalho da exibição resumida exibe a execução de perfil incremental.

Exemplo

Como um analista de dados, você pode exibir os resultados mais recentes de cada coluna em um perfil. Por exemplo, você pode escolher as colunas 1, 2 e 3 para realizar a execução do perfil A e escolher as colunas 3,

4 e 5 para a execução do perfil B. Para exibir os resultados mais recentes de todas as colunas, você pode escolher a execução de perfil consolidado na caixa de diálogo Selecionar Execução de Perfil. A exibição resumida mostra os resultados das colunas 1 e 2 da execução A e exibe os resultados das colunas 3, 4 e 5 da execução B.

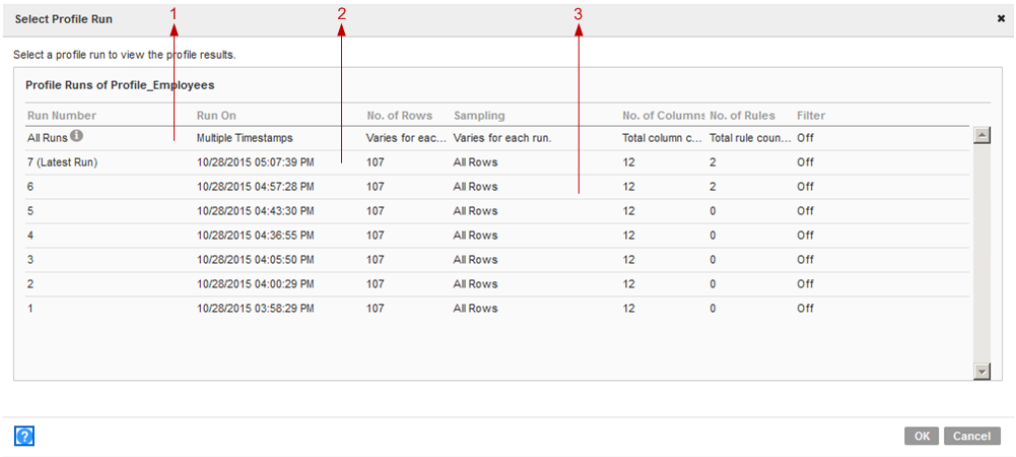
Selecionando uma execução de perfil

Você pode selecionar uma execução de perfil histórico, uma execução de perfil mais recente ou uma execução de perfil consolidado para exibir os resultados do perfil. Você pode exibir os resultados do perfil na exibição resumida e os resultados de coluna na exibição detalhada.

1. No espaço de trabalho **Biblioteca**, selecione o projeto ou a pasta que contém o perfil ou selecione o perfil no painel **Ativos**.
2. Clique em **Ações** > **Abrir** para abrir o perfil.

   A exibição resumida é exibida no espaço de trabalho **Descoberta**.
3. Na exibição resumida, clique em **Ações** > **Escolha a Execução de Perfil**.

   A caixa de diálogo **Selecionar Execução de Perfil** é exibida.
   A seguinte imagem mostra a caixa de diálogo **Selecionar Execução de Perfil**.



   1. Execução de perfil consolidado. Ao escolher essa execução de perfil, você pode visualizar os resultados de perfil mais recentes para cada coluna na exibição resumida.
   2. Execução de perfil mais recente. Ao escolher essa execução de perfil, você pode exibir os resultados do perfil mais recentes na exibição resumida.
   3. Execução de perfil histórico. Ao escolher essa execução de perfil, você pode exibir os resultados do perfil histórico de uma execução de perfil anterior na exibição resumida.
4. Na caixa de diálogo **Selecionar Execução de Perfil**, selecione uma das execuções de perfil para visualizar seus resultados de perfil:
   - Para visualizar os resultados de perfil da execução de perfil mais recente, selecione a execução de perfil mais recente e clique em **OK**.
   - Para visualizar os resultados de perfil para uma execução de perfil histórica, selecione uma execução de perfil diferente da mais recente e clique em **OK**.
   - Para visualizar os resultados de perfil de uma execução de perfil consolidada, selecione **Todas as Execuções** e clique em **OK**. Os resultados de perfil mais recentes de cada coluna são mostrados na exibição resumida.

A ferramenta Analyst realiza uma execução de perfil e mostra os resultados do perfil na exibição resumida.

5. Na exibição resumida, clique em uma coluna para exibir os resultados da coluna.

   A exibição detalhada é exibida.

Visão geral dos resultados de vários perfis de comparação

Você pode comparar os resultados de perfil de duas execuções de perfil. Você pode exibir os resultados de comparação na exibição resumida e os resultados de coluna na exibição detalhada.

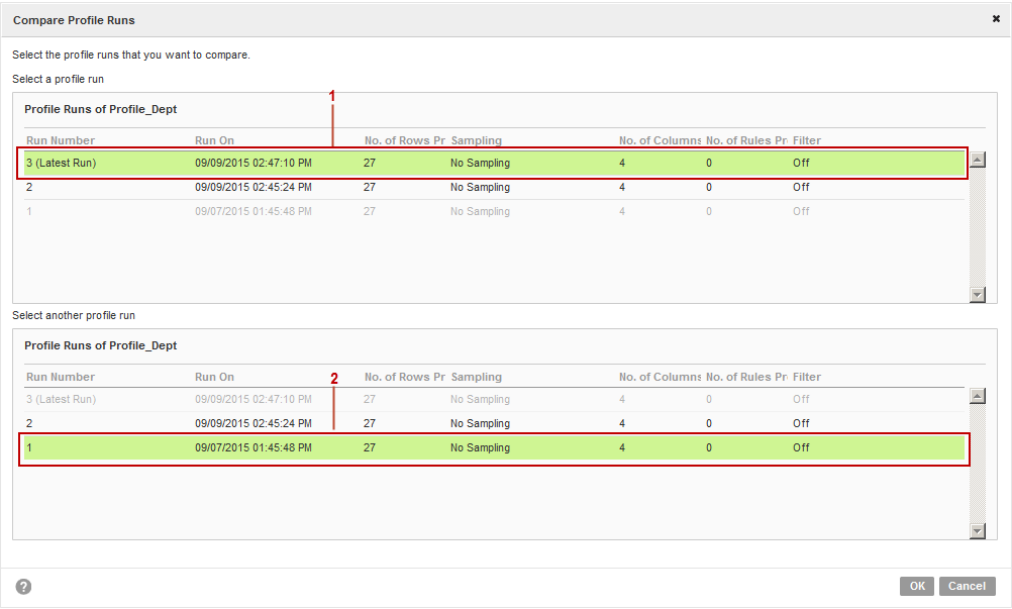
Na exibição resumida, você pode exibir os resultados de comparação de todas as colunas nas execuções de perfil.

Comparando resultados de vários perfis

Ao comparar duas execuções de perfil, você pode exibir a comparação de resultados do perfil na exibição resumida.

1. Na exibição resumida, clique em **Ações** > **Execução de Perfil de Comparação**.

   A seguinte imagem mostra a caixa de diálogo **Comparar Execuções de Perfil**.



   1. Execução A. Escolha uma execução de perfil como Execução A.
   2. Execução B. Escolha uma execução de perfil como Execução B.

   A caixa de diálogo **Execuções de Perfil de Comparação** é exibida.

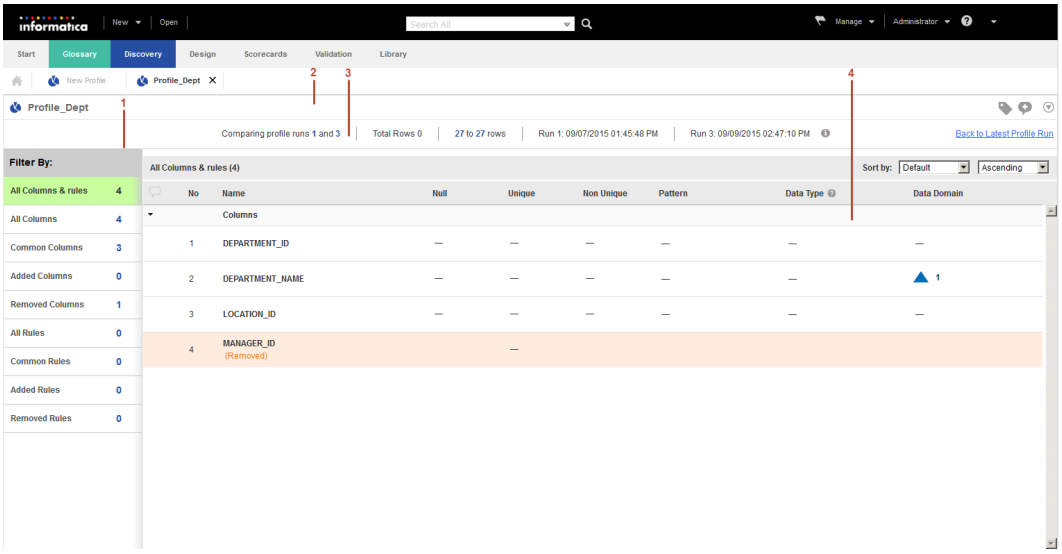
2. Selecione um perfil no painel **Execução A** e selecione outro perfil no painel **Execução B**.
3. Clique em **OK**.

   A exibição resumida mostra uma exibição consolidada dos resultados do perfil.

Exibição Resumida dos Resultados de Perfil de Comparação

Quando você compara duas execuções de perfil, você pode exibir os resultados em um formato de grade na exibição resumida. Você pode usar os filtros padrão na exibição resumida para exibir estatísticas específicas.

A seguinte imagem mostra a comparação de resultados de perfil para duas execuções de perfil na exibição resumida:



1. Filtro padrão. Você pode visualizar os resultados da comparação de perfis na exibição resumida com base nos filtros padrão.
2. Cabeçalho de perfil. Você pode visualizar o nome do perfil no cabeçalho.
3. Cabeçalho da exibição resumida. Você pode visualizar informações específicas do perfil no cabeçalho da exibição resumida. Você pode visualizar as execuções de perfil que são comparadas, o aumento ou a diminuição em linhas entre as execuções de perfil, o número de linhas no perfil e a data e a hora das execuções de perfil.
4. Exibição resumida. Você pode visualizar a comparação entre as colunas em ambas as execuções de perfil.

Propriedades da exibição resumida para comparação de resultados do perfil

As propriedades da exibição resumida para a comparação de resultados de perfil incluem o número e a porcentagem de valores distintos, não distintos e nulos, padrões, tipos de dados inferidos, domínios de dados inferidos e termos comerciais vinculados. A exibição resumida inclui uma representação visual das propriedades. Para classificar os valores das propriedades, clique em cada propriedade do resumo.

Na exibição resumida, o Serviço de Integração de Dados atribui um número em ordem crescente a todas as colunas e regras

Nota: Uma seta para cima com uma contagem numérica exibe um aumento nos valores de uma propriedade de uma execução de perfil a outra. Uma seta para baixo com uma contagem numérica exibe uma diminuição nos valores de uma propriedade de uma execução de perfil a outra.

A seguinte tabela descreve as propriedades do resumo dos resultados do perfil de comparação:

Propriedade	Descrição
Não	Exibe o número da coluna ou regra.
Nome	Exibe o nome da coluna ou regra no perfil.
Nulo	Exibe o aumento ou diminuição em valores nulos.
Distinto	Exibe o aumento ou a diminuição em valores distintos.
Não distinto	Exibe o aumento ou a diminuição em valores não distintos.
Padrão	Exibe a variação em padrões entre as execuções de perfil.
Tipo de dados	Exibe a variação entre os tipos de dados inferidos da coluna ou regra nas duas execuções de perfil.
Domínio de dados	Exibe a variação entre os domínios de dados inferidos associados à coluna ou regra nas duas execuções de perfil.

Filtros padrão para a comparação de resultados de perfil na exibição resumida

Você pode exibir os resultados de perfil com base no filtros padrão na exibição resumida.

Na exibição resumida, você pode exibir os resultados de perfil usando as seguintes opções de filtro padrão:

Opção de filtro padrão	Descrição
Todas as colunas e regras	Exibe os resultados de perfil das colunas de origem, colunas virtuais e colunas de regra. Você pode expandir e recolher as colunas de origem e de regra para exibir os resultados.
Todas as colunas	Exibe os resultados de perfil das colunas de origens e virtuais.
Colunas comuns	Exibe as colunas disponíveis em ambos resultados da execução de perfil.
Colunas adicionadas	Exibe as colunas disponíveis na execução de perfil mais recente. Por exemplo, quando você compara a execução 5 com a execução 3, as colunas adicionadas exibem as colunas disponíveis na execução 5 e não na execução 3.
Colunas removidas	Exibe as colunas disponíveis na execução de perfil histórico. Por exemplo, quando você compara a execução 5 com a execução 3, as colunas removidas exibem as colunas disponíveis na execução 3 e não na execução 5.
Todas as regras	Exibe os resultados de perfil para todas as colunas de regra.
Regras adicionadas	Exibe as regras disponíveis na execução de perfil mais recente. Por exemplo, quando você compara a execução 5 com a execução 3, as regras adicionadas exibem as regras disponíveis na execução 5 e não na execução 3.
Regras removidas	Exibe as regras disponíveis na execução de perfil histórico. Por exemplo, quando você compara a execução 5 com a execução 3, as regras removidas exibem as regras disponíveis na execução 3 e não na execução 5.

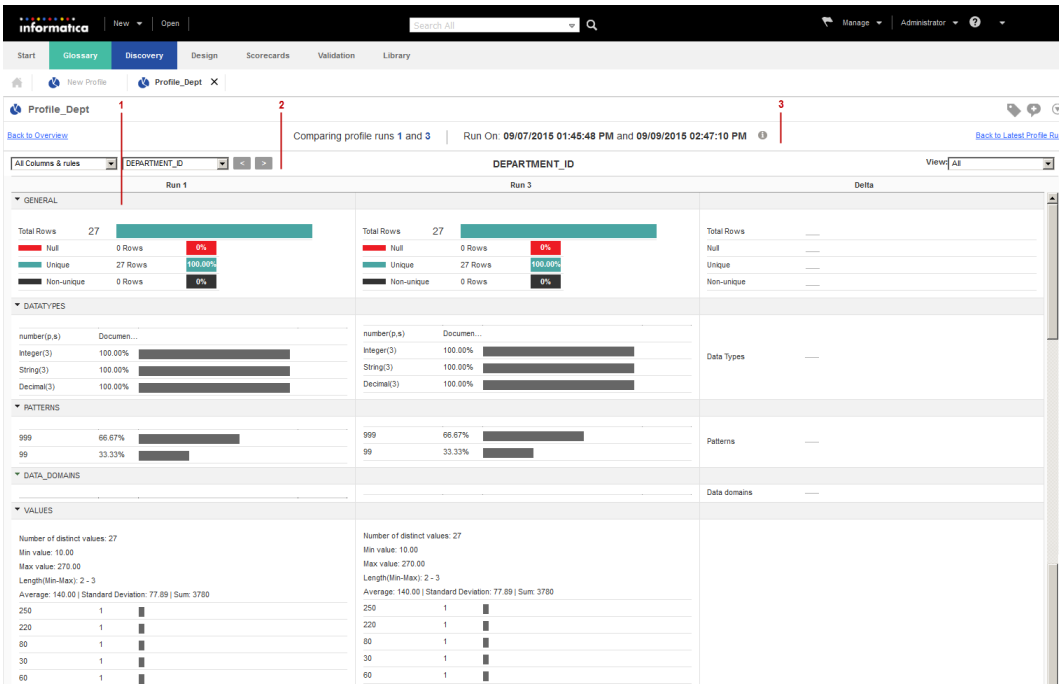
A exibição resumida exibe os resultados de perfil para todas as colunas de origem, colunas virtuais e colunas de regra por padrão.

Exibição detalhada dos resultados do perfil de comparação

Os resultados de coluna são exibidos em um formato de grade na exibição detalhada. Os detalhes da coluna incluem informações gerais, como valores distintos, não distintos e nulos, padrões, tipos de dados, domínios de dados, termos comerciais, valores e a visualização de dados.

A exibição detalhada de uma coluna é exibida quando você clica no nome da coluna. Você pode exibir os resultados das colunas na execução A e na execução B como colunas separadas, e a comparação de dados está disponível na coluna delta.

A seguinte imagem mostra a comparação de resultados de perfil para uma coluna na exibição detalhada:



1. Painéis. Você pode visualizar os resultados de perfil para a coluna nas duas execuções de perfil e visualizar as informações delta para a coluna nas duas execuções de perfil em painéis.
2. Cabeçalho de perfil. Você pode visualizar resultados de coluna selecionando a coluna na lista suspensa ou usando os botões de navegação. Você pode visualizar o nome da coluna e resultados específicos usando as opções na lista suspensa Exibição.
3. Cabeçalho da exibição resumida. Você pode visualizar informações específicas do perfil no cabeçalho da exibição resumida. Você pode visualizar as execuções de perfil que são comparadas, bem como a data e a hora dessas execuções.

Painéis da exibição detalhada para comparação de resultados do perfil

A exibição detalhada exibe os resultados do perfil e de comparação de uma coluna nas duas execuções de perfil em detalhes.

A exibição detalhada exibe os resultados de coluna da execução A e execução B, e a comparação de dados está disponível na coluna delta. Para exibir outros resultados de coluna, você pode selecionar um filtro na lista suspensa de filtro ou selecionar a coluna na lista suspensa de coluna.

Busca Detalhada de Perfil da Coluna

Use as opções de busca detalhada em um perfil de coluna para realizar uma busca detalhada em linhas específicas na fonte de dados com base em um valor de coluna. Você pode escolher a leitura dos dados atuais em uma fonte de dados para busca detalhada ou a leitura de dados de fase no depósito de perfil. Quando você fizer uma busca detalhada em uma linha específica no perfil de dados de fase, a ferramenta Analyst cria um filtro de busca detalhada para um valor de coluna correspondente. Depois de fazer uma busca detalhada, você pode editar, rechamar, redefinir e salvar o filtro de busca detalhada.

É possível selecionar as colunas de busca detalhada, mesmo se você não selecionou essas colunas para a criação de perfil. Você pode escolher a leitura dos dados atuais em uma fonte de dados para busca detalhada ou a leitura de dados de fase no depósito de criação de perfil. Depois de realizar uma busca detalhada em um valor de coluna, você pode exportar os dados da busca detalhada dos valores ou padrões selecionados para um arquivo CSV em uma localização escolhida. Embora o Informatica Analyst exiba os primeiros 200 valores para os dados de busca detalhada, a ferramenta exporta todos os valores para o arquivo CSV.

Fazer Busca Detalhada na Linha de Dados

Depois de executar um perfil, você pode fazer uma busca detalhada em linhas específicas que correspondam ao valor da coluna, tipo de dados ou padrão.

1. Execute um perfil.

   Os resultados do perfil aparecem na exibição resumida.
2. Na exibição resumida, clique em um nome de coluna.

   Os resultados de coluna aparecem na exibição detalhada.
3. Na exibição detalhada, clique com o botão direito do mouse no painel **Valores** e selecione **Busca Detalhada**.

   O painel **Visualização de Dados** exibe os dados de busca detalhada.

Aplicar Filtros aos Dados de Busca Detalhada

É possível filtrar os dados da busca detalhada iterativamente de forma que você possa analisar irregularidades nos dados nos subconjuntos de resultados do perfil.

1. Selecione um valor de coluna na guia **Valores**.
2. Clique com o botão direito do mouse e selecione **Busca Detalhada**.

   Os resultados da busca detalhada aparecem no painel **Visualização de Dados**.
3. Para adicionar uma condição de filtro, clique com o botão direito do mouse em valor de coluna no painel **Visualização de Dados** e selecione **Adicionar ao Filtro**.

   A caixa de diálogo **Filtro de Busca Detalhada** é exibida com a condição de filtro.
4. Adicione as condições de filtro necessárias e clique em **OK**.

   Não é possível aplicar filtros de busca detalhada a tipos de dados inferidos.
5. Para salvar o filtro, clique em **Ações** > **Salvar Filtro**.
6. Para limpar os filtros de busca detalhada, clique em **Ações** > **Atualizar**.
7. Para exportar dados de busca detalhada para uma planilha do Microsoft Excel, clique em **Ações** > **Exportar Dados**.

Curadoria na ferramenta Analyst

Curadoria é o processo de validação e gerenciamento dos metadados descobertos em uma fonte de dados de forma que os metadados estejam prontos para uso e preparação de relatórios. Quando você faz curadoria de metadados na ferramenta Analyst, pode aprovar, rejeitar e redefinir os tipos de dados ou domínios de dados inferidos nos resultados de perfil.

Você pode aprovar um tipo de dados e um domínio de dados de uma coluna. Você pode ocultar os tipos de dados rejeitados ou os domínios de dados para uma coluna. Depois de aprovar ou rejeitar um tipo de dados ou domínio de dados inferido, você pode redefinir o tipo de dados ou domínio de dados para restaurar o status inferido.

Aprovando tipos de dados e domínios de dados

Os resultados do perfil incluem os tipos de dados inferidos e os domínios de dados para cada coluna na fonte de dados. Você pode escolher e aprovar um único tipo de dados e um único domínio de dados para cada coluna na ferramenta Analyst.

1. Execute um perfil.

   Os resultados do perfil são exibidos na exibição resumida.
2. Na exibição resumida, clique em um nome de coluna.

   Os resultados de coluna são exibidos na exibição detalhada.
3. Na exibição detalhada, selecione um tipo de dados no painel **Tipo de Dados** ou um domínio de dados no painel **Domínio de Dados**.
4. Clique em **Ações** > **Aprovar**.
5. Para restaurar o status inferido do tipo de dados ou do domínio de dados, selecione o tipo de dados ou o domínio de dados e clique em **Ações** > **Redefinir**.

Rejeitando tipos de dados e domínios de dados

Na exibição detalhada, você pode rejeitar um tipo de dados ou domínio de dados. Você pode mostrar ou ocultar os tipos de dados e os domínios de dados rejeitados.

1. Execute um perfil.

   Os resultados do perfil são exibidos na exibição resumida.
2. Na exibição resumida, clique em um nome de coluna.

   Os resultados de coluna são exibidos na exibição detalhada.
3. Na exibição detalhada, selecione um tipo de dados no painel **Tipo de Dados** ou um domínio de dados no painel **Domínio de Dados**.
4. Clique em **Ações** > **Rejeitar**.

   A ferramenta Analyst remove o tipo de dados rejeitado da lista de tipos de dados.
5. Para exibir os tipos de dados rejeitados, clique em **Ações** > **Mostrar Rejeitados**.

Arquivos de exportação de perfil de coluna no Informatica Analyst

Você pode exportar resultados de perfis da coluna para um arquivo CSV ou Microsoft Excel com base na sua escolha de uma parte dos resultados do perfil ou na execução dos resultados completos de resumo.

Você pode exportar frequências de valor, frequências padrão, tipos de dados ou dados de busca detalhada dos valores e padrões selecionados para um arquivo CSV. Também pode exportar o resumo dos resultados da criação de perfil para todas as colunas para um arquivo do Microsoft Excel. Use o privilégio do Serviço de Integração de Dados **Busca Detalhada e Exportar Resultados** para determinar, por usuário ou grupo, quem exporta os resultados do perfil.

Resultados de Exportação de Perfil em um Arquivo CSV

Você pode exportar as frequências de valor, as frequências padrão, os tipos de dados ou os dados de busca detalhada para exibir os dados em um arquivo. A ferramenta Analyst salva as informações em um arquivo CSV.

Quando os padrões de coluna inferidos são exportados, a ferramenta Analyst exporta um formato diferente do padrão de coluna. Por exemplo, quando você exporta um padrão de coluna inferida X (5), a ferramenta Analyst exibirá o seguinte formato de padrão de coluna do arquivo CSV: XXXXX.

Resultados de Exportação de Perfil no Microsoft Excel

Quando você exporta o resumo completo dos resultados do perfil, a ferramenta Analyst salva as informações em várias planilhas em um arquivo do Microsoft Excel. A ferramenta Analyst salva o arquivo no formato "xlsx".

A tabela a seguir descreve as informações exibidas em cada planilha no arquivo de exportação:

Guia	Descrição
Perfil de coluna	Informações de resumo exportadas da exibição resumida após a execução do perfil. Exemplos são nomes de coluna, nomes de regra, número de valores distintos, número de valores nulos, tipos de dados inferidos e a data e a hora da última execução de perfil.
Valores	Os valores das colunas e regras e a frequência na qual os valores são exibidos para cada coluna.
Padrões	Padrões de valor das colunas e regras que você executou no perfil e a frequência na qual os padrões são exibidos.
Tipos de dados	Todos os tipos de dados da coluna, a frequência de cada tipo de dados, o valor de porcentagem e o status do tipo de dados, como inferido, aprovado ou rejeitado.
Estatísticas	As estatísticas de cada coluna e regra. Exemplos são médio, tamanho, valores superiores, valores inferiores e desvio padrão.
Propriedades	As informações de exibição de propriedades, incluindo nome de perfil, tipo, diretiva de amostragem e número de linhas.

Exportar Resultados de Perfil do Informatica Analyst

Você pode exportar os resultados de um perfil para um arquivo ".csv" ou ".xlsx" para exibir os dados em um arquivo.

1. No espaço de trabalho **Biblioteca**, selecione o projeto ou a pasta que contém o perfil.
2. Clique no perfil para abri-lo.

   Os resultados do perfil aparecem na exibição resumida.
3. Na exibição resumida, clique em **Ações** > **Exportar Dados**.

   A caixa de diálogo **Exportar dados para um arquivo** é exibida.
4. Na caixa de diálogo **Exportar dados para um arquivo**, insira um nome de arquivo. Como alternativa, use o nome de arquivo padrão.
5. Selecione **Todos (Resumo, Valores, Padrões, Estatísticas, Propriedades)** ou **Resultados da descoberta de domínio de dados** e selecione uma **Página de Código**. Clique em **OK**.

   Os dados são exportados para uma planilha do Microsoft Excel.
6. Clique em uma coluna na exibição resumida.

   Os resultados de coluna aparecem na exibição detalhada.
7. Na exibição detalhada, clique em **Ações** > **Exportar Dados**.

   A caixa de diálogo **Exportar dados para um arquivo** é exibida.
8. Na caixa de diálogo **Exportar dados para um arquivo**, insira um nome de arquivo. Como alternativa, use o nome de arquivo padrão.
9. Selecione uma das seguintes opções:
   - Todos (Resumo, Valores, Padrões, Estatísticas, Propriedades)
   - Frequências de valor para a coluna selecionada.
   - Frequências padrão para a coluna selecionada.
   - Os tipos de dados da coluna selecionada.
   - Busca detalhada de dados nos valores selecionados.
   - Busca detalhada de dados nos padrões selecionados.
   - Busca detalhada para os tipos de dados selecionados.
10. Insira um formato de arquivo. O formato é **Excel** para a opção **Todos** e **CSV** para o restante das opções. Você pode optar por exportar o nome do campo como uma primeira linha do arquivo.
11. Selecione a página de código do arquivo
12. Clique em **OK**.

    Os dados são exportados para o arquivo.

CAPÍTULO 7

Scorecards no Informatica Analyst

Este capítulo inclui os seguintes tópicos:

- Scorecards na Visão Geral do Informatica Analyst, 56
- Processo de Scorecard no Informatica Analyst, 57
- Criando um scorecard no Informatica Analyst, 58
- Adicionando colunas a um scorecard existente, 59
- Executar um Scorecard, 60
- Exibir um Scorecard, 61
- Editar um Scorecard, 61
- Métricas, 62
- Grupos de Métrica, 63
- Buscas Detalhadas em Colunas, 65
- Gráficos de Tendência, 65
- Arquivos de Exportação de Scorecard no Informatica Analyst, 69
- Notificações de Scorecard, 70
- Linhagem de Scorecard, 72

Scorecards na Visão Geral do Informatica Analyst

Um scorecard é a representação gráfica de valores válidos para uma coluna em um perfil. Você pode criar scorecards e fazer uma busca detalhada nos dados ativos ou dados de fase.

Use scorecards para medir o progresso da qualidade de dados. Por exemplo, é possível criar um scorecard para medir a qualidade de dados antes de aplicar regras de qualidade de dados. Depois de aplicar regras de qualidade de dados, crie outro scorecard para comparar o efeito das regras de qualidade de dados.

Os scorecards exibem a frequência de valor para colunas como pontuações. As pontuações refletem a porcentagem de valores válidos nas colunas. Depois de executar um perfil, você pode adicionar colunas a partir do perfil como métricas para um scorecard. Pode também criar grupos de métrica de forma que você possa métricas relacionadas para uma única entidade. Você pode definir limites que especifiquem a faixa de dados inválidos aceitável para colunas em um registro e atribuir pesos de métrica para cada métrica. Quando você executa um scorecard, a ferramenta Analyst gera os valores médios ponderados para cada

grupo de métrica. Para avaliar melhor a qualidade de dados, você também pode atribuir um custo fixo ou variável a cada métrica. Quando você executar o scorecard, a ferramenta Analyst calculará a soma do custo de dados inválidos para cada métrica e exibirá o custo total.

Ao criar ou editar um scorecard, você pode criar filtros de scorecard com base nos dados de origem. Esses filtros de scorecard permitem recalcular pontuações de métricas com base na condição de filtragem. Para identificar registros de dados válidos e registros que não são válidos, faça uma busca detalhada em cada métrica. É possível usar gráficos de tendências para controlar como as pontuações de métricas e o custo de dados inválidos em métricas mudam no decorrer de um período de tempo. Você pode reutilizar os filtros de perfil em um scorecard.

Ao criar ou editar um scorecard, você pode escolher Nativo ou Hadoop como o ambiente de tempo de execução.

Quando o sistema de controle de versão está ativado na ferramenta Analyst, você pode criar várias versões de um scorecard e visualizar o histórico de versões de um scorecard. Por padrão, o scorecard é colocado em check-out após a sua criação. Você deve fazer check-in do scorecard para que outros usuários possam editá-lo.

Você pode exibir os detalhes do scorecard no painel de scorecards.

Você pode configurar e gerenciar notificações por e-mail para scorecards no Informatica Analyst. Use o Serviço de E-mail para gerenciar as notificações por e-mail. O Serviço de E-mail é um serviço do sistema que você pode configurar no Informatica Administrator.

Processo de Scorecard no Informatica Analyst

Você pode criar e editar um scorecard na Developer tool e na ferramenta Analyst. Um scorecard pode ser executado na ferramenta Analyst. Você pode executar o scorecard nos dados atuais no objeto de dados ou nos dados preparados no depósito de criação de perfil.

Você pode exibir um scorecard no espaço de trabalho **Scorecards**. Depois de executar o scorecard, você pode exibir as pontuações no painel **Scorecard**. É possível selecionar o objeto de dados e navegar para o objeto de dados a partir uma pontuação em um scorecard. A ferramenta Analyst abre o objeto de dados em outra guia.

As seguintes tarefas podem ser executadas ao trabalhar com pastas:

1. Criar um scorecard na Developer tool ou na ferramenta Analyst e adicionar colunas de um perfil.
2. Abrir o scorecard na ferramenta Analyst.
3. Depois de executar um perfil, adicionar colunas de perfil como métricas ao scorecard.
4. Opcionalmente, criar filtros de scorecard com base nos dados de origem.
5. Opcionalmente, configure o custo de dados inválidos para cada métrica.
6. Executar o scorecard para gerar pontuações para as colunas.
7. Exibir o scorecard para ver as pontuações para cada coluna em um registro.
8. Fazer uma busca detalhada nas colunas para uma pontuação.
9. Editar um scorecard.
10. Definir os limites para cada métrica em um scorecard.
11. Criar um grupo para adicionar ou mover métricas relacionadas no scorecard.
12. Editar ou excluir um grupo, conforme necessário.

13. Exibir o gráfico de tendências de cada pontuação para monitorar como ela muda com o passar do tempo.
14. Opcionalmente, exibir o gráfico de tendências de custo de cada métrica para monitorar o valor da qualidade de dados.
15. Exibir a linhagem de scorecard para cada métrica ou grupo de métrica.

Criando um scorecard no Informatica Analyst

Crie um scorecard e adicione colunas de um perfil ao scorecard. Antes de adicionar colunas ao scorecard, você deve executar um perfil.

1. No espaço de trabalho **Biblioteca**, selecione o projeto ou a pasta que contém o perfil.
2. Clique no perfil para abri-lo.

   Os resultados do perfil aparecem na exibição resumida, no espaço de trabalho **Descoberta**.
3. Clique em **Ações** > **Adicionar ao scorecard**.

   O assistente **Adicionar ao Scorecard** é exibido.
4. Na tela **Adicionar ao Scorecard**, você pode optar por criar um novo scorecard ou editar um scorecard existente para adicionar as colunas a um scorecard predefinido. A opção **Novo Scorecard** é selecionada por padrão. Clique em **Avançar**.
5. Na tela **Etapa 2 de 8**, insira um nome para o scorecard. Opcionalmente, você pode inserir uma descrição para o scorecard. Selecione o projeto e a pasta em que você deseja salvar o scorecard. Clique em **Avançar**.

   Por padrão, o assistente de scorecard seleciona as colunas e regras definidas no perfil. Não é possível adicionar colunas que não são incluídas no perfil.
6. Na tela **Etapa 3 de 8**, selecione as colunas e regras que você deseja adicionar ao scorecard como métricas. Como alternativa, clique na caixa de seleção no cabeçalho da coluna esquerda para selecionar todas as colunas. Como alternativa, selecione **Nome da Coluna** para classificar nomes de colunas. Clique em **Avançar**.
7. Na tela **Etapa 4 de 8**, você pode adicionar um filtro à métrica.

   Você pode aplicar o filtro criado para o perfil às métricas ou pode criar um novo filtro. Selecione uma métrica no painel **Filtros de Métrica** e clique no ícone **Gerenciar Filtros** para abrir a caixa de diálogo **Editar Filtro: nome da coluna**. Na caixa de diálogo **Editar Filtro: nome da coluna**, você pode optar por realizar uma das seguintes tarefas:

   - Você pode escolher um filtro que criou para o perfil. Clique em **Avançar**.
   - Selecione um filtro existente. Clique no ícone Editar para editar o filtro na caixa de diálogo **Editar Filtro**. Clique em **Avançar**.
   - Clique no ícone de adição (+) para criar filtros na caixa de diálogo **Novo Filtro**. **Clique em Avançar.**

   O filtro aparece no painel **Filtros de Métrica**. É possível aplicar o mesmo filtro a todas as métricas do scorecard.
8. Na tela **Etapa 4 de 8**, clique em **Avançar**.
9. Na tela **Etapa 5 de 8**, selecione cada métrica no painel **Métricas** e configure os valores válidos na lista de todos os valores no painel **Pontuação usando: Valores**. As seguintes tarefas podem ser realizadas na tela **Etapa 5 de 7**:

- Você pode selecionar vários valores no painel **Valores Disponíveis** e clicar no botão de seta à direita para movê-los até o painel **Valores Válidos**. O número total de valores válidos de uma métrica é exibido na parte superior do painel **Valores Disponíveis**.
- No painel **Limites de Métrica**, configure limites de métricas.

  É possível definir limites para as pontuações **Bom**,**Aceitável** e **Inaceitável**.
- Selecione cada métrica e configure o custo de dados inválidos. Para atribuir um valor constante ao custo da métrica, selecione **Custo Fixo**. Opcionalmente, clique em **Alterar Unidade de Custo** para alterar a unidade de custo ou escolha **Nenhuma**. Para anexar uma coluna numérica como um custo variável à métrica, selecione **Custo Variável** e clique em **Selecionar Coluna** para selecionar uma coluna numérica.

10. Na tela **Etapa 6 de 8**, você pode selecionar um grupo de métricas ao qual é possível adicionar as métricas ou criar um novo grupo de métricas. Para criar um novo grupo de métricas, clique no ícone do grupo. Clique em **Avançar**.
11. Na tela **Etapa 8 de 7**, especifique os pesos para as métricas no grupo e os limiares do grupo.
12. Na tela **Etapa 8 de 8**, selecione **Nativo** ou **Hadoop** como ambiente de tempo de execução para executar o scorecard.
13. Clique em **Salvar** para salvar o scorecard ou clique em **Salvar e Executar** para salvar e executar o scorecard.

    O scorecard é exibido no espaço de trabalho **Scorecard**.

Adicionando colunas a um scorecard existente

Depois de executar um perfil, você pode adicionar colunas a um scorecard existente, configurar os valores válidos para as colunas e adicionar o custo de dados inválidos para cada métrica. Se você adicionar uma coluna a um scorecard de um perfil com um filtro ou uma opção de amostragem diferente de **Todas as Linhas**, os resultados de perfil talvez não reflitam os resultados do scorecard.

1. Clique em um perfil para abri-lo.

   Os resultados do perfil aparecem na exibição resumida.
2. Selecione uma coluna. Clique em **Ações** > **Adicionar ao scorecard**.

   O assistente **Adicionar ao Scorecard** é exibido.

   Nota: Use as seguintes regras e diretrizes a seguir antes de adicionar colunas a um scorecard:
   - Você pode adicionar uma coluna a um scorecard se o nome da coluna e do scorecard corresponderem.
   - Você não pode adicionar uma coluna duas vezes para um scorecard, mesmo se alterar o nome da coluna.
3. Selecione **Scorecard Existente** para adicionar as colunas a um scorecard existente. Clique em **Avançar**.
4. Na tela **Etapa 2 de 7**, selecione o scorecard, as métricas e o grupo de métricas aos quais você deseja adicionar as colunas. Clique em **Avançar**.
5. Na tela **Etapa 3 de 7**, selecione as colunas e regras que você deseja adicionar ao scorecard como métricas. Como alternativa, clique na caixa de seleção no cabeçalho da coluna esquerda para selecionar todas as colunas. Clique em **Nome da Coluna** para classificar nomes de colunas. Clique em **Avançar**.
6. Na tela **Etapa 4 de 7**, você pode criar filtros para as métricas. Também pode aplicar às métricas o filtro criado para o perfil.

7. Na tela **Etapa 5 de 7**, você pode realizar as seguintes tarefas:
   - No painel **Métricas**, selecione cada métrica e configure valores de métricas em outros painéis.
   - No painel **Pontuação usando: Valores**, selecione vários valores no painel **Valores Disponíveis** e clique no botão de seta à direita para movê-los até o painel **Valores Válidos**.
     O número total de valores válidos de uma métrica é exibido na parte superior do painel **Valores Disponíveis**.
   - No painel **Limites de Métrica:** , é possível definir limites para as pontuações **Bom**,**Aceitável** e **Inaceitável.**.
   - Em **Custo de dados inválidos**, é possível:
     - Selecionar cada métrica e configurar o custo de dados inválidos para a métrica.
     - Selecionar a opção **Custo Fixo** para atribuir um valor constante ao custo da métrica. Você pode clicar em **Alterar Unidade de Custo** para alterar a unidade de custo.
     - Selecione a opção **Custo Variável** para anexar uma coluna numérica como um custo variável à métrica. Você pode clicar em **Selecionar Coluna** para selecionar uma coluna numérica.
8. Clique em **Avançar**.
9. Na tela **Etapa 7 de 6**, você pode realizar as seguintes tarefas:
   - Selecione o grupo de métricas ao qual você deseja adicionar as métricas.
   - No painel **Padrão - Métricas**, é possível clicar duas vezes no peso da métrica padrão de 0 para alterar o valor.
   - No painel **Limites de Métrica:** , é possível definir limites para as pontuações **Bom**,**Aceitável** e **Inaceitável**.
10. Clique em **Avançar**.
11. Na tela **Etapa 7 de 7**, selecione um ambiente de tempo de execução.
12. Clique em **Salvar** para salvar o scorecard ou clique em **Salvar e Executar** para salvar e executar o scorecard.

Executar um Scorecard

Execute um scorecard para gerar pontuações para colunas.

1. No painel **Ativos**, escolha o scorecard que deseja executar.
2. Clique no scorecard para abri-lo.

   O scorecard será exibido no espaço de trabalho **Scorecards**.
3. Clique em **Ações** > **Executar Scorecard.**
4. Selecione uma pontuação no painel **Métricas** e selecione as colunas no painel **Colunas** para fazer uma busca detalhada neles.
5. Na opção **Busca Detalhada**, escolha uma busca detalhada nos dados ativos ou dados de fase.

   Para obter o melhor desempenho, faça uma busca detalhada nos dados ativos.
6. Clique em **Executar**.

Exibir um Scorecard

Execute um scorecard para ver as pontuações para cada métrica. Um scorecard exibe a pontuação como uma porcentagem e barra. Exiba dados válidos ou inválidos. É possível exibir informações de scorecard, como o peso da métrica, pontuação de grupo de métrica, tendência de pontuação e o nome do objeto de dados.

1. Execute um scorecard para exibir as pontuações.
2. Selecione uma métrica que contenha a pontuação que você deseja exibir.
3. Clique em **Ações** > **Busca Detalhada** para exibir as linhas de dados válidos ou as linhas de dados inválidos da coluna.

   A ferramenta Analyst exibe as linhas de dados inválidos por padrão na seção **Busca Detalhada**.

Editar um Scorecard

Edite os valores válidos para métricas em um scorecard. Execute um scorecard para poder editá-los.

1. No espaço de trabalho **Biblioteca**, clique no scorecard que você deseja editar no painel **Ativos**.

   O scorecard é exibido no espaço de trabalho **Scorecards**.
2. Se o sistema de controle de versão estiver ativado, clique em **Ações** > **Fazer Check-Out**.
3. Clique em **Ações** > **Editar** > **Geral**.

   A caixa de diálogo **Editar Scorecard** é exibida.
4. Na guia **Geral**, você pode editar o nome e a descrição do scorecard conforme necessário.
5. Clique na guia **Métricas**.
6. Selecione uma pontuação no painel **Métricas** e configure os valores válidos na lista de todos os valores no painel **Pontuar usando: Valores**.
7. No painel **Limites de Métrica**, você pode fazer alterações nos limites de pontuação conforme necessário.
8. Revise o custo de dados inválidos para cada métrica e faça alterações conforme necessário.
9. Clique na guia **Filtros de Scorecard**.
10. Você pode adicionar, editar ou excluir filtros.
11. Clique na guia **Grupos de Métrica**.
12. Você pode criar, editar ou remover os grupos de métricas.

    Você também pode editar os pesos e os limites da métrica na guia **Grupos de Métricas**.
13. Clique na guia **Notificações**.
14. Você pode fazer alterações nas configurações de notificação de scorecard conforme necessário.

    Você pode definir as configurações globais e personalizadas para métricas e grupos de métrica.
15. Clique na guia **Ambiente de Tempo de Execução**.

    Você pode selecionar **Nativo** ou **Hadoop** como o ambiente de tempo de execução.
16. Clique em **Salvar** para salvar as alterações no scorecard ou clique em **Salvar e Executar**para salvar as alterações e executar o scorecard.
17. Clique em **Fazer Check-In**.

Métricas

Uma métrica é uma coluna de uma fonte de dados ou saída de uma regra que é parte de um scorecard. Ao criar um scorecard, você pode atribuir um peso para cada métrica. Crie um grupo de métrica para categorizar as métricas relacionadas em um scorecard para um conjunto.

Pesos de Métrica

Ao criar um scorecard, você pode atribuir um peso para cada métrica. O valor padrão de um peso é 0.

Quando você executa um scorecard, a ferramenta Analyst calcula a média ponderada para cada grupo de métrica com base na pontuação e peso da métrica que você atribui a cada métrica.

Por exemplo, você atribui um peso de W1 para métrica M1 e atribui um peso de W2 para métrica M2. A ferramenta Analyst usa a seguinte fórmula para calcular a média ponderada:

```
(M1 X W1 + M2 X W2) / (W1 + W2)
```

Valor da Qualidade de Dados

Medidas de qualidade de dados na fonte de dados são informações críticas no gerenciamento de ativos de dados na organização. O custo de dados inválidos em métricas representado em um scorecard ajuda as organizações a gerar valor no monitoramento da qualidade dos dados de origem. Como analista de dados, você pode querer associar um valor, como uma unidade monetária ou qualquer unidade personalizada, a métricas e grupos de métricas. Em seguida, é possível executar o scorecard para exibir o custo total de dados inválidos nos dados de origem.

Você pode definir a unidade de custo de uma métrica com base nas necessidades comerciais. Você também pode configurar uma variável ou um custo fixo para cada métrica ao criar um scorecard ou editá-lo.

Custo Fixo

O custo corrigido custo é um valor constante que você pode atribuir a uma métrica em um scorecard. Você pode escolher uma unidade de custo predefinida ou criar uma unidade de custo personalizada que atende às necessidades comerciais.

Custo Variável

O custo variável é um valor que você atribui a uma métrica com base nos valores em uma coluna numérica de uma fonte de dados. O Serviço de Integração de Dados calcula o custo variável para a métrica com base na coluna ou na coluna virtual que você atribui ao custo.

Exemplo

Como analista de crédito imobiliário, você precisa fornecer aos seus clientes livros de pagamento para que eles possam enviar os pagamentos do financiamento. Você pode usar um scorecard para medir a precisão dos endereços dos seus clientes para garantir a entrega dos livros de pagamento. Você pode querer definir o custo variável para a coluna Valor do Pagamento Mensal da métrica Precisão do Endereço. Execute o scorecard para calcular o custo total que a organização hipotecária perderá se os clientes não pagarem o valor mensal dentro do prazo.

Definir Limites

Os limites para cada pontuação em um scorecard podem ser definidos. Um limite especifica a faixa na porcentagem de dados inválidos que é aceitável para colunas em um registro. Os limites para uma faixa de dados podem ser definidos como bom, aceitável ou inaceitável. Você pode definir limites para cada coluna quando adicionar colunas a um scorecard, ou quando editar um scorecard.

Conclua uma das seguintes tarefas de pré-requisito antes de definir os limites de colunas em um scorecard:

- Abra um perfil e adicione colunas do perfil ao scorecard na caixa de diálogo **Adicionar ao Scorecard**.
- Como alternativa, clique em um scorecard no espaço de trabalho **Biblioteca** e selecione **Ações** > **Editar** para editar o scorecard na caixa de diálogo **Editar Scorecard**.

1. Na caixa de diálogo **Adicionar ao Scorecard** ou na caixa de diálogo **Editar Scorecard**, selecione cada métrica no painel **Métricas**.
2. No painel **Limites de Métrica**, insira os limites que representam o limite mais alto da faixa inaceitável e o limite mais baixo de uma boa faixa.
3. Clique em **Avançar** ou **Salvar**.

Grupos de Métrica

Crie um grupo de métrica para categorizar as pontuações relacionadas em um scorecard para um conjunto. Por padrão, a ferramenta Analyst classifica todas as pontuações em um grupo de métrica padrão.

Após criar um grupo de métrica, você pode mover pontuações do grupo de métrica padrão para outro grupo de métrica. É possível editar um grupo de métrica para alterar seu nome e descrição, incluindo o grupo de métrica padrão. Você pode excluir grupos de métrica que não são mais usados. Mas não pode excluir o grupo de métrica padrão.

Criar um Grupo de Métrica

Crie um grupo de métrica para adicionar pontuações relacionadas no scorecard para o grupo.

1. No espaço de trabalho **Biblioteca**, clique no scorecard que você deseja editar no painel **Ativos**.

   O scorecard é exibido no espaço de trabalho **Scorecards**.
2. Clique em **Ações** > **Editar**.

   A janela **Editar Scorecard** é exibida.
3. Clique na guia **Grupos de Métrica**.

   O grupo padrão é exibido no painel **Grupos de Métrica** e as pontuações no grupo padrão aparecem no painel **Métricas**.
4. Clique no ícone **Novo Grupo** para criar um grupo de métrica.

   A caixa de diálogo **Grupos de Métrica** é exibida.
5. Insira um nome e descrição opcional.
6. Clique em **OK**.
7. Clique em **Salvar** para salvar as alterações para o scorecard.

Mover Pontuações para um Grupo de Métrica

Depois de criar um grupo de métrica, é possível mover pontuações relacionadas ao grupo de métrica.

1. No espaço de trabalho **Biblioteca**, clique no scorecard que você deseja editar no painel **Ativos**.

   O scorecard é exibido no espaço de trabalho **Scorecards**.
2. Clique em **Ações** > **Editar**.

   A janela **Editar Scorecard** é exibida.
3. Clique na guia **Grupos de Métrica**.

   O grupo padrão é exibido no painel **Grupos de Métrica** e as pontuações no grupo padrão aparecem no painel **Métricas**.
4. Selecione uma métrica do painel **Métricas** e clique no ícone **Mover Métricas**.

   A caixa de diálogo **Mover Métricas** é exibida.

   Nota: Para selecionar várias pontuações, mantenha pressionada a tecla Shift.
5. Selecione o grupo de métrica para mover as pontuações.
6. Clique em **OK**.

Editar um Grupo de Métrica

Edite um grupo de métrica para alterar o nome e descrição. Você pode alterar o nome do grupo de métrica padrão.

1. No espaço de trabalho **Biblioteca**, clique no scorecard que você deseja editar no painel **Ativos**.

   O scorecard é exibido no espaço de trabalho **Scorecards**.
2. Clique em **Ações** > **Editar**.

   A janela **Editar Scorecard** é exibida.
3. Clique na guia **Grupos de Métrica**.

   O grupo de métrica padrão é exibido no painel **Grupos de Métrica** e as métricas no grupo de métrica padrão aparecem no painel **Métricas**.
4. No painel **Grupos de Métrica**, clique no ícone **Editar Grupo**.

   A caixa de diálogo **Editar** é exibida.
5. Insira um nome e uma descrição opcional.
6. Clique em **OK**.

Excluir um Grupo de Métrica

Você pode excluir um grupo de métrica que não é mais válido. Ao excluir um grupo de métrica, você pode mover as pontuações no grupo de métrica para o grupo de métrica padrão. Mas não pode excluir o grupo de métrica padrão.

1. No espaço de trabalho **Biblioteca**, clique no scorecard que você deseja editar no painel **Ativos**.

   O scorecard é exibido no espaço de trabalho **Scorecards**.
2. Clique em **Ações** > **Editar**.

   A janela **Editar Scorecard** é exibida.
3. Clique na guia **Grupos de Métrica**.

O grupo de métrica padrão é exibido no painel **Grupos de Métrica** e as métricas no grupo de métrica padrão aparecem no painel **Métricas**.

4. Selecione um grupo de métrica no painel **Grupos de Métrica**, e clique no ícone **Excluir Grupo**.

   A caixa de diálogo **Excluir Grupos** é exibida.

5. Escolha a opção para excluir as métricas no grupo de métrica ou a opção para mover as métricas para o grupo de métrica padrão antes de excluir o grupo de métrica.
6. Clique em **OK**.

Buscas Detalhadas em Colunas

Faça uma busca detalhada nas colunas para uma pontuação para selecionar colunas que são exibidas quando você exibe as linhas de dados válidos ou linhas de dados inválidos. As colunas selecionadas para fazer uma busca detalhada aparecem no painel **Busca Detalhada**.

1. Execute um scorecard para exibir as pontuações.
2. Selecione uma coluna que contenha a pontuação que você deseja exibir.
3. Clique em **Ações** > **Busca Detalhada** para exibir as linhas de dados válidos ou inválidos para a coluna.
4. Clique em **Ações** > **Colunas de Busca Detalhada.**

   As colunas são exibidas no painel **Busca Detalhada** para a pontuação selecionada. A ferramenta Analyst exibe as linhas de dados válidos para as colunas por padrão. Opcionalmente, clique em **Inválido** para exibir as linhas de dados que não são válidos.

Gráficos de Tendência

Use gráficos de tendências para monitorar como as pontuações de métricas e o custo de dados inválidos em métricas mudam no decorrer de um período de tempo.

Os gráficos de tendências contêm gráficos de pontuação e gráficos de custo que fazem a plotagem dos valores de pontuação e de custo no eixo vertical, comparando-os com todas as execuções de scorecard no eixo horizontal. Por padrão, o gráfico de tendências mostra dados das 10 últimas execuções de scorecard. Você pode exibir o número total de linhas e as linhas inválidas para a métrica no gráfico de tendências. O gráfico de tendências também mostra se as tendências de pontuação e custo permaneceram constantes ou subiram ou desceram com base na última execução de scorecard.

A ferramenta Analyst usa dados históricos de execução de scorecard para cada data e os últimos valores de pontuação válidos para calcular a pontuação. A ferramenta Analyst usa as configurações de limite mais recentes no gráfico para representar a cor dos pontos de pontuação. É possível ver os limites Bom, Aceitável e Inaceitável para a pontuação. Os limites mudam sempre que você executa o scorecard depois de editar os valores das pontuações no scorecard. Quando você exporta um scorecard, a ferramenta Analyst inclui as informações do gráfico de tendências, entre elas as informações de pontuação e de custo, no arquivo exportado.

Gráfico de Tendência de Pontuação

Um gráfico de tendências de pontuação é uma representação gráfica de como as pontuações de métricas mudam no decorrer de várias execuções de perfil. O gráfico de tendências de pontuação faz a plotagem dos valores de pontuação de métricas no eixo vertical, comparando-os com todas as execuções de scorecard no eixo horizontal.

A seguinte imagem mostra uma amostra de gráfico de tendências de pontuação:

Exemplo

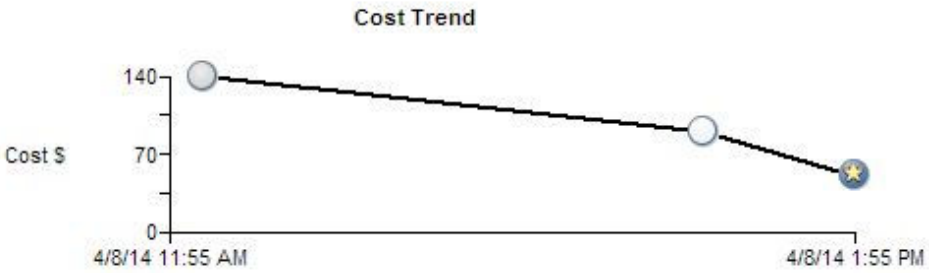
Como analista de dados, você pode monitorar a qualidade de dados para analisar se alterações em mapeamentos e em outros processos provocam aumentos na pontuação de qualidade de dados. Depois de medir a alteração na qualidade de dados, você pode informar essa alteração para que ela seja analisada e utilizada pela organização. Por exemplo, no final de várias execuções de scorecard, a porcentagem de valores válidos em uma coluna Número de Previdência Social pode ter aumentado de 84 para 90. É possível relatar essa alteração na qualidade de dados como um gráfico visual para uma rápida análise.

Gráfico de Tendência de Custo

Um gráfico de tendências de custo é uma representação gráfica de como o custo de dados inválidos em métricas muda ao longo de várias execuções de perfil. O gráfico de tendências de custo pode medir o impacto da qualidade de dados em uma organização. O gráfico de tendências de custo faz a plotagem dos valores de custo no eixo vertical, comparando-os com todas as execuções de scorecard no eixo horizontal. Você também pode exibir o custo total de dados inválidos e os valores válidos para a métrica em uma grade abaixo do gráfico de tendências de custo.

Um gráfico de tendências de custo ajuda a controlar o impacto de dados inválidos em registros de alto valor. Ocasionalmente, ao usar um custo fixo para calcular dados inválidos, o impacto de dados inválidos em registros de alto valor pode passar despercebido. Esse problema ocorre porque os gráficos de tendências podem mostrar uma melhoria no pontuação e uma diminuição no custo total no decorrer de várias execuções de scorecard. No entanto, os menores problemas de qualidade de dados representados no scorecard podem existir em registros de alto valor.

A seguinte imagem mostra uma amostra de gráfico de tendências de custos:



Exemplo

Em uma instituição financeira, existem vários clientes com saldos elevados e grandes depósitos e investimentos, como US$ 10 milhões, no banco. Também existe um grande número de clientes com saldos baixos. O gráfico de tendências de pontuação pode mostrar uma melhoria nas pontuações ao longo de um período de tempo. No entanto, um sexo ou endereço incorreto em algumas contas de cliente com saldo elevado pode afetar os relacionamentos com os clientes mais importantes da organização. É possível definir a coluna Saldo da Conta como a coluna de custo variável para calcular dados inválidos. Se o custo de dados inválidos devido à coluna for alto, você poderá considerar o valor total em risco e tomar ações corretivas imediatas.

Exibir Gráficos de Tendência

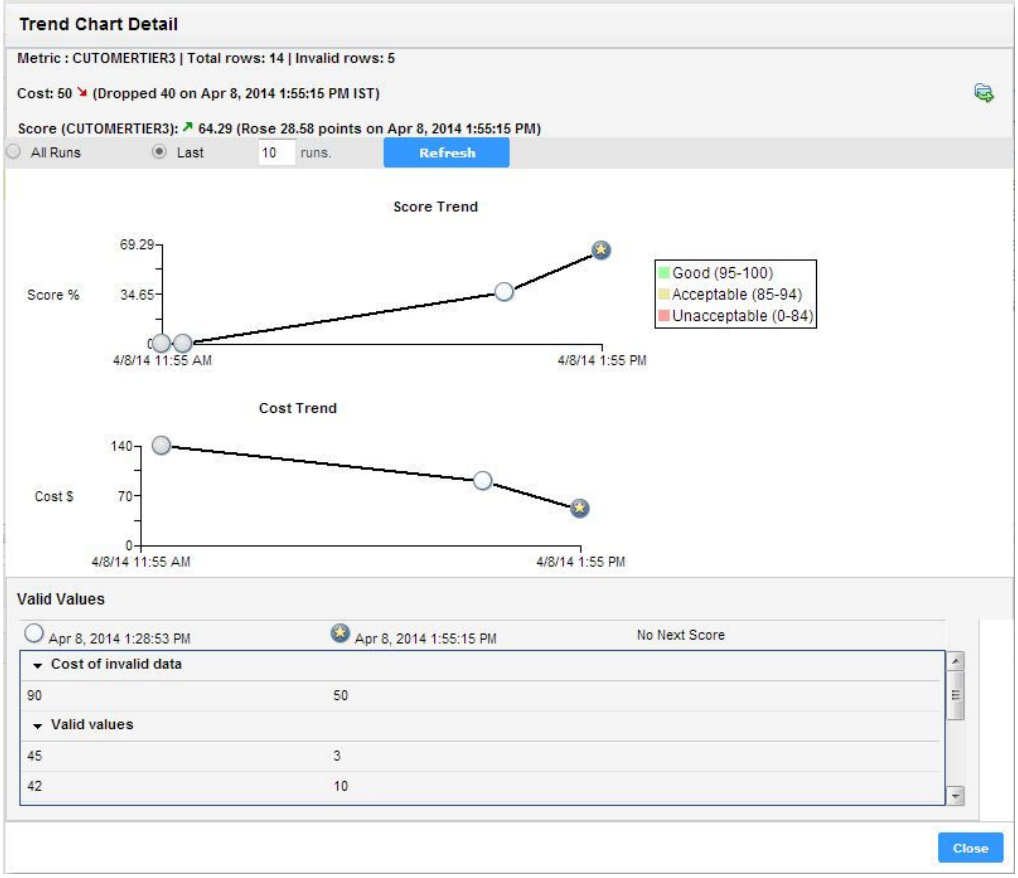
Você pode exibir gráficos de tendências para cada métrica para monitorar como a pontuação ou o custo de dados inválidos muda com o passar do tempo.

1. No espaço de trabalho **Biblioteca**, selecione o projeto ou a pasta que contém o scorecard.
2. Clique no scorecard para abri-lo.

   O scorecard será exibido no espaço de trabalho **Scorecards**.
3. Na exibição **Scorecard**, selecione uma métrica.

4. Clique em **Ações** > **Mostrar Gráfico de Tendência**.

   A caixa de diálogo **Detalhe do Gráfico de Tendência** é exibida.
   A seguinte imagem mostra a caixa de diálogo **Detalhe do Gráfico de Tendência**:



   Você pode exibir valores de pontuação e custo que mudaram com o passar do tempo. Na parte superior da caixa de diálogo, você pode exibir o número total de linhas e o número de linhas inválidas. A ferramenta Analyst usa dados históricos de execução de scorecard para cada data e os últimos valores de pontuação válidos para calcular a pontuação. Abaixo dos gráficos de tendência de pontuação e custo, você pode exibir os valores válidos para a métrica e o custo dos dados inválidos.

Exportando Gráficos de Tendência

Você pode exportar os gráficos de tendências de pontuação e custo para um arquivo ".xlsx" de forma a visualizar os dados em um arquivo.

1. Abra um scorecard.
2. Selecione uma métrica e clique em **Ações** > **Mostrar Gráfico de Tendência**.

   A caixa de diálogo **Detalhes do Gráfico de Tendência** é exibida.
3. Clique no ícone **Exportar Dados**.

   A caixa de diálogo **Exportar dados para um arquivo** é exibida.
4. Digite um nome de arquivo. Como alternativa, use o nome de arquivo padrão.

   O formato de arquivo padrão é Microsoft Excel.
5. Selecione a página de código do arquivo.

6. Clique em **OK**.

Arquivos de Exportação de Scorecard no Informatica Analyst

Você pode exportar os resultados do scorecard para um arquivo do Microsoft Excel. A ferramenta Analyst salva o arquivo no formato "xlsx".

O arquivo exportado contém o resumo do scorecard, os gráficos de tendências, as linhas inválidas e as propriedades do scorecard.

Resultados da Exportação do Scorecard no Microsoft Excel

Quando você exporta os resultados do scorecard, a ferramenta Analyst salva as informações em várias planilhas em um arquivo do Microsoft Excel. A ferramenta Analyst salva o arquivo no formato "xlsx".

A tabela a seguir descreve as informações exibidas em cada planilha no arquivo de exportação:

Guia	Descrição
Resumo do scorecard	Informações de resumo dos resultados de scorecard exportado. As informações incluem o nome do scorecard, o número total de linhas de cada coluna, o número de linhas que não são válidas, a pontuação e o peso da métrica.
Gráfico de tendências	Gráficos de tendências das pontuações.
Linhas inválidas	Os detalhes de linhas que não são válidas de cada coluna. A ferramenta Analyst exporta um máximo de 100 linhas para a planilha.
Propriedades	Propriedades do scorecard, como nome, tipo, descrição e localização.

Exportando Resultados do Scorecard do Informatica Analyst

Você pode exportar os resultados do scorecard para um arquivo ".xlsx" para exibir os dados em um arquivo.

1. Abra um scorecard.
2. Clique em **Ações** > **Exportar Dados.**

   A caixa de diálogo **Exportar Dados para um Arquivo** será exibida.
3. Digite um nome de arquivo. Como alternativa, use o nome de arquivo padrão.

   O formato de arquivo padrão é Microsoft Excel.
4. Selecione a página de código do arquivo.
5. Clique em **OK**.

Notificações de Scorecard

Defina configurações de notificação de scorecard para que a ferramenta Analyst envie e-mails quando pontuações de métricas, pontuações de grupos de métricas ou custos de métricas específicos ultrapassarem limites. Pontuações de métricas ou de grupos de métricas podem ultrapassar limites ou permanecer em faixas de pontuação específicas, como Inaceitável, Aceitável e Bom. Valores de custo de métricas podem ultrapassar os limites de custo superior e inferior que você define.

Você pode configurar notificações de e-mail para pontuações de métricas, grupos de métricas e custos de métricas específicos. Se você usar as configurações globais para pontuações, a ferramenta Analyst enviará e-mails de notificação quando pontuações de métrica específicas ultrapassarem o limite das faixas de pontuação Bom para Aceitável e Aceitável para Ruim. Você também pode receber e-mails de notificação para cada execução de scorecard se a pontuação permanecer na faixa de pontuação Inaceitável ao longo de execuções de scorecard consecutivas. Se você usar as configurações globais para custos de métricas, a ferramenta Analyst enviará e-mails de notificação quando o custo de dados inválidos em métricas selecionadas ultrapassar os limites superior e inferior.

Também pode personalizar as configurações de notificação de forma que os usuários de scorecard recebam notificações de e-mail quando as pontuações passarem das faixas de pontuação Inaceitável para Aceitável e Aceitável para Bom. Você pode optar por enviar notificações de e-mail se uma pontuação de métrica ou um custo de métrica permanecer em faixas específicas para cada execução de scorecard. É possível exibir o custo atual de dados inválidos para cada métrica nas configurações de notificação com base nas quais você pode definir os limites de custo.

Antes de você configurar scorecards para enviar notificações por e-mail, um administrador deve configurar o Serviço de E-mail na ferramenta Administrator.

Modelo de Mensagem de Notificação de E-mail

Também pode configurar o texto da mensagem e estrutura de mensagens de e-mail que a ferramenta Analyst envia para os destinatários como parte das notificações de scorecard. O modelo de e-mail tem uma seção de texto introdutório opcional, seção de corpo da mensagem somente leitura e seção de texto de encerramento opcional.

A tabela a seguir descreve as marcas no modelo de e-mail:

Marca	Descrição
ScorecardName	Nome do scorecard.
ObjectURL	Um link para o scorecard. Forneça o nome de usuário e senha.
MetricGroupName	Nome do grupo de métrica ao qual a métrica pertence.
CurrentWeightedAverage	Valor da média ponderada para o grupo de métrica na execução de scorecard atual.
CurrentRange	A faixa de pontuação, como Inaceitável, Aceitável e Bom, para o grupo de métrica na execução de scorecard atual.
PreviousWeightedAverage	Valor da média ponderada para o grupo de métrica na execução de scorecard anterior.

Marca	Descrição
PreviousRange	A faixa de pontuação, como Inaceitável, Aceitável e Bom, para o grupo de métrica na execução de scorecard anterior.
MetricName	Nome da métrica.
MetricGroupName	Nome do grupo de métricas.
CurrentScore	Pontuação com base na última execução de scorecard.
CurrentRange	Faixa de pontuação na qual a pontuação atual permanece com base na última execução de scorecard.
PreviousScore	Pontuação com base na execução de scorecard anterior.
PreviousRange	Faixa de pontuação com base na execução de scorecard anterior.
CurrentCost	Custo de dados inválidos na métrica com base na última execução de scorecard.
PreviousCost	Custo de dados inválidos na métrica com base na execução de scorecard anterior.
ColumnName	Nome da coluna de origem para a qual a métrica está sendo atribuída.
ColumnType	Tipo de coluna de origem.
RuleName	Nome da regra.
RuleType	Tipo de regra.
DataObjectName	Nome do objeto de dados de origem.

Configurar as Notificações de Scorecard

Você pode configurar as notificações de scorecard nos níveis de métrica e grupo de métrica. As configurações de notificação globais se aplicam a essas métricas e grupos de métrica que não têm configurações de notificação individuais,

1. Execute um scorecard na ferramenta Analyst.
2. Clique em **Ações** > **Editar**.
3. Clique na guia **Notificações**.
4. Selecione **Habilitar Notificações** para iniciar a configuração de notificações de scorecard.
5. Selecione uma métrica ou um grupo de métricas.
6. Clique na caixa de seleção **Notificações** para ativar as configurações globais para a métrica ou o grupo de métricas.
7. Selecione **Usar configurações personalizadas** para alterar as configurações da métrica ou do grupo de métricas.

   Você pode optar por enviar um e-mail de notificação quando a pontuação estiver nas faixas **Inaceitável**, **Aceitável** e **Bom** ultrapassar os limites. Você também pode enviar um e-mail de notificação quando o custo da métrica ultrapassar o limite superior ou inferior.

8. Para editar as configurações globais para notificações de scorecard, clique no ícone **Editar Configurações Globais**.

   A caixa de diálogo **Editar Configurações Globais** é exibida onde você pode editar as configurações, incluindo o modelo de e-mail.

Definir as Configurações Globais para Notificações de Scorecard

Se você escolher as configurações globais de notificação de scorecard, a ferramenta Analyst enviará e-mails aos usuários de destino quando a pontuação estiver na faixa **Inaceitável**. Você também pode definir as configurações de notificação para enviar e-mails quando as pontuações de métricas ou os custos de métricas ultrapassarem limites. É possível configurar o modelo de e-mail, incluindo os endereços de e-mail e o texto de mensagem para um scorecard.

1. Execute um scorecard na ferramenta Analyst.
2. Clique em **Ações** > **Editar** > **Notificações** para abrir a caixa de diálogo **Editar Scorecard**.
3. Selecione **Habilitar Notificações** para iniciar a configuração de notificações de scorecard.
4. Clique no ícone **Editar Configurações Globais**.

   A caixa de diálogo **Editar Configurações Globais** é exibida onde você pode editar as configurações, incluindo o modelo de e-mail.
5. Escolha quando você deseja enviar notificações de e-mail para pontuações de métricas usando as caixas de seleção **Pontuação em** e **Pontuação move**.
6. Escolha quando você deseja enviar notificações de e-mail para custos de métrica usando as caixas de seleção **Custo atinge**.
7. No campo **E-mail para**, insira o ID de e-mail do destinatário. Use ponto-e-vírgula para separar vários IDs de e-mail.

   O ID de e-mail do remetente padrão é o **Endereço de E-mail do Remetente** configurado nas propriedades de SMTP do domínio.
8. Insira o texto para o assunto do e-mail.
9. No campo **Corpo**, adicione o texto de introdução e fechamento e o texto da mensagem de e-mail.
10. Para aplicar as configurações globais, selecione **Aplicar configurações para todas as métricas e grupos de métrica**.
11. Clique em **OK**.

Linhagem de Scorecard

A linhagem de scorecard mostra a origem dos dados, descreve o caminho e mostra como os dados fluem para uma métrica ou grupo de métricas. Você pode usar a linhagem de scorecard para analisar a causa raiz de uma variação inaceitável na pontuação em métricas ou grupo de métricas. Exiba a linhagem de scorecard na ferramenta Analyst.

Conclua as seguintes tarefas para exibir a linhagem de scorecard:

1. No Informatica Administrator, associe um Serviço do Metadata Manager ao Analyst Service.
2. Selecione um projeto e exporte os objetos de scorecard dele para um arquivo XML usando a opção Exportar Arquivo de Recursos para o Metadata Manager na ferramenta Developer ou o comando infacmd oie exportResources.

3. No Metadata Manager, use o arquivo XML exportado para criar um recurso e carregá-lo.

   Nota: O nome do arquivo de recursos que você criar e carregar no Metadata Manager deverá usar a seguinte convenção de nomenclatura: <MRS name>_<project name>. Para obter mais informações sobre como criar e carregar um arquivo de recursos, consulte *Guia do Usuário do Metadata Manager do Informatica PowerCenter*.

4. Na ferramenta Analyst, abra o scorecard e selecione uma métrica ou grupo de métricas.
5. Exiba a linhagem de scorecard.

Exibir a Linhagem de Scorecard o Informatica Analyst

Você pode exibir um diagrama de linhagem de scorecard para uma métrica ou grupo de métrica. Antes de exibir o diagrama de linhagem de scorecard na ferramenta Analyst, carregue a linhagem do scorecard e os metadados no Metadata Manager.

1. No espaço de trabalho **Biblioteca**, clique no scorecard que você deseja exibir no painel **Ativos**.

   O scorecard é exibido no espaço de trabalho **Scorecards**.

2. Na exibição **Scorecard**, selecione uma métrica ou grupo de métrica.
3. Clique duas vezes e selecione **Exibir a Linhagem**.

   O diagrama de linhagem de scorecard é exibido em uma nova janela.

   Importante: Se você não criar e carregar um recurso no Metadata Manager com um arquivo XML exportado de objetos do scorecard, uma mensagem de erro informando que o recurso não está disponível no catálogo pode ser exibida. Para obter mais informações sobre exportação de um arquivo XML para a linhagem de scorecard, consulte"Exportar um Arquivo de Recursos para a Linhagem de Scorecard" na página 95.

Parte III: Criação de perfil com o Informatica Developer

Esta parte contém os seguintes capítulos:

- Perfis de Objeto de Dados, 75
- Perfis de coluna em fontes de dados semiestruturadas, 79
- Regras no Informatica Developer, 86
- Coluna de Resultados de Perfil no Informatica Developer, 88
- Scorecards no Informatica Developer, 94
- Criação de Perfil de Mapplet e Mapeamento, 97

CAPÍTULO 8

Perfis de Objeto de Dados

Este capítulo inclui os seguintes tópicos:

- Perfis de Coluna no Informatica Developer, 75
- Criando um Único Perfil de Objeto de DadosCriando um Perfil, 76
- Criando Vários Perfis de Objeto de Dados, 77
- Sincronizando um Objeto de Dados de Arquivo Simples, 78
- Sincronizar um Objeto de Dados Relacionais, 78

Perfis de Coluna no Informatica Developer

Use um perfil de coluna para analisar as características de colunas em uma fonte de dados, como porcentagens de valores e padrões de valores. Você pode adicionar filtros para determinar as linhas que o perfil lê no tempo de execução. O perfil não processa linhas que não atendem aos critérios do filtro.

Você pode descobrir os seguintes tipos de informações sobre as colunas nas quais executa um perfil:

- O número de vezes que um valor é exibido em uma coluna.
- Frequência de ocorrência de cada valor em uma coluna, expressa como porcentagem ou número de linhas.
- Os padrões de caracteres dos valores em uma coluna.
- As estatísticas, como os tamanhos mínimo e máximo dos valores em uma coluna, e o primeiro e último valor.
- Tipos de dados inferidos, frequência, critérios de conformidade para descoberta do domínio de dados e status da inferência de tipo de dados.

Você pode definir um perfil da coluna para um objeto de dados em um mapeamento ou mapplet, ou em um objeto no repositório do Modelo. O objeto no repositório pode estar em um único perfil de objeto de dados, em vários perfis de objetos de dados ou no perfil de descoberta empresarial.

É possível escolher opções de amostragem, opções de busca detalhada e um ambiente de tempo de execução para um perfil de coluna. Regras e filtros podem ser adicionados a um perfil de coluna.

Opções de Filtragem

Você pode adicionar filtros para determinar as linhas que um perfil da coluna usa ao executar operações de criação de perfil. O perfil não processa linhas que não atendam aos critérios de filtro.

1. Crie ou abra um perfil da coluna.

2. Selecione a exibição **Filtro**.
3. Clique em **Adicionar**.
4. Selecione um tipo de filtro e clique em **Avançar.**
5. Insira um nome para o filtro. Opcionalmente, insira uma descrição de texto do filtro.
6. Selecione **Definir como Ativo** para aplicar o filtro ao perfil. Clique em **Avançar**.
7. Defina os critérios de filtro.
8. Clique em **Concluir.**

Opções de Amostragem

Configure as opções de amostragem para determinar o número de linhas lidas pelo perfil durante uma operação de criação de perfil.

A seguinte tabela descreve as opções de amostragem:

Propriedade	Descrição
Todas as Linhas	Selecione todas as linhas no objeto de dados.
Primeiro	O número de linhas que você deseja executar em relação ao perfil. A Developer tool escolhe entre as primeiras linhas na origem.
Amostragem Aleatória de	O número de linhas de uma amostra aleatória na qual executar o perfil.
Amostragem Aleatória (Automática)	O tamanho de amostra aleatória com base no número de linhas no objeto de dados.
Exclui tipos de dados e domínios de dados aprovados da inferência de tipo de dados e de domínio de dados nas execuções de perfil subsequentes	Exclui o tipo de dados ou o domínio de dados aprovado da inferência de tipo de dados e domínio dados a partir da próxima execução de perfil.

Propriedade	Descrição
Todas as Linhas	Lê todas as linhas da origem. Por padrão, essa opção fica ativada.
Primeiro	Lê da primeira linha até a linha especificada.

Criando um Único Perfil de Objeto de DadosCriando um Perfil

Você pode criar um único perfil de objeto de dados para uma ou mais colunas em um objeto de dados e armazenar o objeto de perfil no repositório do Modelo.Você pode criar um perfil para uma ou mais colunas em um objeto de dados e armazenar o objeto de perfil no repositório do Modelo.

1. Na exibição **Object Explorer**, selecione o objeto de dados cujo perfil você deseja criar.

2. Clique em **Arquivo** > **Novo** > **Perfil** para abrir o assistente de perfil.
3. Selecione **Perfil** e clique em **Avançar**.
4. Insira um nome para o perfil e verifique a localização do projeto. Se necessário, navegue para uma nova localização.
5. Opcionalmente, insira uma descrição de texto do perfil.
6. Verifique se o nome do objeto de dados selecionado aparece na seção **Objetos de Dados**.
7. Clique em **Avançar**.
8. Configure as operações de perfil que você deseja executar. Você pode configurar as seguintes opções:
   - Criação de perfil de coluna
   - Descoberta de chave primária
   - Descoberta de dependência funcional
   - Descoberta do domínio de dados

   Nota: Para habilitar um perfil de operação, selecione **Ativado como parte da ação "Executar Perfil"** para esta operação. A criação de perfil da coluna é habilitada por padrão.
9. Configure as opções de perfil de coluna, como as opções de seleção de coluna, filtro, amostragem e busca detalhada.
10. Consulte as opções de perfil.

    A seleção de coluna pode ser editada para todos os tipos de perfil. Consulte o filtro e as opções de amostragem para a coluna de perfis. Você pode examinar as opções de inferência de chave primária, dependência funcional e descoberta do domínio de dados. Você também pode analisar a seleção de domínio de dados para a descoberta do domínio de dados.
11. Consulte as opções de busca detalhada e edite-as, se necessário. Por padrão, a opção **Ativar Busca Detalhada de Linha** está selecionada. Você pode editar as opções de busca detalhada dos perfis de coluna. As opções também determinam se as operações de busca detalhada leem da fonte de dados ou dos dados de preparação, e se o perfil armazena os dados de resultado das execuções de perfil anteriores.
12. Na seção **Executar Configurações**, configure as definições de validação e de tempo de execução. Você pode escolher a conexão **Nativo**, **Hive** ou **Hadoop** como o ambiente de tempo de execução.
13. Clique em **Concluir**.

Criando Vários Perfis de Objeto de Dados

Operações de criação de perfil com vários objetos de dados usam opções padrão para gerar perfis de coluna para um ou mais objetos de dados.

1. Na exibição **Object Explorer**, selecione os objetos de dados cujo perfil você deseja criar.
2. Clique em **Arquivo** > **Novo** > **Perfil** para abrir o assistente de **Novo Perfil**.
3. Selecione a opção **Vários Perfis** e clique em **Avançar**.
4. Selecione a localização onde você deseja criar os perfis. Você pode criar cada perfil no mesmo local do objeto cujo perfil foi criado ou pode especificar uma localização comuns para os perfis.
5. Verifique se os nomes dos objetos de dados selecionados aparecem na seção **Objetos de Dados**.

   Opcionalmente, clique em **Adicionar** para adicionar outro objeto de dados.

6. Opcionalmente, especifique o número de linhas para criar o perfil e escolha se deseja executar o perfil quando o assistente for concluído.
7. Clique em **Concluir**.
8. Opcionalmente, insira strings de prefixo e sufixo a serem adicionadas aos nomes dos perfis.
9. Clique em **OK**.

Sincronizando um Objeto de Dados de Arquivo Simples

Você pode sincronizar as alterações em uma fonte de dados de arquivo simples externa com seu objeto de dados no Informatica Developer. Use o assistente **Sincronizar Arquivo Simples** para sincronizar os objetos de dados.

1. Na exibição **Object Explorer**, selecione um objeto de dados de arquivo simples.
2. Clique com o botão direito e selecione **Sincronizar**.

   O assistente **Sincronizar Objeto de Dados de Arquivo Simples** é exibido.
3. Verifique o caminho do arquivo simples no campo **Selecionar arquivo simples existente**.
4. Clique em **Avançar**.
5. Como opção, selecione a página de código, o formato, as propriedades de formato delimitado e as propriedades de coluna.
6. Clique em **Concluir** e em**OK**.

Sincronizar um Objeto de Dados Relacionais

Você pode sincronizar alterações de fontes de dados externas de uma fonte de dados relacionais com seu objeto de dados no Informatica Developer. Alterações de fontes de dados externas incluem adicionar, alterar e remover colunas, bem como alterações em regras.

1. Na exibição **Object Explorer**, selecione um objeto de dados relacionais.
2. Clique com o botão direito e selecione **Sincronizar**.

   Uma mensagem solicita que você confirme a ação.
3. Para concluir o processo de sincronização, clique em **OK**. Clique em **Cancelar** para cancelar o processo.

   Se você clicar em **OK**, será exibida uma mensagem de status do processo de sincronização.
4. Quando a mensagem for exibida, clique em **OK**.

   A mensagem exibe um resumo das alterações de metadados feitas para o objeto de dados.

CAPÍTULO 9

Perfis de coluna em fontes de dados semiestruturadas

Este capítulo inclui os seguintes tópicos:

- Visão geral de perfis de coluna em fontes de dados semiestruturadas, 79
- Perfil de coluna em uma fonte de dados JSON ou XML, 80
- Perfil de coluna em uma fonte de dados semiestruturada no HDFS, 81
- Criando uma conexão HDFS, 81
- Criando um objeto de dados de arquivo complexo em um arquivo JSON ou XML no HDFS, 82
- Criando um objeto de dados em uma fonte de dados Avro ou Parquet, 83
- Executando um perfil de coluna em fontes de dados semiestruturadas, 84

Visão geral de perfis de coluna em fontes de dados semiestruturadas

É possível criar objetos de dados a partir de fontes de dados Avro, JSON, Parquet e XML e, em seguida, criar um perfil de coluna nesses objetos de dados.

Os formatos Avro, JSON, Parquet e XML são fontes de dados semiestruturadas. Para usar as fontes de dados semiestruturadas para criar um perfil de coluna, você pode realizar as seguintes tarefas:

1. Criar um objeto de dados físicos na fonte de dados semiestruturada.
2. Criar e executar um perfil de coluna no objeto de dados físicos.

É possível criar objetos de dados de arquivo simples para fontes de dados JSON ou XML. Você pode criar objetos de dados de arquivo complexos para fontes de dados Avro, JSON, Parquet e XML no Hadoop Distributed File System (HDFS).

Perfil de coluna em uma fonte de dados JSON ou XML

Você pode criar um objeto de dados de arquivo simples ou um objeto de dados de arquivo complexo a partir de uma fonte de dados JSON ou XML. Também pode criar e executar um perfil de coluna no objeto de dados.

Crie um arquivo de texto que contenha o caminho da fonte de dados XML ou JSON e use o arquivo de texto como fonte de dados para criar um objeto de dados de arquivo simples. Também é possível adicionar o caminho do arquivo para várias fontes de dados XML ou JSON ao arquivo de texto.

Você pode criar um objeto de dados de arquivo complexo a partir de uma fonte de dados XML ou JSON com um leitor de arquivo complexo. O leitor de arquivo complexo fornece entrada para uma transformação de Processador de Dados que analisa o arquivo e converte os dados de origem em registros de valores simples separados por vírgula.

Nota: A Developer tool não oferece suporte a uma fonte de dados JSON com codificação UTF-8.

Criando um objeto de dados em uma fonte de dados JSON ou XML

Você pode criar um objeto de dados de arquivo simples ou um objeto de dados de arquivo complexo em uma fonte de dados JSON ou XML.

1. Na exibição **Object Explorer** da Developer tool, selecione o projeto no qual você deseja criar o objeto de dados e o perfil de coluna.
2. Clique em **Arquivo** > **Novo** > **Objeto de Dados**.

   A caixa de diálogo **Novo** é exibida.
3. Você pode optar por criar um objeto de dados de arquivo simples ou um objeto de dados de arquivo complexo.
   - Para criar um objeto de dados de arquivo simples, realize as seguintes tarefas:
     1. Selecione **Objetos de Dados Físicos** > **Objeto de Dados de Arquivo Simples** e clique em **Avançar**.
        A caixa de diálogo **Novo Objeto de Dados de Arquivo Simples** é exibida.
     2. Selecione **Criar de um arquivo simples existente** e clique em **Procurar** para escolher o arquivo de texto. Clique em **Avançar**.
     3. Verifique se a página de código é **MS Windows Latin 1 (ANSI), superconjunto de Latin 1** e se o formato está delimitado. Clique em **Avançar**.
     4. Verifique se o delimitador está definido como **vírgula**. Clique em **Concluir**.
   - Para criar um objeto de dados de arquivo complexo, realize as seguintes tarefas:
     1. Selecione **Objetos de Dados Físicos** > **Objeto de Dados de Arquivo Complexo** e clique em **Avançar**.
        A caixa de diálogo **Novo Objeto de Dados de Arquivo Complexo** é exibida.
     2. Adicione um nome para o objeto de dados. Selecione o tipo de acesso, como **Arquivo**.
     3. Clique em **Procurar** para escolher um arquivo JSON ou XML. Clique em **Concluir**.
        Quando o servidor do Developer está no Linux, é necessário atualizar o caminho de arquivo da fonte de dados para a localização nesse servidor. Para atualizar o caminho de arquivo, selecione o objeto de dados de arquivo complexo, clique em **Ler** na guia **Operações de Objeto de Dados** e adicione o caminho de arquivo na guia **Avançado** do painel **Detalhes da Operação de Objeto de Dados**.

   O objeto de dados é exibido na pasta de projeto.

Perfil de coluna em uma fonte de dados semiestruturada no HDFS

É possível criar e executar um perfil de coluna em um arquivo Avro, JSON, Parquet ou XML que utiliza o HDFS. Para ler o arquivo JSON ou XML no HDFS, use um leitor de arquivo complexo para transmitir a entrada XML ou JSON para a transformação de Processador de Dados.

Perfil de coluna em uma fonte de dados XML ou JSON no HDFS

É possível criar um objeto de dados de arquivo complexo a partir de um arquivo JSON ou XML. Também pode criar e executar um perfil de coluna no objeto de dados.

Crie uma conexão com o HDFS antes de criar os objetos de dados para arquivos JSON ou XML no HDFS.

Você pode usar um dos seguintes métodos para criar um objeto de dados de um arquivo JSON ou XML no HDFS:

- Criar um objeto de dados de arquivo complexo em um arquivo JSON ou XML.
- Criar um objeto de dados de arquivo complexo em uma pasta que contém vários arquivos JSON ou XML.

Depois de criar o objeto de dados, você pode criar e executar um perfil de coluna nesse objeto.

Perfil de coluna em uma fonte de dados Avro ou Parquet no HDFS

É possível criar um objeto de dados de arquivo complexo em uma fonte de dados Avro ou Parquet no HDFS. Esse objeto de dados pode ser usado para criar e executar um perfil de coluna.

É possível criar um objeto de dados de arquivo complexo em um arquivo Avro ou Parquet ou em uma pasta que contém vários arquivos Avro ou Parquet. Você pode criar um objeto de dados de arquivo complexo para uma fonte de dados Avro e Parquet com o tipo de acesso de arquivo ou conexão e o formato de recurso como Binário, Avro ou Parquet. É necessário criar uma conexão HDFS antes de criar um objeto de dados de arquivo complexo em fontes de dados Avro e Parquet.

Nota: Você pode escolher o Formato do Recurso como **Avro** ou **Parquet** somente para fontes de dados estruturadas simples Avro e Parquet.

É possível escolher uma das seguintes opções ao criar um objeto de dados a partir de arquivos Avro e Parquet no HDFS:

- Selecionar o tipo de acesso como arquivo e o formato do recurso como Binário.
- Selecionar o tipo de acesso como arquivo e o formato do recurso como Avro ou Parquet.
- Selecionar o tipo de acesso como conexão e o formato do recurso como Avro ou Parquet.

Criando uma conexão HDFS

Configure a conexão HDFS no Informatica Developer para criar um perfil de coluna em fontes de dados Avro, JSON, Parquet e XML no HDFS. Você pode criar um objeto de dados de arquivo complexo depois de criar uma conexão HDFS.

1. Clique em **Janela** > **Preferências**.
2. Selecione **Informatica** > **Conexões**.

3. Expanda o domínio.
4. Selecione o tipo de conexão **Sistemas de Arquivos** > **Sistema de Arquivos Hadoop** e clique em **Adicionar**.
5. Insira um nome de conexão.
6. Opcionalmente, insira uma descrição de conexão.
7. Clique em **Avançar**.
8. Configure as propriedades da conexão.
9. Clique em **Testar Conexão** para verificar a conexão com o HDFS.
10. Clique em **Concluir**.

Criando um objeto de dados de arquivo complexo em um arquivo JSON ou XML no HDFS

Você pode criar um objeto de dados físicos complexo para o arquivo de origem JSON ou XML que usa o HDFS e criar um perfil de coluna nesse objeto físico.

1. Na exibição **Object Explorer** na Developer tool, selecione o projeto no qual deseja criar um objeto de dados físicos e um perfil de coluna.
2. Clique em **Arquivo** > **Novo** > **Objeto de Dados**.

   A caixa de diálogo **Novo** é exibida.
3. Selecione **Objetos de Dados Físicos** > **Objeto de Dados de Arquivo Complexo** e clique em **Avançar**.

   A caixa de diálogo **Novo Objeto de Dados de Arquivo Complexo** é exibida.
4. Adicione um nome para o objeto de dados. Selecione o tipo de acesso, como **Conexão**.
5. É possível criar um objeto de dados em um arquivo JSON ou XML ou em uma pasta que contém vários arquivos JSON ou XML.
   - Para continuar criando um objeto de dados de arquivo complexo em um arquivo JSON ou XML, realize as seguintes tarefas:
     1. Clique em **Procurar** para selecionar uma conexão. Na caixa de diálogo **Adicionar Recurso**, clique em **Adicionar** para escolher um arquivo JSON ou XML. Clique em **Concluir**.
     2. O objeto de dados é exibido na pasta de projeto.
   - Para continuar criando um objeto de dados de arquivo complexo em uma pasta com vários arquivos JSON ou XML, realize as seguintes tarefas:
     1. Clique em **Procurar** para selecionar uma conexão. Na caixa de diálogo **Adicionar Recurso**, clique em **Adicionar** para escolher um arquivo JSON ou XML na pasta. Clique em **Concluir**.
     2. O objeto de dados é exibido na pasta de projeto.
     3. Selecione o objeto de dados na pasta do projeto e clique em **Avançado** > **Tempo de Execução: Leitura** > **Diretório do arquivo de origem**.
     4. Remova o nome de arquivo e mantenha o nome da pasta no caminho de arquivo.

Criando um objeto de dados em uma fonte de dados Avro ou Parquet

É possível criar um objeto de dados físicos complexo para uma fonte de dados Avro ou Parquet com **Arquivo** ou **Conexão** como o tipo de acesso. É possível criar um perfil de coluna no objeto de dados físicos.

1. Na exibição **Object Explorer**, selecione um projeto.
2. Clique em **Arquivo** > **Novo** > **Objeto de Dados**.

   A caixa de diálogo **Novo** é exibida.
3. Selecione **Objetos de Dados Físicos** > **Objeto de Dados de Arquivo Complexo** e clique em **Avançar**.

   A caixa de diálogo **Novo Objeto de Dados de Arquivo Complexo** é exibida.
4. Insira um nome para o objeto de dados.
5. Você pode escolher o tipo de acesso como **Conexão** ou **Arquivo**.
   - Se escolher o tipo de acesso como **Conexão**, realize as seguintes tarefas:
     1. Clique em **Procurar** para escolher uma conexão HDFS.
        A caixa de diálogo **Escolher Conexão** é exibida.
     2. Na caixa de diálogo **Escolher Conexão**, escolha uma fonte de dados e clique em **OK**.
     3. Na caixa de diálogo **Novo Objeto de Dados de Arquivo Complexo**, clique em **Concluir**.
        O objeto de dados é exibido na pasta de projeto.
   - Se você escolher o Tipo de Acesso como **Arquivo** e o Formato do Recurso como **Binário**, realize as seguintes tarefas:
     1. Clique em **Procurar** para escolher um arquivo Avro ou Parquet na máquina local.
     2. Na caixa de diálogo **Novo Objeto de Dados de Arquivo Complexo**, clique em **Concluir**.
        O objeto de dados é exibido na pasta de projeto.
     3. Selecione o objeto de dados na pasta do projeto e clique na exibição **Operações de Objeto de Dados**.
     4. Na exibição **Operações de Objeto de Dados**, clique na guia **Leitura** > **Avançado**.
     5. Na guia **Avançado**, insira o caminho do arquivo da fonte de dados na máquina Linux ou Windows no campo **Caminho do arquivo**.
     6. Insira o Formato do Arquivo como **Entrada Personalizada**.
     7. Insira **com.informatica.avro.AvroToXML** no campo **Formato de Entrada** para fontes de dados Avro e **com.informatica.parquet.ParquetToXML** no campo **Formato de Entrada** para fontes de dados Parquet. Quando você adiciona o formato de entrada, a Transformação de Processador de Dados processa e transforma as fontes de dados no formato Avro ou Parquet em uma fonte de dados no formato XML em tempo de execução.
   - Se você escolher o Tipo de Acesso como **Arquivo** e o Formato do Recurso como **Avro** ou **Parquet**, realize as seguintes tarefas:
     1. Clique em **Procurar** para escolher um arquivo Avro ou Parquet na máquina local.
     2. Na caixa de diálogo **Novo Objeto de Dados de Arquivo Complexo**, clique em **Concluir**.
        O objeto de dados é exibido na pasta de projeto.
     3. Depois de criar o objeto de dados, navegue até a guia **Operações de Objeto de Dados** > **Leitura** > **Avançado** e verifique se o caminho do arquivo no campo **Caminho de arquivo** corresponde à fonte de dados na máquina Linux ou Windows.

Nota: Você pode escolher o Formato do Recurso como **Avro** ou **Parquet** somente para fontes de dados estruturadas simples Avro e Parquet.

Você pode escolher uma pasta com vários arquivos Avro ou Parquet para criar um objeto de dados. Depois de criar o objeto de dados, navegue até a guia **Operações de Objeto de Dados** > **Leitura** > **Avançado** e verifique se o caminho do arquivo no campo **Caminho de arquivo** corresponde à pasta das fontes de dados na máquina Linux ou Windows.

Executando um perfil de coluna em fontes de dados semiestruturadas

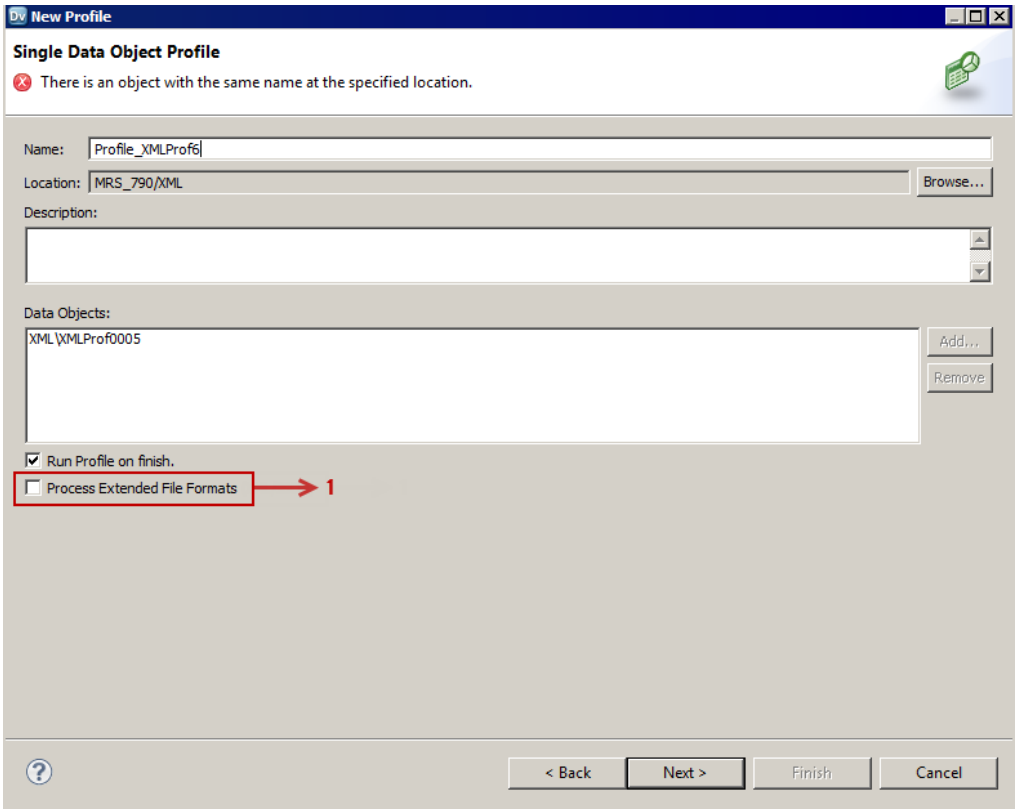
Depois de criar um objeto de dados de arquivo simples ou complexo a partir de fontes de dados Avro, JSON, Parquet ou XML, você pode criar e executar um perfil de coluna no objeto de dados.

1. Na exibição **Object Explorer**, selecione o objeto de dados físicos para o arquivo Avro, JSON, Parquet ou XML.
2. Clique em **Arquivo** > **Novo** > **Perfil**.

   A caixa de diálogo **Novo** é exibida.
3. Selecione **Perfil**. Clique em **Avançar**.

   A caixa de diálogo **Novo Perfil** é exibida.
4. Na caixa de diálogo **Novo Perfil**, adicione um nome para o perfil e uma descrição opcional.
5. Selecione a opção **Processar Formatos de Arquivo Estendidos**. Clique em **Avançar**.

   A seguinte imagem mostra o assistente para **Novo Perfil** com a opção **Processar Formatos de Arquivo Estendidos**:



1. Processar Formatos de Arquivo Estendidos. Selecione essa opção para processar fontes de dados no formato Avro, JSON, Parquet ou XML.

Nota: A opção **Processar Formatos de Arquivo Estendidos** não aparece para fontes de dados Avro e Parquet quando você escolhe o Formato de Recurso como **Avro** ou **Parquet**.

6. Na página **Perfil de Objeto de Dados Simples**, selecione as colunas e opções em **Seleção de Colunas** e **Descoberta de Domínio de Dados** conforme necessário. Clique em **Concluir**.

   Nota: Se a Developer tool estiver instalada em uma máquina Linux, e o objeto de dados físicos JSON ou XML for um objeto de dados de arquivo simples com um arquivo de texto, realize as seguintes tarefas:

   1. Na guia **Visão Geral**, atualize o valor de **Precisão** para incluir o número de caracteres no caminho de arquivo da fonte de dados no servidor.
   2. Atualizar o caminho de arquivo da fonte de dados para a localização no servidor depois de criar um perfil no objeto de dados de arquivo simples. Para atualizar o caminho de arquivo, clique em **Tempo de Execução: Leitura** > **Diretório do arquivo de origem** na guia **Avançado** e adicione o caminho de arquivo.

7. Clique com o botão direito do mouse no perfil e selecione **Executar Perfil**.

   Os resultados do perfil são exibidos.

CAPÍTULO 10

Regras no Informatica Developer

Este capítulo inclui os seguintes tópicos:

- Visão Geral de Regras no Informatica DeveloperDiretrizes para Regras, 86
- Criar uma Regra no Informatica Developer, 87
- Aplicando uma Regra no Informatica DeveloperAplicando uma Regra, 87

Visão Geral de Regras no Informatica DeveloperDiretrizes para Regras

Uma regra é uma lógica comercial que define condições aplicadas a dados de origem quando você executa um perfil. Você pode criar regras reutilizáveis a partir de mapplets na ferramenta Developer. Você pode reutilizar essas regras na ferramenta Analyst para validar dados de origem.Você pode criar regras reutilizáveis a partir de mapplets na ferramenta Developer. Essas regras podem ser usadas em perfis para validar dados de origem.

Crie um mapplet e valide como uma regra. Essa regra aparece como uma regra reutilizável na ferramenta Analyst. Aplique a regra a um perfil da coluna na ferramenta Developer ou na ferramenta Analyst.

Uma regra deve atender aos requisitos a seguir:

- Deve conter uma transformação de Entrada e Saída. Você não pode usar fontes de dados em uma regra.
- Pode conter transformações de Expressão, transformações de Pesquisa e transformações passivas de qualidade de dados. Não pode conter qualquer outro tipo de transformação. Por exemplo, uma regra não pode conter uma transformação de Correspondência, que é uma transformação ativa.
- Ele não especifica cardinalidade entre grupos de entrada.

Uma regra deve atender aos requisitos a seguir:

- Deve conter uma transformação de Entrada e Saída. Você não pode usar fontes de dados em uma regra.
- Ela pode conter transformações de Expressão, transformações de Pesquisa e transformações passivas. Não pode conter qualquer outro tipo de transformação. Por exemplo, uma regra não pode conter uma transformação de Classificador, pois ela é uma transformação ativa.
- Ele não especifica cardinalidade entre grupos de entrada.

Criar uma Regra no Informatica Developer

Valide um mapplet como uma regra para criar uma regra na ferramenta Developer.

Crie um mapplet na ferramenta Developer.

1. Clique com o botão direito do mouse no editor de mapplet.
2. Selecione **Validar como** > **Regra**.

Aplicando uma Regra no Informatica DeveloperAplicando uma Regra

É possível adicionar uma regra a um perfil de colunas salvo. Não é possível adicionar uma regra a um perfil configurado para análise de associação. Você pode adicionar uma regra a um perfil de coluna salvo.

1. Procure a exibição **Object Explorer** e localize o perfil que você precisa.
2. Clique o botão direito do mouse no perfil e selecione **Abrir**.

   O perfil será aberto no editor.
3. Clique na guia **Definição** e selecione Regras.
4. Clique em **Adicionar**.

   A caixa de diálogo **Aplicar Regra** é aberta.
5. Clique em **Procurar** para localizar a regra que você deseja aplicar.

   Selecione uma regra de um projeto de repositório e clique em **OK**.
6. Clique na coluna **Valor** em **Valores de Entrada** para selecionar uma porta de entrada para a regra.
7. Opcionalmente, clique na coluna **Valor** em **Valores de Saída** para editar o nome da porta de saída da regra.

   A regra aparece na guia **Definição**.

CAPÍTULO 11

Coluna de Resultados de Perfil no Informatica Developer

Este capítulo inclui os seguintes tópicos:

- Resultados do Perfil de Colunas no Informatica DeveloperResultados do Perfil de Colunas, 88
- Propriedades de Valor de Coluna, 89
- Propriedades de Padrão de Coluna, 90
- Propriedades de Estatísticas de Coluna, 90
- Propriedades de tipo de dados da coluna, 91
- Curadoria no Informatica DeveloperCuradoria no Informatica Developer, 92
- Exportar Resultados de Perfil do Informatica Developer, 93

Resultados do Perfil de Colunas no Informatica DeveloperResultados do Perfil de Colunas

Um analista de perfil de coluna fornece informações sobre a qualidade de dados realçando frequências de valores, padrões e estatísticas de dados.

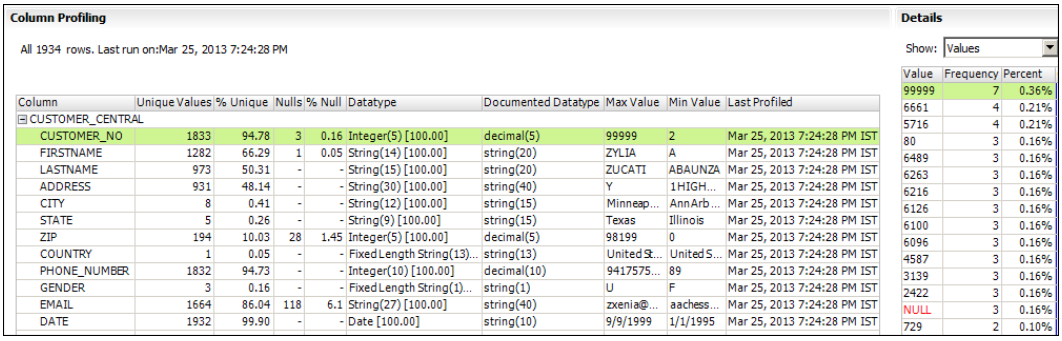
A tabela a seguir descreve os resultados do perfil para cada tipo de análise:

A análise da criação de perfil de colunas gera os seguintes resultados do perfil:

Tipo de perfil	Resultados do perfil
Perfil de coluna	- Estatísticas de porcentagem e contagem para valores nulos e únicos - Tipos de dados inferidos - O tipo de dados que a fonte de dados declara para os dados - O número máximo e mínimo de valores - A data e hora da última execução do perfil - Estatísticas de porcentagem e contagem para cada elemento de dados único na coluna - Estatísticas de porcentagem e contagem para cada padrão de caractere único na coluna
Perfil de chave primária	- Chaves primárias Inferidas - Violações de chave
Perfil de dependência funcional	- Dependências funcionais Inferidas - Violações de dependência funcional

- Estatísticas de porcentagem e contagem para valores nulos e únicos
- Tipos de dados inferidos
- O tipo de dados que a fonte de dados declara para os dados
- O número máximo e mínimo de valores
- A data e hora da última execução do perfil
- Estatísticas de porcentagem e contagem para cada elemento de dados único na coluna
- Estatísticas de porcentagem e contagem para cada padrão de caractere único na coluna

A seguinte imagem mostra os resultados do perfil de coluna:



Column Profiling

All 1934 rows. Last run on:Mar 25, 2013 7:24:28 PM

Column	Unique Values	% Unique	Nulls	% Null	Datatype	Documented Datatype	Max Value	Min Value	Last Profiled
CUSTOMER_CENTRAL									
CUSTOMER_NO	1833	94.78	3	0.16	Integer(5) [100.00]	decimal(5)	99999	2	Mar 25, 2013 7:24:28 PM IST
FIRSTNAME	1282	66.29	1	0.05	String(14) [100.00]	string(20)	ZYLIA	A	Mar 25, 2013 7:24:28 PM IST
LASTNAME	973	50.31	-	-	String(15) [100.00]	string(20)	ZUCATI	ABAUNZA	Mar 25, 2013 7:24:28 PM IST
ADDRESS	931	48.14	-	-	String(30) [100.00]	string(40)	Y	1HIGH...	Mar 25, 2013 7:24:28 PM IST
CITY	8	0.41	-	-	String(12) [100.00]	string(15)	Minneap...	AnnArb...	Mar 25, 2013 7:24:28 PM IST
STATE	5	0.26	-	-	String(9) [100.00]	string(15)	Texas	Illinois	Mar 25, 2013 7:24:28 PM IST
ZIP	194	10.03	28	1.45	Integer(5) [100.00]	decimal(5)	98199	0	Mar 25, 2013 7:24:28 PM IST
COUNTRY	1	0.05	-	-	Fixed Length String(13)...	string(13)	United St...	United S...	Mar 25, 2013 7:24:28 PM IST
PHONE_NUMBER	1832	94.73	-	-	Integer(10) [100.00]	decimal(10)	9417575...	89	Mar 25, 2013 7:24:28 PM IST
GENDER	3	0.16	-	-	Fixed Length String(1)...	string(1)	U	F	Mar 25, 2013 7:24:28 PM IST
EMAIL	1664	86.04	118	6.1	String(27) [100.00]	string(40)	zxenia@...	aachess...	Mar 25, 2013 7:24:28 PM IST
DATE	1932	99.90	-	-	Date [100.00]	string(10)	9/9/1999	1/1/1995	Mar 25, 2013 7:24:28 PM IST

Details

Show: Values

Value	Frequency	Percent
99999	7	0.36%
6661	4	0.21%
5716	4	0.21%
80	3	0.16%
6489	3	0.16%
6263	3	0.16%
6216	3	0.16%
6126	3	0.16%
6100	3	0.16%
6096	3	0.16%
4587	3	0.16%
3139	3	0.16%
2422	3	0.16%
NULL	3	0.16%
729	2	0.10%

Propriedades de Valor de Coluna

As propriedades de valor de coluna mostram os valores nas colunas de perfil e a frequência com que cada valor é exibido em cada coluna. As frequências são mostradas como um número, uma porcentagem e um gráfico de barras.

Para exibir propriedades de valores de coluna, selecione Valores na lista **Mostrar**. Clique duas vezes no valor de coluna para fazer uma busca detalhada das linhas que contêm o valor.

A tabela a seguir descreve as propriedades de valores de coluna:

Propriedade	Descrição
Valores	Lista de todos os valores para a coluna no perfil.
Frequência	Número de vezes que um valor é exibido em uma coluna.
Porcentagem	Número de vezes que um valor é exibido em uma coluna, expresso como uma porcentagem de todos os valores na coluna.
Gráfico	Gráfico de barras para a porcentagem.

Propriedades de Padrão de Coluna

As propriedades do padrão de coluna mostram os padrões de dados nas colunas de perfil e a frequência com que os padrões são exibidos em cada coluna. Os padrões são mostrados como um número, uma porcentagem e um gráfico de barras.

Para exibir informações de padrão, selecione Padrões na lista **Mostrar**. Clique duas vezes no padrão para fazer uma busca detalhada das linhas que contêm o padrão.

A tabela a seguir descreve as propriedades para os padrões de valor de coluna:

Propriedade	Descrição
Padrões	Padrão para a coluna selecionada
Frequência	Número de vezes que um padrão é exibido em uma coluna.
Porcentagem	Número de vezes que um padrão é exibido em uma coluna, expresso como uma porcentagem de todos os valores na coluna.
Gráfico	Gráfico de barras para a porcentagem.

Propriedades de Estatísticas de Coluna

Estatísticas de coluna incluem propriedades, como tamanhos mínimo e máximo de valores e primeiro e último valores.

Para exibir informações estatísticas, selecione Estatísticas na lista **Mostrar**.

A tabela a seguir descreve as propriedades de estatísticas de coluna:

Propriedade	Descrição
Tamanho Máximo	Tamanho do valor mais longo na coluna.
Tamanho Mínimo	Tamanho do valor mais curto na coluna.
Inferior	Últimos cinco valores na coluna.
Superior	Primeiros cinco valores na coluna.
Soma	Soma de todos os valores na coluna com um tipo de dados numérico.

Nota: O perfil também exibe as estatísticas de desvio padrão e médio para colunas do tipo Número inteiro.

Propriedades de tipo de dados da coluna

Os tipos de dados de coluna incluem todos os tipos de dados inferidos de cada coluna nos resultados de perfil.

Para exibir informações do tipo de dados, selecione **Tipos de dados** na lista **Mostrar**. Clique duas vezes em um tipo de dados para realizar uma busca detalhada nas linhas que contêm o tipo de dados.

A seguinte tabela descreve as propriedades dos tipos de dados de coluna:

Propriedade	Descrição
Tipo de dados	Lista de todos os tipos de dados inferidos para a coluna no perfil.
Frequência	Número de vezes que um tipo de dados de uma coluna é exibido, expresso como um número.
% de conformidade	Porcentagem na qual o tipo de dados de uma coluna é exibido.
Status	Indica o status do tipo de dados. Os status são inferido, aprovado ou rejeitado. **Inferido** Indica o tipo de dados da coluna inferido pela Developer tool. **Aprovado** Indica um tipo de dados aprovado da coluna. Ao aprovar um tipo de dados, você confirma esse tipo de dados no repositório do modelo. **Rejeitado** Indica um tipo de dados rejeitado da coluna.

Curadoria no Informatica DeveloperCuradoria no Informatica Developer

Curadoria é o processo de validação e gerenciamento dos metadados descobertos em uma fonte de dados de forma que os metadados estejam prontos para uso e preparação de relatórios. Ao fazer a curadoria de metadados no Informatica Developer, você pode aprovar, rejeitar e redefinir os tipos de dados ou domínios de dados inferidos nos resultados do perfil.

Você pode aprovar um tipo de dados ou domínio de dados de uma coluna. Você pode ocultar os tipos de dados rejeitados ou os domínios de dados para uma coluna. Depois de aprovar ou rejeitar um tipo de dados ou domínio de dados inferido, você pode redefinir o tipo de dados ou domínio de dados para restaurar o status inferido.

Aprovando tipos de dadosAprovando tipos de dados no Informatica Developer

Os resultados do perfil incluem os tipos de dados inferidos, a frequência, a porcentagem de conformidade e o status de inferência para cada coluna na fonte de dados. É possível escolher e aprovar um único tipo de dados para cada coluna.

1. Na exibição **Object Explorer**, selecione e abra um perfil.
2. Verifique se você está na guia **Resultados**.
3. Na exibição **Criação de Perfil de Coluna**, selecione uma coluna para exibir frequências de valor, padrões, tipos de dados e estatísticas no painel direito.
4. No painel **Detalhes**, selecione **Tipos de dados** na lista **Mostrar**.

   Os tipos de dados inferidos da coluna são exibidos.
5. Clique com o botão direito do mouse na coluna que você deseja aprovar e clique em **Aprovar**.

   O status do tipo de dados muda para **Aprovado**.
6. Para restaurar o status inferido do tipo de dados, clique com o botão direito do mouse no tipo de dados e depois clique em **Redefinir**.

Rejeitando tipos de dadosRejeitando tipos de dados no Informatica Developer

O Informatica Developer exibe tipos de dados inferidos nos resultados do perfil por padrão. Você pode rejeitar tipos de dados inferidos ou aprovados. Você pode optar por exibir ou ocultar os tipos de dados rejeitados.

1. Na exibição **Object Explorer**, selecione um perfil.
2. Clique duas vezes no perfil para abri-lo.

   O perfil é exibido em uma guia.
3. Na exibição **Criação de Perfil de Coluna**, selecione uma linha.
4. Para rejeitar tipos de dados da coluna inferidos, selecione a exibição **Tipos de dados** no painel direito. Selecione o tipo de dados inferido que você deseja rejeitar, clique com o botão direito do mouse na linha e selecione **Rejeitar**.

   O Informatica Developer esmaece o tipo de dados rejeitado na lista de tipos de dados.

5. Para ocultar os tipos de dados rejeitados, clique com o botão direito do mouse na linha e selecione **Ocultar Rejeitados**.
6. Para exibir os tipos de dados rejeitados, clique com o botão direito do mouse em uma das linhas e selecione **Mostrar Rejeitados**.

Exportar Resultados de Perfil do Informatica Developer

É possível exportar resultados do perfil de coluna para um arquivo .csv ou do Microsoft Excel. Quando você exporta os resultados do perfil para um arquivo do Microsoft Excel, a Developer tool salva as informações em um arquivo .xlsx.

1. Na exibição **Object Explorer**, abra um perfil.
2. Opcionalmente, execute o perfil para atualizar os resultados do perfil.
3. Selecione a exibição de **Resultados**.
4. Selecione uma coluna.
5. Em **Detalhes**, selecione **Valores**, **Padrões** ou **Tipos de dados** e clique no ícone **Exportar**.

   A caixa de diálogo **Exportar os dados para um arquivo** será exibida.
6. Aceite ou altere o nome de arquivo padrão.
7. Selecione o tipo de dados a ser exportado. Você pode selecionar **Valores para a coluna selecionada**, **Padrões para a coluna selecionada**, **Tipos de dados para a coluna selecionada** ou **Tudo (Resumo, Valores, Padrões, Tipos de Dados, Estatísticas, Propriedades)**.
8. Clique em **Procurar** para selecionar uma localização e salvar o arquivo localmente no computador.
9. Se você não quiser exportar nomes de campo como a primeira linha, desmarque a caixa de seleção **Exportar nomes de campo como primeira linha**.
10. Clique em **OK**.

CAPÍTULO 12

Scorecards no Informatica Developer

Este capítulo inclui os seguintes tópicos:

- Visão Geral dos Scorecards no Informatica Developer, 94
- Criar um Scorecard, 94
- Exportar um Arquivo de Recursos para a Linhagem de Scorecard, 95
- Exibir a Linhagem de Scorecard do Informatica Developer, 95

Visão Geral dos Scorecards no Informatica Developer

Um scorecard é uma representação gráfica das medidas de qualidade em um perfil. É possível exibir os scorecards na ferramenta Developer. Depois de criar um scorecard na ferramenta Developer, conecte-se à ferramenta Analyst para abrir o scorecard para edição. Execute o scorecard nos dados atuais no objeto de dados ou nos dados preparados no depósito de criação de perfil.

Você pode editar e executar um scorecard e exibir a linhagem de do scorecard para uma métrica ou grupo de métrica na ferramenta Analyst.

Criar um Scorecard

Crie um scorecard e adicione colunas a partir de um perfil para o scorecard. Antes de adicionar colunas ao scorecard, execute um perfil.

1. Na exibição do **Object Explorer**, selecione o projeto ou pasta onde você deseja criar o scorecard.
2. Clique em **Arquivo** > **Novo** > **Scorecard**.

   A caixa de diálogo **Novo Scorecard** é exibida.
3. Clique em **Adicionar**.

   A caixa de diálogo **Selecionar Perfil** é exibida. Selecione o perfil que contém as colunas que você deseja adicionar.
4. Clique em **OK** e em **Avançar**.

5. Selecione as colunas que você deseja adicionar ao scorecard.

   Por padrão, o assistente de scorecard seleciona as colunas e regras definidas no perfil. Não é possível adicionar colunas que não são incluídas no perfil.

6. Clique em **Concluir**.

   A ferramenta Developer cria o scorecard.

7. Opcionalmente, clique em **Abrir com o Informatica Analyst** para estabelecer uma conexão com a ferramenta Analyst e abra o scorecard na ferramenta Analyst.

Exportar um Arquivo de Recursos para a Linhagem de Scorecard

Você pode exportar um projeto que contém os scorecards e objetos dependentes como um arquivo de recursos do Metadata Manager. Use o recurso de arquivo exportado no formato XML para criar e carregar um recurso para a linhagem de scorecard no Metadata Manager.

1. Para abrir o assistente **Exportar**, clique em **Arquivo** > **Exportar**.
2. Selecione **Informatica** > **Arquivo de Recursos do Metadata Manager**.
3. Clique em **Avançar**.
4. Clique em **Procurar** para selecionar um projeto que contenha os objetos de scorecard e a linhagem que você precisa exportar.
5. Clique em **Avançar**.
6. Selecione os objetos de scorecard que você deseja exportar.
7. Insira o nome do arquivo de exportação e a localização do arquivo.
8. Para exibir os objetos dependentes que o assistente **Exportar** exporta com os objetos que você selecionou, clique em **Avançar**.

   O assistente **Exportar** exibe objetos dependentes.

9. Clique em **Concluir**.

   A ferramenta Developer exporta os objetos para o arquivo XML.

Exibir a Linhagem de Scorecard do Informatica Developer

Para exibir a linhagem de scorecard para uma métrica ou grupo de métrica da ferramenta Developer, inicie a ferramenta Analyst.

1. Na exibição do **Object Explorer**, selecione o projeto ou pasta que contém o scorecard.
2. Clique duas vezes no scorecard para abri-lo.

   O scorecard é exibido em uma guia.

3. Clique em **Abrir com o Informatica Analyst**.

A ferramenta Analyst é exibida na janela do navegador.

4. Na exibição do **Scorecard** da ferramenta Analyst, selecione uma métrica ou grupo de métrica.
5. Clique duas vezes e selecione **Exibir a Linhagem**.

   O diagrama de linhagem do scorecard é exibido em uma caixa de diálogo.

CAPÍTULO 13

Criação de Perfil de Mapplet e Mapeamento

Este capítulo inclui os seguintes tópicos:

- Visão Geral da Criação de Perfil de Mapplet e MapeamentoPerfis de Mapplet e Mapeamento, 97
- Executar um Perfil em um Mapplet ou Objeto de Mapeamento, 97
- Comparar Perfis de Objetos de Mapplet ou de Mapeamento, 98
- Gerar um Mapeamento de Perfil, 98

Visão Geral da Criação de Perfil de Mapplet e MapeamentoPerfis de Mapplet e Mapeamento

Você pode definir um perfil da coluna para um objeto em um mapplet ou mapeamento. Execute um perfil em um objeto de mapplet ou de mapeamento para verificar o design do mapplet ou do mapeamento sem salvar os resultados do perfil. Você também pode gerar um mapeamento de um perfil.

Executar um Perfil em um Mapplet ou Objeto de Mapeamento

Quando você executa um perfil em um mapplet ou em um objeto de mapeamento, o perfil é executado em todas as colunas de dados e permite operações de busca detalhada nos dados que são preparados para o objeto de dados. É possível executar um perfil em um mapplet ou objeto de mapeamento com várias portas de saída.Quando você executa um perfil em um mapplet ou objeto de mapeamento, esse perfil é executado em todas as colunas de dados e permite operações de busca detalhada nesses dados. Você pode executar um perfil em um mapplet ou objeto de mapeamento com várias portas de saída.

O perfil rastreia os dados de origem por meio de mapeamento para as portas de saída do objeto selecionado. O perfil analisa os dados que apareceriam nessas portas se você executasse o mapeamento.

1. Abra um mapplet ou mapeamento.
2. Verifique se o mapplet ou mapeamento é válido.

3. Clique com o botão direito do mouse em um objeto de dados ou transformação e selecione **Criar Perfil Agora**.

   Se a transformação tiver vários grupos de saída, a caixa de diálogo **Selecionar Grupo de Saída** é exibida. Se a transformação tem um único grupo de saída, os resultados de perfil são exibidos na guia **Resultados** do perfil.
4. Se a transformação tiver vários grupos de saída, selecione os grupos de saída conforme necessário.
5. Clique em **OK**.

   Os resultados do perfil aparecem na guia **Resultados** do perfil.

Comparar Perfis de Objetos de Mapplet ou de Mapeamento

Você pode criar um perfil que analisa dois objetos em um mapplet ou mapeamento e compara os resultados dos perfis da coluna para esses objetos.

Como os perfis de objetos únicos de mapeamento ou de mapplet, as comparações de perfis são executadas em todas as colunas de dados e ativam operações de busca detalhada nos dados de fase dos objetos de dados.

Como perfis de objetos únicos de mapplet ou mapeamento, comparações de perfil são executadas em todas as colunas de dados.

1. Abra um mapplet ou mapeamento.
2. Verifique se o mapplet ou mapeamento é válido.
3. Pressione a tecla **CTRL** e clique em dois objetos no editor.
4. Clique com o botão direito do mouse em um dos objetos e selecione **Comparar Perfis**.
5. Opcionalmente, configure a comparação de perfil para corresponder colunas de um objeto com outro.
6. Opcionalmente, corresponda as colunas clicando em uma coluna em um objeto e arrastando-a para uma coluna em outro objeto.
7. Opcionalmente, escolha se o perfil analisa todas as colunas ou somente as colunas correspondentes.
8. Clique em **OK**.

Gerar um Mapeamento de Perfil

É possível criar um objeto de mapeamento de um perfil. Use o objeto de mapeamento que você cria para desenvolver um mapeamento válido. O mapeamento que você cria tem uma fonte de dados com base no objeto de perfil e pode conter transformações com base na lógica da regra de perfil. Depois de criar o mapeamento, adicione objetos para concluir.

1. Na exibição do **Object Explorer**, localize o perfil no qual o mapeamento será criado.
2. Clique com o botão direito do mouse no nome do perfil e selecione **Gerar Mapeamento**.

   A caixa de diálogo **Gerar Mapeamento** é exibida.
3. Insira um nome de mapeamento. Se quiser, insira uma descrição para o mapeamento.

4. Confirme a localização da pasta para o mapeamento.

   Por padrão, a ferramenta Developer cria o mapeamento na pasta **Mapeamentos** no mesmo projeto que o perfil. Clique em **Procurar** para selecionar uma localização diferente para o mapeamento.

5. Confirme a definição de perfil que a ferramenta Developer usa para criar o mapeamento. Para usar outro perfil, clique em **Selecionar Perfil**.

6. Clique em **Concluir**.

   O mapeamento aparece no **Object Explorer**.

Adicione objetos ao mapeamento para concluir.

ÍNDICE

C

criação de perfil
  ferramentas 13
  gerenciamento de versões e bloqueios 16
  processo 13
Criação de Perfil de Mapplet e Mapeamento
  Visão Geral 97
criando um perfil de coluna
  perfis 25
criando uma regra de expressão
  regras 33
curadoria
  conceitos 18
  Informatica Analyst 53
  Informatica Developer 92
  tarefas 18

E

exceção
  detectando 42
exportar
  linhagem de scorecard para XML 95

G

gráficos de tendência
  custo 66
  exibindo 67
  exportar do Informatica Analyst 68
  pontuação 66

I

Informatica Analyst
  gerenciamento de versões e bloqueios 25
  regras 30
  resultados de perfil de coluna 34
  resultados do perfil de coluna 46
  visão geral de perfis da coluna 21, 48
Informatica Developer
  regras 86

L

linhagem de scorecard
  exibir no Informatica Analyst 73
  exibir no Informatica Developer 95

O

objeto de dados de arquivo simples
  sincronizando 28
objeto de dados de tabela
  sincronizando 29
objeto de mapeamento
  executar um perfil 97

P

perfil
  componentes 13
  formatos Avro ou Parquets 81
  formatos XML e JSON 80, 81
perfil da coluna
  busca detalhada 52
perfil de coluna
  Informatica Developer 75
  opções 16
  perfil do sistema operacional 24
  processo 22
  visão geral 15
perfis
  criando um perfil de coluna 25
  editar um perfil de coluna 26
  executando 27, 47, 48
  introdução 12
Perfis de Mapplet e Mapeamento
  Visão Geral 97
perfis de objeto de dados
  criando vários perfis 77
  criar um perfil único 76

R

regras
  aplicando no PowerCenter Express 87
  aplicar no Informatica Developer 87
  aplicar uma regra predefinida 31
  criando uma regra de expressão 33
  criar no Informatica Developer 87
  expressão 32
  pré-requisitos 86
  predefinido 31
regras de expressão
  processo 32
regras predefinidas
  processo 31
resultados de perfil
  tipos de dados de coluna 91
resultados de perfil da coluna na ferramenta Analyst
  interface 50

resultados de perfil de coluna na ferramenta Analyst
detalhes da coluna 37, 51
interface 49
resumo 35
resultados de scorecard
exportando 69
exportar do Informatica Analyst 69
exportar para o Excel 69
resultados do perfil
aprovando tipos de dados 53
aprovando tipos de dados no Informatica Developer 92
Excel 54
exibição detalhada 38
exibição resumida 36
exportando 54
exportando do Informatica Analyst 55
exportar no Informatica Developer 93
fazer busca detalhada 52
padrões de coluna 42
rejeitando tipos de dados 53
rejeitando tipos de dados na Developer tool 92
resumo 49, 51
tipos de dados de coluna 40
valores de coluna 43
resultados do perfil de coluna
Informatica Developer 88
perfil de coluna 88
resultados do perfil de coluna na ferramenta Analyst
interface 36

S

scorecard
configurar notificações 71
definir as configurações globais de notificação 72
scorecards
adicionando colunas a um scorecard 59
criar um grupo de métrica 63
custo de dados inválidos 62
custo fixo 62
custo variável 62
definir limites 63
editando 61
editar um grupo de métrica 64
excluir um grupo de métrica 64
executando 60
exibindo 61
fazer busca detalhada 65
gráfico de tendências 65
grupos de métrica 63
Informatica Analyst 56
Informatica Developer 94
métricas 62
mover pontuações 64
notificações 70
pesos de métrica 62
processo do Informatica Analyst 57
visão geral 17