Informatica®
10.1.1

Guia de Descoberta de Dados

Informatica Guia de Descoberta de Dados
10.1.1
Dezembro 2016

© Copyright Informatica LLC 2011, 2018

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

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

Partes desta documentação e/ou software estão sujeitas a copyright 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.

As informações contidas neste documento estão sujeitas a alteração sem aviso prévio. Se você encontrar problemas nesta documentação, informe-nos por escrito e envie para Informatica LLC 2100 Seaport Blvd. Redwood City, CA 94063.

A INFORMATICA LLC FORNECE AS INFORMAÇÕES NESTE DOCUMENTO "COMO ESTÃO" SEM GARANTIA DE QUALQUER TIPO, EXPRESSA OU IMPLÍCITA, INCLUINDO, SEM QUAISQUER GARANTIAS DE COMERCIALIZAÇÃO, ADEQUAÇÃO A UM DETERMINADO FIM E QUALQUER GARANTIA OU CONDIÇÃO DE NÃO-VIOLAÇÃO.

Data da Publicação: 2018-07-03

Conteúdo

Prefácio. 14

Recursos da Informatica. 14

Rede da Informatica. 14

Base de Dados de Conhecimento da Informatica. 14

Documentação da Informatica. 15

Matrizes de Disponibilidade de Produto Informatica. 15

Informatica Velocity. 15

Informatica Marketplace. 15

Suporte global a clientes Informatica. 15

Parte I: Introdução à Descoberta de Dados. 16

Capítulo 1: Introdução à Criação de Perfil. 17

Visão Geral da Criação de Perfil. 17

Arquitetura de Criação de Perfil. 18

Processo de Descoberta de Dados. 20

Capítulo 2: Descoberta de Dados. 22

Visão Geral da Descoberta de Dados. 22

Tipos de perfil e análise. 22

Componentes de Criação de Perfil. 23

Resultados do Perfil. 24

Capítulo 3: Conceitos de Perfil de Coluna. 26

Visão Geral de Conceitos de Perfil de ColunaVisão Geral de Perfis. 26

Opções de Perfil de Coluna. 27

Gerenciamento de Perfis com Controle de Versão e Bloqueios de Perfis de Repositório. 27

Scorecards. 28

Capítulo 4: Conceitos de Descoberta do Domínio de Dados. 29

Conceitos de Descoberta do Domínio de Dados - Visão Geral. 29

Domínios de Dados. 30

Grupos de Domínios de Dados. 30

Glossário do Domínio de Dados. 31

Processo de Descoberta do Domínio de Dados. 31

Capítulo 5: Conceitos de Curadoria. 32

Visão Geral dos Conceitos de Curadoria. 32

Curadoria para Analistas e Desenvolvedores. 32

Tarefas de Curadoria. 33

Parte II: Descoberta de Dados com o Informatica Analyst. 35

Capítulo 6: Perfis da Coluna no Informatica Analyst. 36

Visão Geral dos Perfis de Coluna no Informatica Analyst. 36

Processo de Criação de Perfil de Coluna. 37

Opções de Perfil. 37

Opções de Amostragem. 38

Opções de Busca Detalhada. 38

Ambiente de tempo de execução. 38

Ambiente Nativo. 39

Ambiente Hadoop. 39

Visão geral de perfis de sistema operacional no Informatica Analyst. 40

Selecionando um perfil do sistema operacional. 40

Visão Geral de Bloqueios de Ativos de Repositório e do Desenvolvimento Baseado em Equipe. . . . 40

Criando um perfil de coluna no Informatica Analyst. 41

Editar um Perfil de Coluna. 42

Executar um Perfil. 43

Sincronizando um Objeto de Dados de Arquivo Simples. 44

Sincronizando um Objeto de Dados Relacional. 45

Capítulo 7: Regras no Informatica Analyst. 46

Regras na Visão Geral do Informatica Analyst. 46

Regras em um perfil de coluna. 46

Regras Predefinidas. 47

Processo de Regras Predefinidas. 47

Aplicar uma Regra Predefinida. 47

Regras de Expressão. 48

Processo de Regras de Expressão. 48

Criando uma Regra de Expressão. 49

Capítulo 8: Filtros no Informatica Analyst. 50

Filtros na visão geral do Informatica Analyst. 50

Criando um filtro. 50

Criando um filtro simples. 51

Criando um filtro avançado. 52

Criando um filtro SQL. 53

Gerenciando filtros. 53

Capítulo 9: Resultados do Perfil da Coluna no Informatica Analyst. 55

Resultados de perfil de coluna na visão geral do Informatica Analyst. 55

Exibição Resumida. 56

Propriedades da exibição resumida. 57

Filtros padrão na exibição resumida. . . . 57
Exibição Detalhada. . . . 58
Painéis da exibição detalhada. . . . 59
Estatísticas. . . . 60
Visualização de Dados. . . . 61
Tipos de dados. . . . 61
Exceções. . . . 62
Padrões. . . . 63
Valores. . . . 64
Tipos de execução de perfil. . . . 67
Execução de perfil mais recente. . . . 67
Execução de perfil histórico. . . . 67
Execução de perfil consolidada. . . . 67
Selecionando uma execução de perfil. . . . 68
Visão geral dos resultados de vários perfis de comparação. . . . 69
Comparando resultados de vários perfis. . . . 69
Exibição Resumida dos Resultados de Perfil de Comparação 70
Exibição detalhada dos resultados do perfil de comparação 72
Busca Detalhada de Perfil da Coluna. . . . 73
Fazer Busca Detalhada na Linha de Dados. . . . 74
Aplicar Filtros aos Dados de Busca Detalhada. . . . 74
Curadoria na ferramenta Analyst. . . . 74
Aprovando tipos de dados e domínios de dados. . . . 75
Rejeitando tipos de dados e domínios de dados. . . . 75
Arquivos de exportação de perfil de coluna no Informatica Analyst. . . . 75
Resultados de Exportação de Perfil em um Arquivo CSV. . . . 76
Resultados de Exportação de Perfil no Microsoft Excel. . . . 76
Exportar Resultados de Perfil do Informatica Analyst. . . . 76

Capítulo 10: Termos comerciais, comentários e marcas no Informatica Analyst. . . . 78

Termos comerciais, comentários e marcas na visão geral do Informatica Analyst. . . . 78
Termos Comerciais. . . . 78
Atribuindo termos comerciais a colunas. . . . 79
Comentários. . . . 79
Adicionando comentários a um perfil ou colunas. . . . 79
Marcas. . . . 80
Atribuindo marcas a um perfil ou colunas. . . . 80

Capítulo 11: Scorecards no Informatica Analyst. . . . 81

Scorecards na Visão Geral do Informatica Analyst. . . . 81
Processo de Scorecard no Informatica Analyst. . . . 82
Criando um scorecard no Informatica Analyst. . . . 83

Adicionando colunas a um scorecard existente. 84
Executar um Scorecard. 85
Exibir um Scorecard. 86
Editar um Scorecard. 86
Métricas. 87
Pesos de Métrica. 87
Valor da Qualidade de Dados. 87
Definir Limites. 88
Grupos de Métrica. 88
Criar um Grupo de Métrica. 88
Mover Pontuações para um Grupo de Métrica. 89
Editar um Grupo de Métrica. 89
Excluir um Grupo de Métrica. 90
Buscas Detalhadas em Colunas. 90
Gráficos de Tendência. 90
Gráfico de Tendência de Pontuação. 91
Gráfico de Tendência de Custo. 91
Exibir Gráficos de Tendência. 92
Exportando Gráficos de Tendência. 93
Painel de scorecards no Informatica Analyst. 94
Scorecards por Projeto. 95
Tendência de Execução do Scorecard. 96
Objetos de Dados com Scorecards. 96
Tendências de Métricas Cumulativas. 97
Arquivos de Exportação de Scorecard no Informatica Analyst. 99
Resultados da Exportação do Scorecard no Microsoft Excel. 99
Exportando Resultados do Scorecard do Informatica Analyst. 99
Notificações de Scorecard. 99
Modelo de Mensagem de Notificação de E-mail. 100
Configurar as Notificações de Scorecard. 101
Definir as Configurações Globais para Notificações de Scorecard. 102
Linhagem de Scorecard. 102
Exibir a Linhagem de Scorecard o Informatica Analyst. 103

Capítulo 12: Descoberta do Domínio de Dados no Informatica Analyst. 104

Descoberta do Domínio de Dados no Informatica Analyst - Visão Geral. 104
Glossário do Domínio de Dados no Informatica Analyst. 104
Criando um Grupo do Domínio de Dados no Informatica Analyst. 105
Criando um Domínio de Dados no Informatica Analyst. 105
Criando um Domínio de Dados a Partir de Resultados do Perfil no Informatica Analyst. 106
Localizar Domínios de Dados e Grupos de Domínios de Dados no Informatica Analyst. 106
Opções de Descoberta do Domínio de Dados no Informatica Analyst. 107
Seleção de Colunas de Domínio de Dados no Informatica Analyst. 107

Seleção de Domínio de Dados no Informatica Analyst. 108
Opções de Inferência de Domínio de Dados no Informatica Analyst. 108
Criando um Perfil para Executar Descoberta do Domínio de Dados no Informatica Analyst. 110
Editando um Perfil no Informatica Analyst. 111
Executando um Perfil para Realizar a Descoberta do Domínio de Dados. 112
Resultados da Descoberta do Domínio de Dados no Informatica Analyst. 112
Aprovando Domínios de Dados. 113
Rejeitando Domínios de Dados. 113
Arquivos de Exportação de Descoberta do Domínio de Dados no Informatica Analyst. 113
Resultados da Descoberta do Domínio de Dados no Microsoft Excel. 114
Exportando Resultados da Descoberta do Domínio de Dados do Informatica Analyst. 114

Capítulo 13: Descoberta Empresarial no Informatica Analyst. 115

Descoberta Empresarial no Informatica Analyst - Visão Geral. 115
Processo de descoberta empresarial no Informatica Analyst. 116
Opções de Configuração para Descoberta Empresarial. 116
Configurações da Descoberta do Domínio de Dados. 116
Configurações de Perfil de Coluna. 117
Criando um perfil de descoberta empresarial no Informatica Analyst. 118
Editando Opções de Descoberta Empresarial. 119

Capítulo 14: Resultados da Descoberta Empresarial no Informatica Analyst 121

Visão geral dos resultados da descoberta empresarial no Informatica Analyst. 121
Exibição Resumida. 121
Resultados de Perfil da Exibição Resumo. 122
Exibindo Resultados da Descoberta do Domínio de Dados. 123
Exibindo Resultados do Perfil de Coluna. 123
Conflito de Tipo de Dados. 124
Exibindo Conflitos de Tipos de Dados. 124
Exibição Perfis. 125
Exibindo Propriedades do Perfil. 125

Capítulo 15: Pesquisa de Descoberta no Informatica Analyst. 126

Pesquisa de Descoberta no Informatica Analyst - Visão Geral. 126
Pré-requisitos de Pesquisa de Descoberta. 127
Processo de Pesquisa de Descoberta no Informatica Analyst. 127
Opções de Pesquisa de Descoberta. 128
Critérios de Pesquisa de Descoberta. 129
Pesquisando um Ativo. 129
Resultados da Pesquisa de Descoberta no Informatica Analyst. 129
Painel de Resultados da Pesquisa de Descoberta. 130
Filtrando os Resultados da Pesquisa de Descoberta. 131
Tipos de Correspondência. 131

Correspondência Direta. 132
Correspondência Indireta. 132
Exibindo Informações de Correspondência. 132
Abrindo Ativos dos Resultados da Pesquisa de Descoberta. 132
Ativos Relacionados. 133
Ativos relacionados a Cada Tipo de Ativo. 133
Exibindo Ativos Relacionados. 134
Perguntas Frequentes. 134

Capítulo 16: Business Glossary Desktop no Informatica Analyst. 136

Termos Comerciais. 136
Gerenciar Termos Comerciais no Glossário Comercial do Metadata Manager. 137
Pesquisar um Termo Comercial no Glossário Comercial Desktop. 137

Parte III: Descoberta de Dados com o Informatica Developer. 138

Capítulo 17: Perfis do Informatica Developer. 139

Perfis do Informatica Developer - Visão Geral. 139
Exibições de Perfil do Informatica DeveloperExibições de Perfil. 141
Bloqueios de Objetos de Repositório e Desenvolvimento Baseado em Equipe com Objetos com Controle de Versão. 142

Capítulo 18: Perfis de Objeto de Dados. 143

Visão Geral dos Perfis de Objeto de Dados. 143
Perfis de Coluna no Informatica Developer. 144
Opções de Filtragem. 145
Opções de Amostragem. 146
Ambiente de tempo de execução. 146
Ambiente Nativo. 146
Ambiente Hadoop. 147
Descoberta de Chave Primária. 148
Propriedades de Inferência de Chave Primária. 148
Propriedades de Chaves Primárias Inferidas. 149
Propriedades de Violações de Chave. 149
Descoberta de Dependência Funcional. 149
Propriedades de Inferência de Dependência Funcional. 150
Propriedades de Dependências Funcionais Inferidas. 150
Propriedades de Violações de Dependência Funcional. 151
Perfis de sistema operacional no Informatica Developer. 151
Selecionando um perfil do sistema operacional. 151
Criando um perfil de objeto de dados único no Developer Informatica. 152
Criando Vários Perfis de Objetos de Dados no Informatica Developer. 153
Editando um Perfil. 153

Sincronizando um Objeto de Dados de Arquivo Simples. 153
Sincronizar um Objeto de Dados Relacionais. 154
Comentários. 154
Adicionando comentários no Informatica Developer. 154

Capítulo 19: Perfis de coluna em fontes de dados semiestruturadas. 156

Visão geral de perfis de coluna em fontes de dados semiestruturadas. 156
Perfil de coluna em uma fonte de dados JSON ou XML. 157
Criando um objeto de dados em uma fonte de dados JSON ou XML. 157
Perfil de coluna em uma fonte de dados semiestruturada no HDFS. 158
Perfil de coluna em uma fonte de dados XML ou JSON no HDFS. 158
Perfil de coluna em uma fonte de dados Avro ou Parquet no HDFS. 158
Criando uma conexão HDFS. 158
Criando um objeto de dados de arquivo complexo em um arquivo JSON ou XML no HDFS. 159
Criando um objeto de dados em uma fonte de dados Avro ou Parquet. 160
Executando um perfil de coluna em fontes de dados semiestruturadas. 161

Capítulo 20: Coluna de Resultados de Perfil no Informatica Developer. 163

Resultados do Perfil de Colunas no Informatica DeveloperResultados do Perfil de Colunas. 163
Propriedades de Valor de Coluna. 164
Propriedades de Padrão de Coluna. 165
Propriedades de Estatísticas de Coluna. 165
Propriedades de tipo de dados da coluna. 166
Curadoria no Informatica DeveloperCuradoria no Informatica Developer. 167
Aprovando tipos de dadosAprovando tipos de dados no Informatica Developer. 167
Rejeitando tipos de dadosRejeitando tipos de dados no Informatica Developer. 167
Exportar Resultados de Perfil do Informatica Developer. 168

Capítulo 21: Regras no Informatica Developer. 169

Visão Geral de Regras no Informatica DeveloperDiretrizes para Regras. 169
Criar uma Regra no Informatica Developer. 170
Aplicando uma Regra no Informatica DeveloperAplicando uma Regra. 170

Capítulo 22: Scorecards no Informatica Developer. 171

Visão Geral dos Scorecards no Informatica Developer. 171
Criar um Scorecard. 171
Exportar um Arquivo de Recursos para a Linhagem de Scorecard. 172
Exibir a Linhagem de Scorecard do Informatica Developer. 172

Capítulo 23: Criação de Perfil de Mapplet e Mapeamento. 174

Visão Geral da Criação de Perfil de Mapplet e MapeamentoPerfis de Mapplet e Mapeamento. . . . 174
Executar um Perfil em um Mapplet ou Objeto de Mapeamento. 174
Comparar Perfis de Objetos de Mapplet ou de Mapeamento. 175

Gerar um Mapeamento de Perfil. 175

Capítulo 24: Descoberta do Domínio de Dados no Informatica Developer. . . 177

Descoberta do Domínio de Dados no Informatica Developer - Visão Geral. 177
Glossário do Domínio de Dados no Informatica Developer. 178
Criando um Grupo do Domínio de Dados no Informatica Developer. 178
Criando um Domínio de Dados no Informatica Developer. 179
Criando um Domínio de Dados a Partir de Resultados do Perfil no Informatica Developer. . . . 179
Localizar Domínios de Dados no Informatica Developer. 180
Importando Domínios de Dados. 180
Exportando Domínios de Dados. 181
Opções de Descoberta do Domínio de Dados no Informatica Developer. 182
Seleção de Domínio de Dados no Informatica Developer. 182
Seleção de Colunas de Domínio de Dados no Informatica Developer. 183
Opções de Inferência de Domínio de Dados no Informatica Developer. 183
Criando um Perfil para Executar Descoberta do Domínio de Dados no Informatica Developer. . . . 185
Editando um Perfil no Informatica Developer. 185
Executando um Perfil para Realizar a Descoberta do Domínio de Dados no Informatica Developer 186
Resultados da Descoberta do Domínio de Dados no Informatica Developer. 186
Exibindo por grupos de domínios de dados. 187
Exibindo por colunas. 187
Verificando os resultados. 188
Aprovando Domínios de Dados. 188
Rejeitando Domínios de Dados. 189
Exportando Resultados da Descoberta do Domínio de Dados do Informatica Developer. 189

Capítulo 25: Descoberta Empresarial no Informatica Developer. 190

Descoberta Empresarial no Informatica Developer - Visão Geral. 190
Processo de Descoberta Empresarial. 191
Opções de Perfil para Descoberta Empresarial. 192
Seleção de Domínio de Dados para Descoberta Empresarial. 192
Opções de amostragem de perfil de coluna para descoberta empresarial. 193
Opções de ambientes de tempo de execução. 193
Opções de Inferência de Chave Primária para Descoberta Empresarial. 194
Opções de Inferência de Chave Estrangeira para Descoberta Empresarial. 194
Parâmetros de curadoria automáticos para inferência de chaves externas. 195
Criando um Perfil de Descoberta Empresarial no Informatica Developer. 197
Editando um Perfil. 198
Executando um Perfil de Descoberta Empresarial. 199
Descoberta de Chave Externa. 200
Definindo Relacionamentos de Objeto Pai e Filho. 200
Descobrindo Relacionamentos de Chave Externa entre Objetos de Dados. 200
Resultados de Análise de Chave Externa. 201

Análise de Associação. 201
Criando um Perfil de Associação. 202
Resultados da Análise de Associação. 202
Exportando Resultados do Perfil de Associação para um Arquivo. 203
Descoberta de Sobreposição. 203
Resultados da descoberta de sobreposição. 203
Descobrindo Dados de Sobreposição. 204
Arquivos de Script DDL. 205
Criando Scripts DDL de um Perfil de Descoberta Empresarial. 205

Capítulo 26: Resultados da Descoberta Empresarial. 206

Resultados da Descoberta Empresarial - Visão Geral. 206
Exibição Relacionamentos. 207
Pesquisando um Objeto de Dados. 207
Navegando até a Exibição Criação de Perfil de Chave Externa. 208
Exibição Criação de Perfil de chave Externa. 208
Exibindo Relacionamentos de Objeto de Dados. 208
Ampliando e Reduzindo o Zoom da Exibição. 209
Localizando um Objeto de Dados. 209
Exibindo Relacionamentos de Colunas. 210
Salvando o Diagrama de Relacionamento de Entidades como uma Imagem. 210
Exibindo Resultados do Perfil de Objeto de Dados na Exibição Criação de Perfil de Chave Externa. 210
Exibição Tabular. 211
Painel Detalhes da Tabela. 211
Verificando os Resultados da Descoberta Empresarial. 211
Fazendo a curadoria de relacionamentos de coluna. 212
Confirmando os Resultados no Repositório do Modelo. 212
Exibição Domínios de Dados. 213
Exibindo Resultados da Descoberta do Domínio de Dados. 213
Verificando Resultados da Descoberta do Domínio de Dados. 213
Fazendo Buscas Detalhadas em Linhas. 213
Exibindo Resultados do Perfil de Objeto de Dados na Exibição Domínios de Dados. 214
Exibição Perfil de Coluna. 214
Exibindo Resultados do Perfil de Objeto de Dados. 214
Exibindo Resultados do Perfil de Coluna Durante a Execução da Descoberta Empresarial. 214
Exibindo Resultados da Descoberta do Domínio de Dados Durante a Execução da Descoberta Empresarial. 215
Exibindo o Status da Descoberta Empresarial em Tempo de Execução. 215
Arquivos de Exportação de Descoberta Empresarial. 216
Exportando Resultados da Descoberta Empresarial. 216

Capítulo 27: Glossário Comercial Desktop no Informatica Developer. 217

Pesquisa no Business Glossary. 217

Pesquisar um Termo Comercial. 217

Personalizando Teclas de Acesso para Pesquisar um Termo Comercial. 218

Índice. 219

Prefácio

O *Guia do Informatica Data Explorer* foi concebido para usuários do Informatica Analyst e do Informatica Developer. Ele contém informações sobre como você pode usar perfis para analisar o conteúdo, a qualidade e a estrutura de fontes de dados e realizar a descoberta de dados para descobrir os metadados de sistemas de origem que incluem conteúdo e estrutura.

Use perfis para descobrir problemas de qualidade em uma fonte de dados e para compreender os relacionamentos entre colunas em uma ou mais fontes 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.

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 à Descoberta de Dados

Esta parte contém os seguintes capítulos:

- Introdução à Criação de Perfil, 17
- Descoberta de Dados, 22
- Conceitos de Perfil de Coluna, 26
- Conceitos de Descoberta do Domínio de Dados, 29
- Conceitos de Curadoria, 32

CAPÍTULO 1

Introdução à Criação de Perfil

Este capítulo inclui os seguintes tópicos:

- Visão Geral da Criação de Perfil, 17
- Arquitetura de Criação de Perfil, 18
- Processo de Descoberta de Dados, 20

Visão Geral da Criação de Perfil

Use a criação de 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 tanto na Developer tool quanto 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.
- Execute a descoberta do domínio de dados. Você pode descobrir características de dados críticas dentro de uma empresa.
- Faça curadoria dos resultados do perfil incluindo tipos de dados, domínios de dados, chaves primárias e chaves externas.
- Crie scorecards para monitorar a qualidade dos dados.
- Escolha um perfil do sistema operacional para criar e executar perfis de coluna, perfis de descoberta empresarial e scorecards com base nas permissões do usuário do sistema operacional que você define nesse perfil do sistema operacional.
- Use bloqueios de ativos de repositório para impedir que outros usuários sobrescrevam trabalhos.
- Use o sistema de controle de versão para salvar várias versões de um perfil.

- Crie e atribua marcas aos objetos de dados.
- Pesquise o significado do nome de um objeto como um termo comercial no Business Glossary Desktop. Por exemplo, você pode pesquisar o significado de um nome de coluna ou nome de perfil para entender seu requisito de negócios e a implementação atual.

Você pode executar as seguintes tarefas na Developer tool:

- Descobrir o grau de possíveis associações entre duas colunas de dados de uma fonte de dados.
- Determinar a porcentagem de dados de sobreposição em pares de colunas dentro de uma fonte de dados ou várias fontes de dados.
- Comparar os resultados da criação de perfil de coluna.
- Gerar um objeto de mapeamento a partir de um perfil.
- Descobrir chaves primárias em uma fonte de dados.
- Descobrir chaves externas em um conjunto de uma ou mais fontes de dados.
- Descobrir dependência funcional entre colunas de uma fonte de dados.
- Executar tarefas de descoberta de dados em um grande número de fontes de dados em várias conexões. As tarefas de descoberta de dados incluem perfil de coluna, inferência de relacionamentos de chave primária e chave externa, descoberta do domínio de dados e geração de um resumo gráfico consolidado dos relacionamentos de dados.

Você pode executar as seguintes tarefas na ferramenta Analyst:

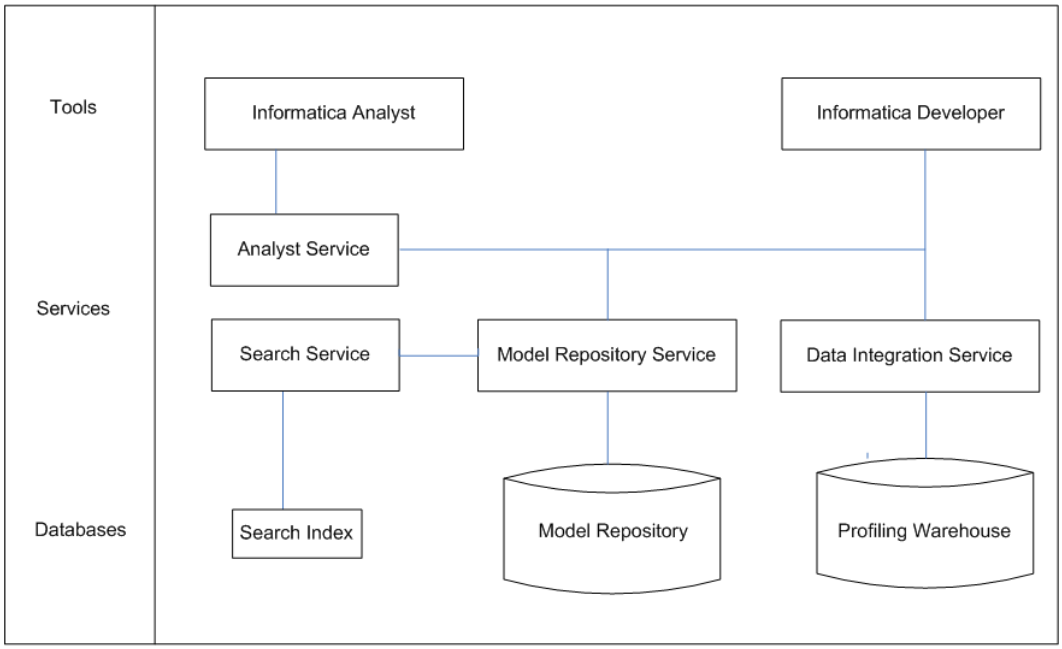
- Executar a descoberta empresarial em um grande número de fontes de dados em várias conexões. Você pode exibir um resumo consolidado dos resultados da descoberta de metadados da coluna e domínios de dados.
- Executar a pesquisa de descoberta para encontrar onde os dados e metadados existem na empresa. Você pode procurar ativos específicos, como objetos de dados, regras e perfis. A pesquisa de descoberta localiza ativos e identifica relacionamentos com outros ativos nos bancos de dados e esquemas da empresa.
- Exiba os resultados do perfil de uma execução de perfil histórico.
- Compare os resultados de perfil para duas execuções de perfil em um perfil de coluna.
- Exiba a linhagem para cada métrica e grupo de métricas do scorecard.
- Exiba o painel de scorecards.
- Adicione comentários a um perfil ou a colunas em um perfil.
- Atribua marcas a um perfil ou a colunas em um perfil.
- Atribua termos comerciais a colunas em um perfil.

Arquitetura de Criação de Perfil

A arquitetura de criação de perfil consiste em ferramentas, serviços e bancos de dados. O componente de ferramentas consiste em aplicativos cliente. O componente de serviços tem serviços de aplicativo necessários para gerenciar as ferramentas, executar as tarefas de integração de dados e gerenciar os

metadados dos objetos de perfil. O componente de bancos de dados consiste no repositório do Modelo e no depósito de criação de perfil.

A imagem a seguir mostra os componentes de arquitetura para criação de perfil:



Quando você executa um perfil, o Serviço Analyst ou a Developer tool recebe a definição de perfil do Serviço de Repositório do Modelo. Em seguida, o Serviço Analyst ou a Developer tool invoca o plug-in de criação de perfil no Serviço de Integração de Dados. Em seguida, o plug-in de criação de perfil processa o trabalho de perfil e envia o trabalho para o Serviço de Integração de Dados. O Serviço de Integração de Dados gera os resultados de criação de perfil. Em seguida, o Serviço de Integração de Dados grava os resultados da criação de perfil no depósito de criação de perfil.

A pesquisa de descoberta usa o Serviço de Pesquisa. O Serviço de Pesquisa realiza cada pesquisa em um índice de pesquisa, em vez de no repositório do Modelo ou no depósito de criação de perfil. O Serviço de Pesquisa gera o índice de pesquisa com base no conteúdo do repositório do Modelo e do depósito de criação de perfil. O Serviço de Pesquisa contém extratores para extrair o conteúdo de cada repositório.

A seguinte tabela descreve os componentes da arquitetura:

Componente	Descrição
Informatica Analyst	Um aplicativo cliente baseado na Web que você pode usar para descobrir, analisar e relatar sobre dados e metadados de fontes de dados.
Informatica Developer	Um aplicativo cliente usado para executar a descoberta de dados avançada, como a descoberta de chaves primárias, a descoberta de chaves externas e a descoberta empresarial.
Serviço Analyst	Um serviço de aplicativo que executa a ferramenta Analyst e gerencia conexões entre os componentes do serviço e os usuários da ferramenta Analyst.
Serviço de Pesquisa	Um serviço de aplicativo que gerencia pesquisas na ferramenta Analyst. Por padrão, o Serviço de Pesquisa retorna resultados de pesquisa do repositório do Modelo, como objetos de dados, perfis, especificações de mapeamento, tabelas de referência, regras e scorecards.

Componente	Descrição
Índice de Pesquisa	Um sistema de arquivos em um diretório personalizado que armazena conteúdo indexado que o Serviço de Pesquisa extrai do repositório do Modelo e do depósito de criação de perfil.
Serviço de Repositório do Modelo	Um serviço de aplicativo que gerencia o repositório do Modelo.
Serviço de Integração de Dados	Um serviço de aplicativo que executa tarefas de integração de dados para a ferramenta Analyst, a Developer tool e clientes externos.
Repositório do Modelo	Um banco de dados relacional que armazena os metadados para projetos criados na ferramenta Analyst ou na Developer tool.
Depósito de criação de perfil	Um banco de dados que armazena informações de criação de perfil, como resultados de perfil e de scorecard.

Processo de Descoberta de Dados

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, qualidade e 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, uma descoberta de chave primária, uma descoberta de chave externa e uma descoberta do domínio de dados. Você revela e documenta problemas de qualidade de dados. Conclua as seguintes tarefas para realizar a descoberta de dados:

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. Descubra a estrutura dos dados. Inclui chaves, dependências funcionais e chaves externas.
3. Analise e valide os resultados de perfil.
4. Faça uma busca detalhada nos resultados do perfil.
5. Faça curadoria dos resultados de perfil.
6. Crie dados de referência.
7. Documente problemas de dados.
8. Crie e execute regras.
9. Crie scorecards para monitorar a qualidade dos dados.

É possível usar as seguintes ferramentas para gerenciar o processo de descoberta:

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 nesta ferramenta para localizar e analisar os metadados de um ou mais fontes de dados, descobrindo inclusive os relacionamentos entre colunas. Você cria perfis usando um assistente.

Informatica Analyst

Você pode executar um perfil de coluna, executar a descoberta de domínio de dados e executar a descoberta empresarial em objetos de dados na ferramenta Analyst. Depois de executar um perfil, você pode fazer uma busca detalhada nas linhas de dados em uma fonte de dados.

CAPÍTULO 2

Descoberta de Dados

Este capítulo inclui os seguintes tópicos:

- Visão Geral da Descoberta de Dados, 22
- Tipos de perfil e análise, 22
- Componentes de Criação de Perfil, 23
- Resultados do Perfil, 24

Visão Geral da Descoberta de Dados

Descoberta de dados é o processo de descobrir os metadados de sistemas de origem que incluem conteúdo e estrutura. O conteúdo se refere a valores de dados, frequências e tipos de dados. A estrutura inclui chaves candidatas, chaves primárias, chaves externas e dependências funcionais. Você pode criar e executar perfis para descobrir o conteúdo e estrutura das fontes de dados.

É possível definir um perfil para analisar dados em um único objeto de dados ou em vários. Adicione comentários aos perfis para poder controlar o processo de criação de perfil de forma eficaz.

Execute um perfil para avaliar a estrutura de dados e verificar se as colunas de dados contêm os tipos de informações que você espera. Você pode fazer uma busca detalhada nas linhas dos dados cujo perfil você criou. Se os resultados do perfil revelarem problemas nos dados, você poderá aplicar regras para corrigir o conjunto de resultados. É possível criar scorecards para controlar e medir a qualidade dos dados antes e depois de aplicar as regras. Se os metadados de origem externa de um perfil ou scorecard mudarem, você poderá sincronizar as alterações com seu objeto de dados.

Tipos de perfil e análise

Crie um perfil com base no tipo de análise que você precisa executar. O tipo de perfil que você criar corresponderá ao tipo de análise que executar. Por exemplo, para executar uma análise de chave primária, você cria um perfil de chave primária.

Você pode criar os seguintes perfis para executar análise e descoberta de dados:

Perfil de Coluna

Analisa a qualidade dos dados em colunas selecionadas em uma tabela ou arquivo. Você pode definir perfis para análise de coluna na ferramenta Analyst e na Developer tool.

Descoberta do Domínio de Dados

Descobre características de dados críticas em uma empresa. A descoberta do domínio de dados identifica todos os domínios de dados associados a uma coluna com base no valor ou no nome da coluna. Como parte do processo de descoberta, você pode criar manualmente as regras de dados e regras de nome de coluna para verificar se um valor ou nome de coluna pertence a um domínio de dados. Você pode então associar essas regras ao criar um domínio de dados. Você também pode criar domínios de dados a partir dos valores e padrões dos resultados do perfil de coluna.

Perfil de Chave Primária

Descobre relacionamentos de chave primária entre colunas em uma tabela ou arquivo. Você pode definir perfis para análise de chave primária na Developer tool.

Perfil de Dependência Funcional

Descobre dependências funcionais entre colunas em uma tabela ou arquivo. Você pode definir perfis para análise de dependência funcional na Developer tool.

Perfil de Chave Externa

Descobre relacionamentos de chave externa entre colunas em várias tabelas ou vários arquivos. Você pode definir perfis para análise de chave externa na Developer tool.

Perfil de Associação

Descobre o grau de possíveis associações entre colunas de uma fonte de dados ou entre várias fontes de dados. Você pode definir perfis para análise de associação na Developer tool. Os resultados são exibidos em um diagrama de Venn.

Descoberta de Sobreposição

Determina a porcentagem de dados de sobreposição em pares de colunas dentro de uma fonte de dados ou várias fontes de dados. Você pode executar a tarefa de descoberta de sobreposição no editor na Developer tool. Você pode validar os resultados e exibi-los em um diagrama de Venn.

Descoberta Empresarial

Executa várias tarefas de descoberta de dados em um grande número de fontes de dados e gera um resumo consolidado dos resultados do perfil. Inclui a execução de um perfil de coluna, descoberta do domínio de dados e descoberta de relacionamentos de chave primária e chave externa. A descoberta empresarial automatiza o processo de perfil para um grande número de fontes de dados.

Nota: As alterações feitas em perfis na ferramenta Analyst não aparecerão na Developer tool até que você atualize a conexão da Developer tool com o repositório do modelo.

Componentes de Criação 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 objetos de grupo de acordo com seu uso comercial. Atribua marcas a um perfil ou a colunas em um perfil na ferramenta Analyst.

Comentário

Descrição sobre o perfil. Use comentários para compartilhar informações sobre perfis com outros usuários da ferramenta Analyst e da Developer tool. Adicione comentários a um perfil ou a colunas em um perfil na ferramenta Analyst.

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 de dados.

Resultados do Perfil

Você pode exibir os resultados após executar um perfil. É possível exibir um resumo de valores, padrões e estatísticas de colunas e regras no perfil. Você pode exibir as propriedades das colunas e regras do perfil. É possível visualizar os dados do perfil.

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

Tipo de perfil	Resultados
Perfil de coluna	- Número e porcentagem de valores nulos, distintos e não distintos nas colunas e os tipos de dados inferidos para valores de colunas. - Os padrões de frequência e caractere dos valores de dados em uma coluna selecionada e um resumo estatístico para a coluna. - Tipos de dados inferidos ao analisar dados de colunas. - Tipo de dados documentado para os dados. - Valores máximo e mínimo. - Data e hora da execução de perfil.
Perfil de chave primária	- O número e a porcentagem de valores exclusivos, duplicados e nulos para os candidatos a chave primária inferidos. - Número de violações de chave nos candidatos a chave primária inferidos.
Perfil de dependência funcional	- Dependências funcionais inferidas. - Número de violações de dependência funcional.
Perfil de chave externa	- As colunas de chave primária e externa que atendem aos critérios de inferência de chave primária-externa definidos por você. - O número de valores de dados com correspondência entre as chaves primária e externa, expresso como uma porcentagem. - O tipo de relacionamento definido para as colunas de chave primária e externa antes de o perfil ser executado.
Perfil de associação	- Diagrama de Venn que mostra os relacionamentos entre colunas. - O número e a porcentagem de valores órfãos, nulos e associados nas colunas.

Tipo de perfil	Resultados
Descoberta de sobreposição	- Porcentagem de sobreposição entre duas colunas. - Diagrama de Venn que mostra a sobreposição entre colunas.
Descoberta do Domínio de Dados	- Nome da coluna e dados que correspondem a domínios de dados predefinidos. - Grupo de domínio de dados ao qual a coluna pertence e seu tipo de dados.
Descoberta empresarial	- Resultados do perfil de coluna. - Resultados da descoberta do domínio de dados. - Resultados da descoberta de chave primária. - Resultados do perfil de chave externa em exibições gráfica e tabular.

Você pode usar ferramentas de relatório de terceiros para ler os resultados do perfil no warehouse de criação de perfil. A Informatica fornece um conjunto de exibições de perfil que pode ser personalizado para as estatísticas de perfil que você deseja ler. Essas exibições são baseadas em tipos comuns de estatísticas de perfil e análise de resultados do perfil.

CAPÍTULO 3

Conceitos de Perfil de Coluna

Este capítulo inclui os seguintes tópicos:

- Visão Geral de Conceitos de Perfil de ColunaVisão Geral de Perfis, 26
- Opções de Perfil de Coluna, 27
- Gerenciamento de Perfis com Controle de Versão e Bloqueios de Perfis de Repositório, 27
- Scorecards, 28

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 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 4

Conceitos de Descoberta do Domínio de Dados

Este capítulo inclui os seguintes tópicos:

- Conceitos de Descoberta do Domínio de Dados - Visão Geral, 29
- Domínios de Dados, 30
- Grupos de Domínios de Dados, 30
- Glossário do Domínio de Dados, 31
- Processo de Descoberta do Domínio de Dados, 31

Conceitos de Descoberta do Domínio de Dados - Visão Geral

Você precisa identificar e entender o significado dos dados de origem críticos para poder adotar medidas para trabalhar neles com eficácia. Descoberta do domínio de dados é o processo de descobrir o significado funcional dos dados nas fontes com base na semântica dos dados.

Crie um perfil para realizar a descoberta do domínio de dados e poder identificar características de dados críticas dentro da empresa. Você poderá então aplicar mais políticas de gerenciamento de dados, como qualidade ou mascaramento, aos dados. Por exemplo, descubra códigos ou descrições de produto para avaliar quais regras de análise ou padronização da qualidade dos dados é necessário aplicar para tornar os dados úteis e dignos de confiança. Outro exemplo consiste em localizar dados confidenciais de clientes, como números de cartão de crédito, IDs de e-mail e números de telefone. Você pode então querer mascarar essas informações para protegê-las.

Você pode criar e executar um perfil para realizar a descoberta do domínio de dados nas ferramentas Analyst e Developer. É possível definir um perfil para realizar a descoberta do domínio de dados com base nas seguintes regras:

- Regra de dados. Localiza colunas com dados que correspondem à lógica específica definida na regra.
- Regra de nome da coluna. Localiza colunas que correspondem à lógica de nome da coluna definida na regra.

Você pode criar domínios de dados a partir dos valores e padrões dos resultados do perfil de coluna. Você pode então usar esses domínios de dados para descobrir dados críticos em vários sistemas de dados ou empresas.

Você pode criar um perfil com uma opção de amostragem e filtros para realizar a descoberta do domínio de dados. Ao executar o perfil, você aplica a opção de amostragem e os filtros à fonte de dados e gera um conjunto de dados. O processo de descoberta do domínio de dados usa o conjunto de dados para descobrir domínios de dados.

Domínios de Dados

Um domínio de dados é um objeto de repositório do Modelo predefinido ou definido pelo usuário com base na semântica de dados da coluna ou em um nome de coluna. Por exemplo, número de previdência social, número de cartão de crédito, ID de e-mail e número de telefone podem ser domínios de dados individuais.

Um domínio de dados ajuda você a localizar dados importantes que permanecem não descobertos em uma fonte de dados. Por exemplo, você pode ter sistemas de dados herdados que contenham números de previdência social em um campo Comentários. É necessário localizar essas informações e protegê-las antes migrá-las para novos sistemas de dados.

Você pode escolher uma porcentagem mínima de linhas de origem ou o número mínimo de linhas de origem como critérios de conformidade para a correspondência de domínio de dados. Também pode excluir valores nulos ao realizar a descoberta do domínio de dados em um perfil de coluna.

Você pode agrupar domínios de dados lógicos em grupos. Um glossário do domínio de dados lista todos os domínios de dados e grupos de domínios de dados. Use o menu Preferências na ferramenta Developer para importar e exportar domínios de dados de e para o glossário do domínio de dados.

Você usa regras para definir padrões de dados e nomes de coluna que correspondam aos dados e metadados de origem. Quando você cria um domínio de dados, a ferramenta Analyst ou a ferramenta Developer copia regras associadas e outros objetos dependentes para o glossário do domínio de dados. Use a ferramenta Developer para gerenciar domínios de dados que incluem a importação e exportação de domínios de dados de e para o glossário do domínio de dados. Você também pode usar a ferramenta Developer para gerenciar a lógica da regra dos domínios de dados.

Nota: Você pode querer salvar todas as regras de domínio de dados em um único projeto ou pasta. Esta etapa ajuda depois que você exporta domínios de dados e tem uma necessidade de editar as regras e outros objetos de dados associados.

Grupos de Domínios de Dados

Os grupos de domínios de dados ajudam você a categorizar domínios de dados em grupos específicos. Por exemplo, você pode agrupar os domínios de dados first_name, last_name e account_number sob o grupo do domínio de dados Informações Pessoais de Saúde (PHI).

Você pode criar um grupo do domínio de dados Informações de Identificação Pessoal (PII) que inclua o número de previdência social, nome e sobrenome. Um domínio de dados pode ser uma parte de vários grupos de domínios de dados. Por exemplo, o número de previdência social pode pertencer aos grupos de domínios de dados PCI (Payment Card Industry) e PII. Os grupos de domínios de dados podem conter domínios de dados e não outros grupos de domínios de dados.

Nota: Se você importar o arquivo de domínio de dados `Informatica_IDE_DataDomain.xml` após a instalação, o glossário de domínios de dados exibirá domínios de dados e grupos de domínios de dados predefinidos. Você poderá, então, criar mais grupos de domínios de dados conforme necessário. Para exibir e alterar as regras associadas aos domínios de dados, importe o arquivo `Informatica_IDE_DataDomainRule.xml`.

Glossário do Domínio de Dados

O glossário do domínio de dados é um contêiner de todos os domínios de dados e grupos de domínios de dados. Você pode usar o glossário do domínio de dados para criar, gerenciar e remover domínios de dados e grupos de domínios de dados.

Você pode pesquisar domínios e grupos de domínios específicos no glossário do domínio de dados. Pode também exportar domínios de dados de um arquivo XML e importar domínios de dados de um arquivo XML para o glossário do domínio de dados. O glossário do domínio de dados contém regras copiadas e todos os dados de referência associados a domínios de dados. Não é possível editar as regras no glossário do domínio de dados.

Você pode exibir o glossário do domínio de dados no menu Preferências da ferramenta Developer e no menu Gerenciar da ferramenta Analyst. Use o privilégio **Gerenciar Domínios de Dados** do Model Repository Service para determinar quem cria, edita e exclui domínios de dados e grupos de domínios de dados.

Processo de Descoberta do Domínio de Dados

Você pode definir e executar um perfil para realizar a descoberta do domínio de dados na ferramenta Analyst ou na ferramenta Developer com base na função do seu cargo. Após configurar as opções de descoberta do domínio de dados e executar o perfil, você pode verificar e fazer uma busca detalhada nos resultados. Se você executar a descoberta do domínio de dados no editor, poderá adicionar os resultados a um modelo de dados.

Conclua as etapas a seguir para realizar a descoberta do domínio de dados:

1. Crie ou importe domínios de dados e grupos de domínios.
2. Opcionalmente, consolide domínios de dados em grupos de domínios apropriados.
3. Crie um perfil para realizar a descoberta do domínio de dados. Você começa escolhendo se deseja executar um perfil de coluna junto com a descoberta do domínio de dados ou apenas esta última.
4. Selecione as colunas, domínios e as opções de amostragem apropriadas.
5. Execute o perfil.
6. Verifique, faça uma busca detalhada nos resultados do perfil e adicione os resultados a um modelo de perfil, conforme necessário.

CAPÍTULO 5

Conceitos de Curadoria

Este capítulo inclui os seguintes tópicos:

- Visão Geral dos Conceitos de Curadoria, 32
- Curadoria para Analistas e Desenvolvedores, 32
- Tarefas de Curadoria, 33

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.

Curadoria para Analistas e Desenvolvedores

Como um analista de dados ou administrador de dados, você pode fazer curadoria dos resultados do perfil da coluna e dos resultados da descoberta do domínio de dados na ferramenta Analyst. Você pode fazer

curadoria dos resultados do perfil para deixar as informações de perfil precisas prontas para pesquisa de descoberta e para validação posterior dos ativos de dados.

Como um desenvolvedor ou arquiteto de dados, você pode fazer curadoria dos resultados de perfil de coluna, resultados da descoberta do domínio de dados, resultados da descoberta de chave primária e resultados da descoberta de chave externa na Developer tool.

Exemplos de Curadoria

Quando você executa a descoberta empresarial como um desenvolvedor, a Developer tool processa os domínios de dados selecionados de todo o conjunto de dados. Essa ação pode resultar em várias inferências de domínios de dados, como os dados de número de telefone inferidos como o domínio de dados do número de Previdência Social. As várias inferências de domínio de dados ocorrem quando partes dos dados em uma coluna correspondem a diferentes domínios de dados. Por exemplo, um número de telefone de 10 dígitos que está sem um dígito pode ter o mesmo padrão que um número de Previdência Social. Essa ocorrência indica possíveis problemas de qualidade de dados em uma coluna ou um padrão de correspondência em vários domínios de dados. Nesse caso, a Developer tool pode inferir o domínio de dados do número de telefone e o domínio de dados do número de Previdência Social. Você pode fazer curadoria dos resultados do perfil de forma que possa selecionar o domínio de dados mais adequado e aprová-lo. No exemplo, o número de telefone é o domínio de dados relevante porque a inferência do domínio de dados do número de Previdência Social ocorre devido a um problema de qualidade de dados.

Quando você executa a descoberta empresarial, a Developer tool pode inferir vários tipos de dados, como dados, cadeia e varchar, para uma coluna de data. Como um arquiteto de dados, talvez você deseje selecionar e aprovar o tipo de dados de data, que é o tipo de dados mais relevante de uma coluna de dados.

A descoberta empresarial na Developer tool pode inferir todos os relacionamentos de objeto de dados com base nos dados da coluna. Alguns desses relacionamentos de objeto de dados incluem relacionamentos indesejados de objeto de dados nas chaves candidatas descobertas. Por exemplo, a Developer tool pode inferir colunas que representam uma sequência de chaves e descobrir relacionamentos com outras tabelas que têm colunas semelhantes. Esses relacionamentos de objeto de dados podem não formar relacionamentos válidos no banco de dados. Nesses casos, você pode avaliar, verificar e aprovar os resultados de perfil inferidos mais adequados como parte da curadoria.

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: Descoberta de Dados com o Informatica Analyst

Esta parte contém os seguintes capítulos:

- Perfis da Coluna no Informatica Analyst, 36
- Regras no Informatica Analyst, 46
- Filtros no Informatica Analyst, 50
- Resultados do Perfil da Coluna no Informatica Analyst, 55
- Termos comerciais, comentários e marcas no Informatica Analyst, 78
- Scorecards no Informatica Analyst, 81
- Descoberta do Domínio de Dados no Informatica Analyst, 104
- Descoberta Empresarial no Informatica Analyst, 115
- Resultados da Descoberta Empresarial no Informatica Analyst, 121
- Pesquisa de Descoberta no Informatica Analyst, 126
- Business Glossary Desktop no Informatica Analyst, 136

CAPÍTULO 6

Perfis da Coluna no Informatica Analyst

Este capítulo inclui os seguintes tópicos:

- Visão Geral dos Perfis de Coluna no Informatica Analyst, 36
- Processo de Criação de Perfil de Coluna, 37
- Opções de Perfil, 37
- Ambiente de tempo de execução, 38
- Visão geral de perfis de sistema operacional no Informatica Analyst, 40
- Visão Geral de Bloqueios de Ativos de Repositório e do Desenvolvimento Baseado em Equipe, 40
- Criando um perfil de coluna no Informatica Analyst, 41
- Editar um Perfil de Coluna, 42
- Executar um Perfil, 43
- Sincronizando um Objeto de Dados de Arquivo Simples, 44
- Sincronizando um Objeto de Dados Relacional, 45

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.

Ambiente de tempo de execução

Você pode escolher 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.

Ambiente Nativo

Quando você executa um perfil no ambiente de tempo de execução 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 na mesma máquina em que ele é executado e grava os resultados do perfil no depósito de criação de perfil. Por padrão, todos os perfis são executados no ambiente de tempo de execução nativo.

Você pode usar origens nativas para criar e executar perfis no ambiente nativo. Uma fonte de dados nativa é uma origem não Hadoop, como um arquivo simples, uma origem relacional ou uma origem de mainframe. Você também pode executar um perfil em uma especificação de mapeamento ou em uma fonte de dados lógica com uma fonte de dados Hive ou HDFS no ambiente nativo.

Ambiente Hadoop

Depois que você escolher a opção Hive e selecionar uma conexão do Hadoop, o Serviço de Integração de Dados enviará a lógica do perfil ao mecanismo Hive no cluster Hadoop para executar perfis. Depois que você escolher a opção Hadoop e selecionar uma conexão do Hadoop, o Serviço de Integração de Dados enviará a lógica do perfil ao mecanismo Blaze no cluster Hadoop para executar perfis.

Quando você executa um perfil no ambiente Hadoop, 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 envia os mapeamentos ao ambiente Hadoop por meio da conexão do Hadoop. O mecanismo Hive ou Blaze processa os mapeamentos, e o Serviço de Integração de Dados grava os resultados do perfil no depósito de criação de perfil.

Perfis de coluna para fontes de dados Sqoop

Você pode executar um perfil de coluna em objetos de dados que usam o Sqoop. Você pode selecionar o ambiente de tempo de execução Hive ou Hadoop para executar os perfis de coluna.

No mecanismo Hive, para executar um perfil de coluna em um objeto de dados relacional que usa o Sqoop, você deve definir o argumento Sqoop argument **m** como 1 na conexão JDBC. Utilize a seguinte sintaxe:

```
-m 1
```

Ao executar um perfil de coluna em um objeto de dados lógicos ou um objeto de dados personalizados, você pode configurar o argumento num-mappers para obter paralelismo e otimizar o desempenho. Você também deve configurar o argumento split-by para especificar a coluna com base no qual o programa Sqoop deve dividir as unidades de trabalho.

Utilize a seguinte sintaxe:

```
--split-by <nome_coluna>
```

Se a chave primária não tiver uma distribuição uniforme dos valores entre o intervalo mínimo e máximo, você poderá configurar o argumento split-by para especificar outra coluna que tenha uma distribuição equilibrada dos dados para dividir as unidades de trabalho.

Se você não definir a coluna split-by, o programa Sqoop dividirá as unidades de trabalho com base nos seguintes critérios:

- Se o objeto de dados contiver uma única chave primária, o programa Sqoop a usará como a coluna split-by.

- Se o objeto de dados contiver uma chave primária composta, o programa Sqoop assumirá como padrão o comportamento Sqoop de lidar com chaves primárias compostas sem o argumento split-by. Consulte a documentação do Sqoop para obter mais informações.
- Se um objeto de dados contiver duas tabelas com uma coluna idêntica, você deverá definir a coluna split-by com um nome de tabela qualificado. Por exemplo, se o nome da tabela for CUSTOMER e o nome da coluna for FULL_NAME, defina a coluna split-by da seguinte maneira:
  `--split-by CUSTOMER.FULL_NAME`
- Se o objeto de dados não contiver uma chave primária, o valor dos argumentos m e mappers-num assumirá 1 como padrão.

Quando você usa o Cloudera Connector Powered by Teradata ou o Hortonworks Connector for Teradata, e a tabela Teradata não contém uma chave primária, o argumento split-by é necessário.

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 no Informatica Analyst

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 a opção Hive ou 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 7

Regras no Informatica Analyst

Este capítulo inclui os seguintes tópicos:

- Regras na Visão Geral do Informatica Analyst, 46
- Regras Predefinidas, 47
- Regras de Expressão, 48

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:

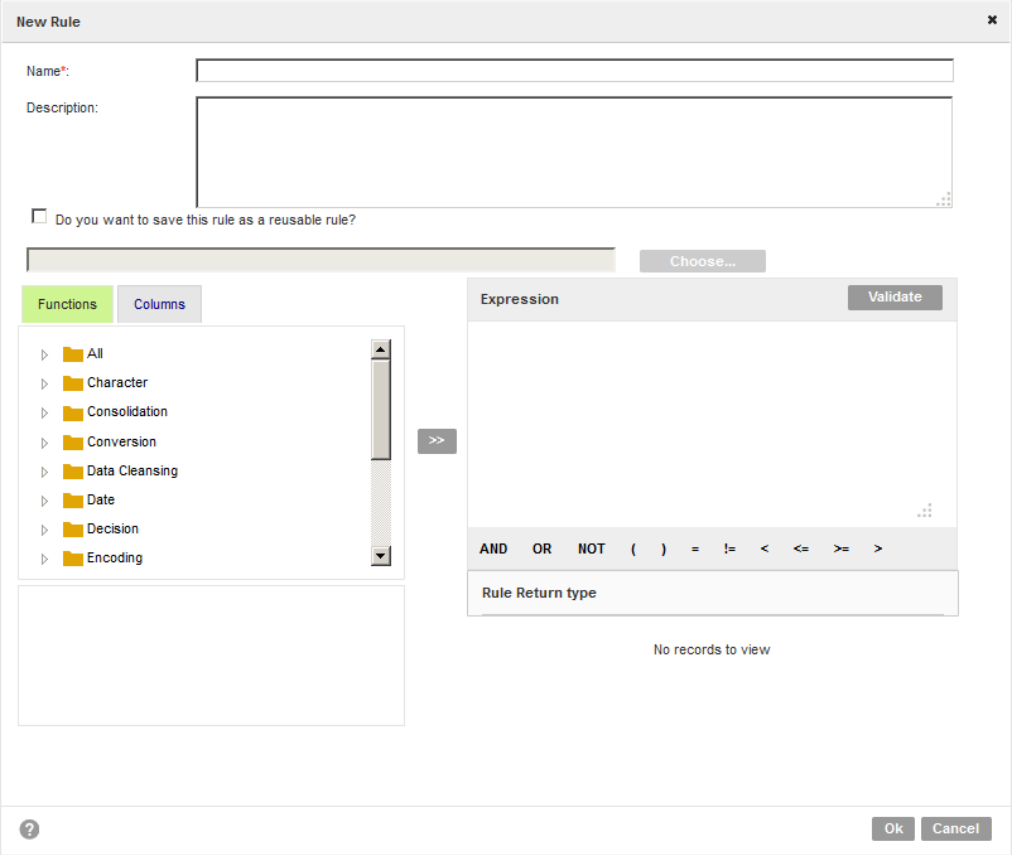
1. Abra um perfil.
2. Configure a lógica da regra usando funções de expressão e colunas como parâmetros.
3. Configure as opções de criação de perfil.

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 8

Filtros no Informatica Analyst

Este capítulo inclui os seguintes tópicos:

- Filtros na visão geral do Informatica Analyst, 50
- Criando um filtro, 50
- Gerenciando filtros, 53

Filtros na visão geral do Informatica Analyst

Você pode criar um filtro que pode ser transformado em um subconjunto da fonte de dados original que atende aos critérios de filtro. Em seguida, você pode executar um perfil nos dados filtrados.

Você pode criar um filtro para exibir os resultados do perfil que atenda aos critérios de filtro. Você pode exibir os resultados do perfil com os filtros padrão que estão disponíveis na exibição resumida.

Criando um filtro

Você pode criar um filtro que pode ser transformado em um subconjunto da fonte de dados original que atende aos critérios de filtro.

1. Abra um perfil.
2. Na exibição resumida, clique em **Ações** > **Editar Perfil**.

   O assistente de **Perfil** é exibido.
3. Clique em **Especificar Regras e Filtros**.
4. No painel **Filtros**, clique em **Ações** > **Adicionar um Filtro**.

   A caixa de diálogo **Novo Filtro** é exibida.
5. Crie um filtro SQL, simples ou avançado.

   Nota: Para um filtro simples ou avançado em uma coluna de data, especifique a condição no formato YYYY/MM/DD HH:MM:SS.

   O painel **Visualização de Dados** exibe o subconjunto da fonte de dados original que atende os critérios de filtro.
6. Clique em **OK**.

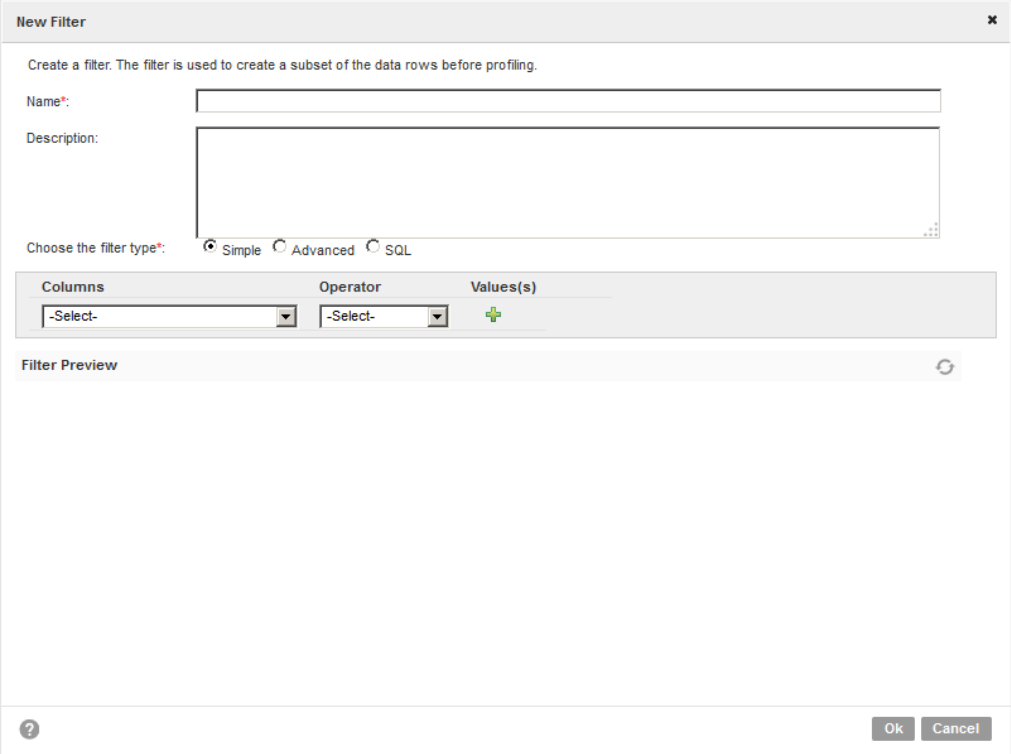
   A tela **Especificar Regras e Filtros** é exibida com o filtro no painel **Filtros**.

Criando um filtro simples

Você pode criar um filtro simples com operadores condicionais, como =, !=, >, <. Use o filtro para criar um subconjunto da fonte de dados original.

1. Na caixa de diálogo **Novo Filtro**, clique em **Simples**.

   A seguinte imagem mostra as opções que você pode usar para criar um filtro simples na caixa de diálogo **Novo Filtro**:



2. Insira um nome e uma descrição opcional.
3. Selecione uma coluna.
4. Selecione um operador condicional.
5. Insira um valor.
6. Opcionalmente, clique no ícone de adição (+) para adicionar mais filtros.
7. Clique em **OK**.

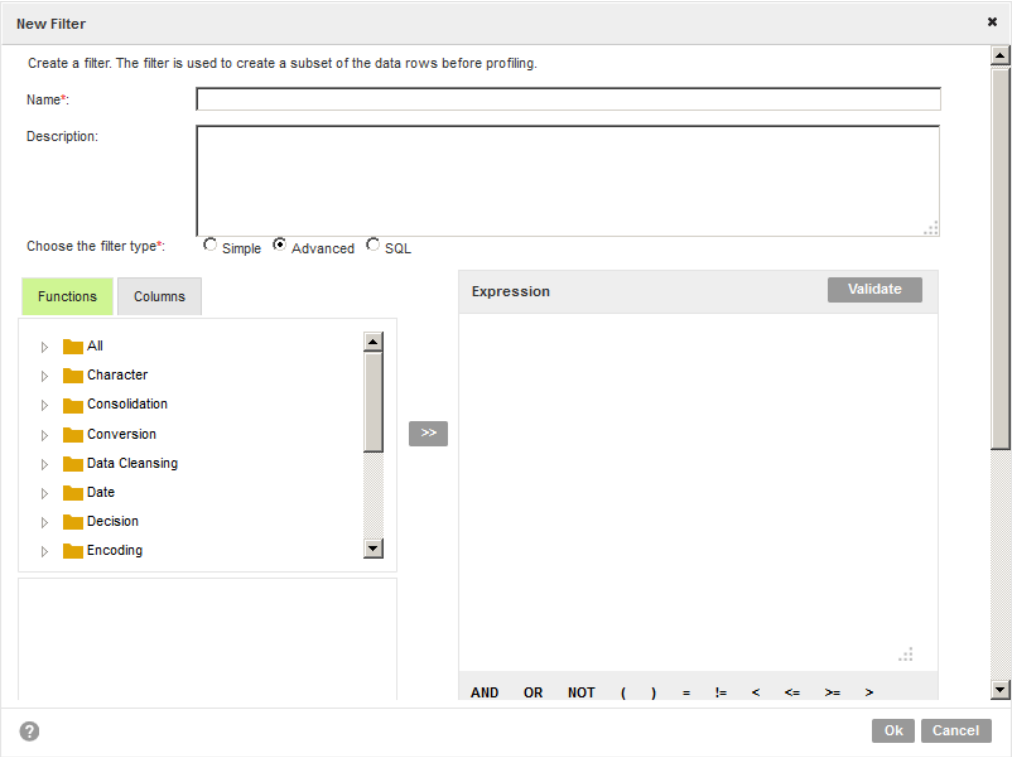
   A página **Especificar Regras e Filtros** é exibida com o filtro no painel Filtros.

Criando um filtro avançado

Você pode criar um filtro avançado com expressões, como AND, OR e NOT para fazer um subconjunto da fonte de dados original.

1. Na caixa de diálogo **Novo Filtro**, clique em **Avançado**.

   A seguinte imagem mostra opções de filtro avançado na caixa de diálogo **Novo Filtro**.



2. Insira um nome e uma descrição opcional para o filtro avançado.
3. Você pode criar um filtro avançado com os painéis Funções ou 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 um filtro avançado.
4. Para verificar o filtro avançado, clique em **Validar**.
5. Clique em **OK**.

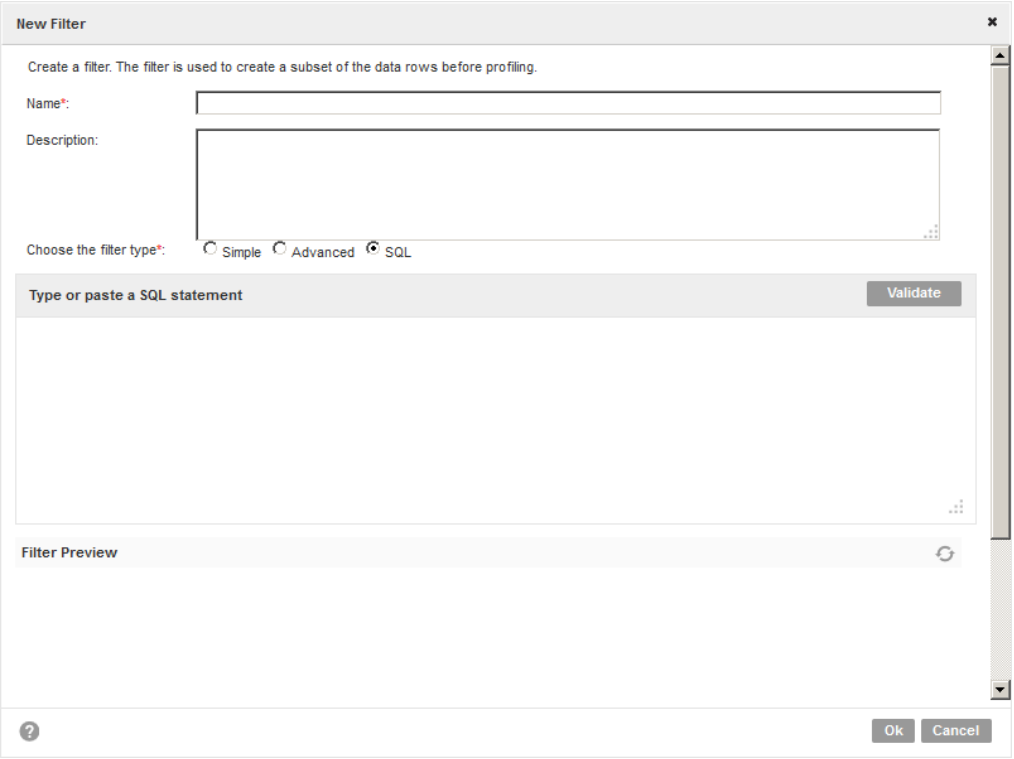
   A tela **Especificar Regras e Filtros** é exibida com o filtro no painel Filtros.

Criando um filtro SQL

Você pode criar um filtro SQL com consultas SQL. Você pode criar um filtro SQL para fontes de dados relacionais.

1. Na caixa de diálogo **Novo Filtro**, clique em **SQL**.

   A seguinte imagem mostra opções de filtro SQL na caixa de diálogo **Novo Filtro**:



2. Insira um nome e uma descrição opcional para o filtro SQL.
3. Na caixa de texto, digite ou cole uma consulta SQL.
4. Clique em **Validar** para verificar a consulta SQL.
5. Clique em **OK**.

   A página **Especificar Regras e Filtros** é exibida com o filtro SQL no painel Filtros.

Gerenciando filtros

Você pode editar e excluir filtros.

1. No espaço de trabalho **Biblioteca**, selecione o projeto que contém o perfil ou selecione o perfil que você deseja filtrar no painel **Ativos**.
2. Abra um perfil.
3. Na exibição resumida, clique em **Ações** > **Editar Perfil** para abrir o assistente de **Perfil**.
4. Clique em **Especificar Regras e Filtros**.
5. No painel Filtros, selecione um filtro e clique em **Ações** > **Editar Filtro**.

A caixa de diálogo **Editar Filtro** é exibida.

6. Edite as configurações de filtro e clique em **OK**.
7. Para excluir um filtro, selecione-o e clique em **Ações** > **Excluir Filtro**.

CAPÍTULO 9

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, 55
- Exibição Resumida, 56
- Exibição Detalhada, 58
- Estatísticas, 60
- Tipos de execução de perfil, 67
- Visão geral dos resultados de vários perfis de comparação, 69
- Busca Detalhada de Perfil da Coluna, 73
- Curadoria na ferramenta Analyst, 74
- Arquivos de exportação de perfil de coluna no Informatica Analyst, 75

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.

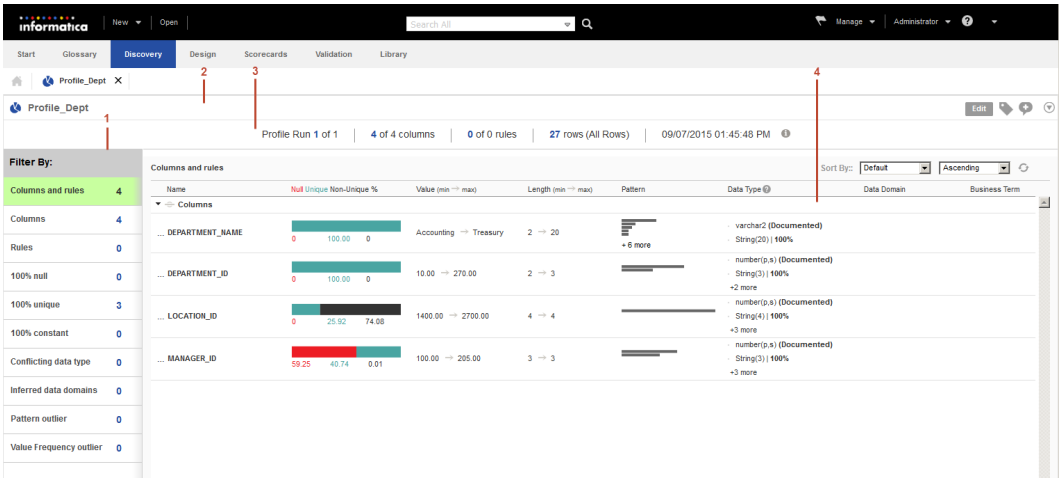
Nota: Você pode exibir e executar um perfil em fontes de dados Avro, JSON, Parquet e XML. Você pode exibir os resultados do perfil de execução mais recentes, execução de histórico e execução consolidada, e comparar os resultados do perfil de duas execuções de perfil.

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:

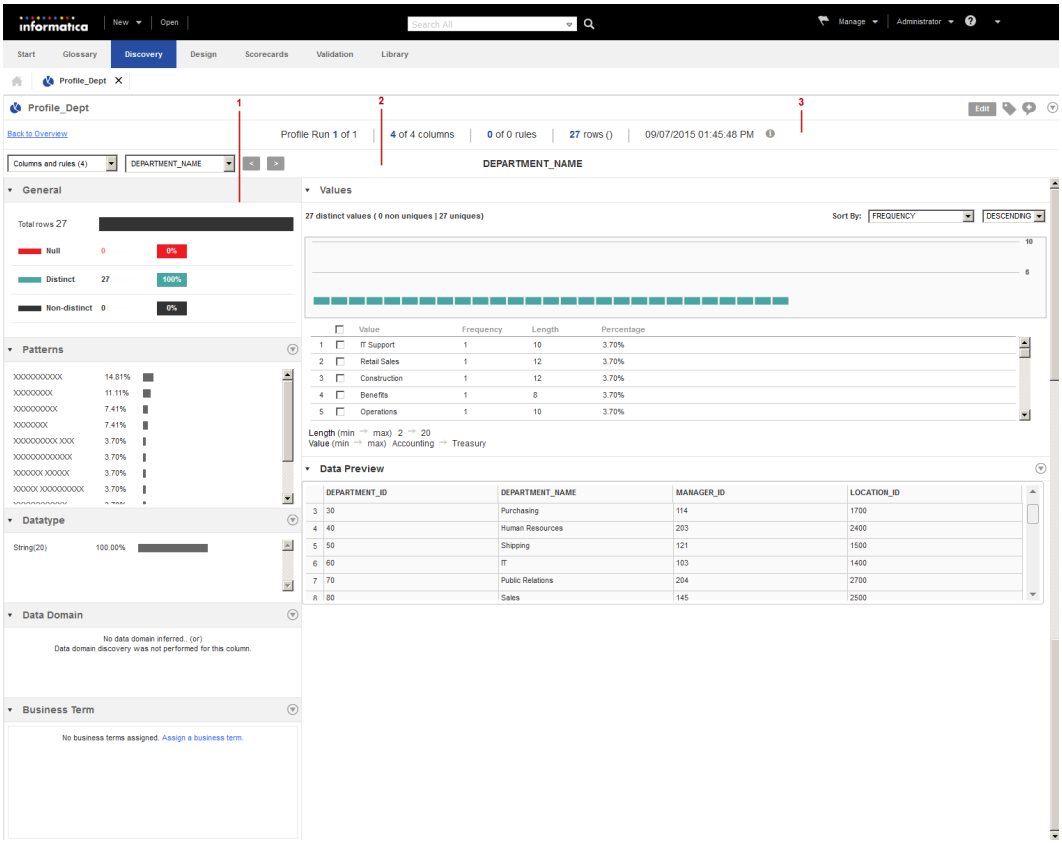
Opção de filtro padrão	Descrição
Colunas e regras	Exibe os resultados do perfil das colunas de origem e de regra. Você pode expandir e recolher as colunas de origem e de regra para exibir os resultados.
Colunas	Exibe os resultados do perfil das colunas de origem.
Regras	Exibe os resultados do perfil 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.

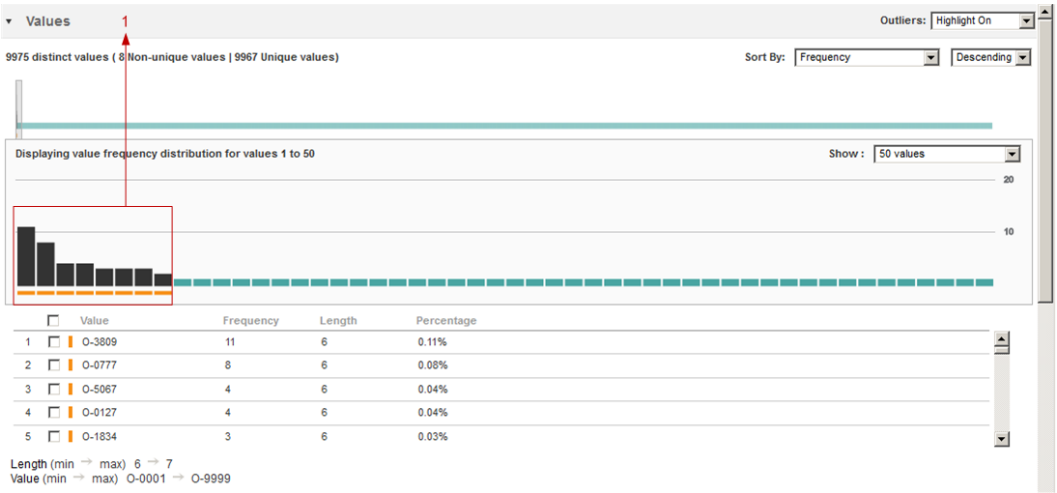
Exceções

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

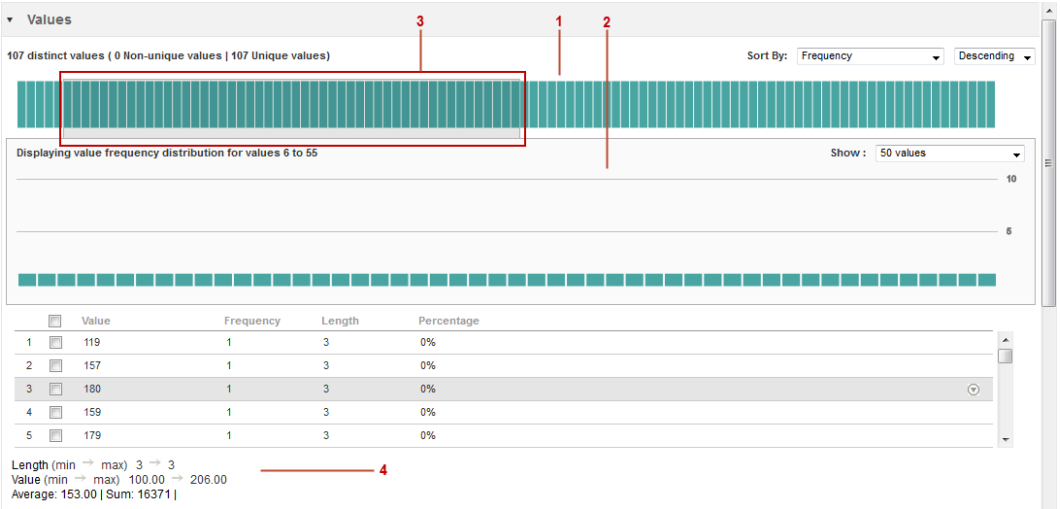
Valores na exibição detalhada

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.

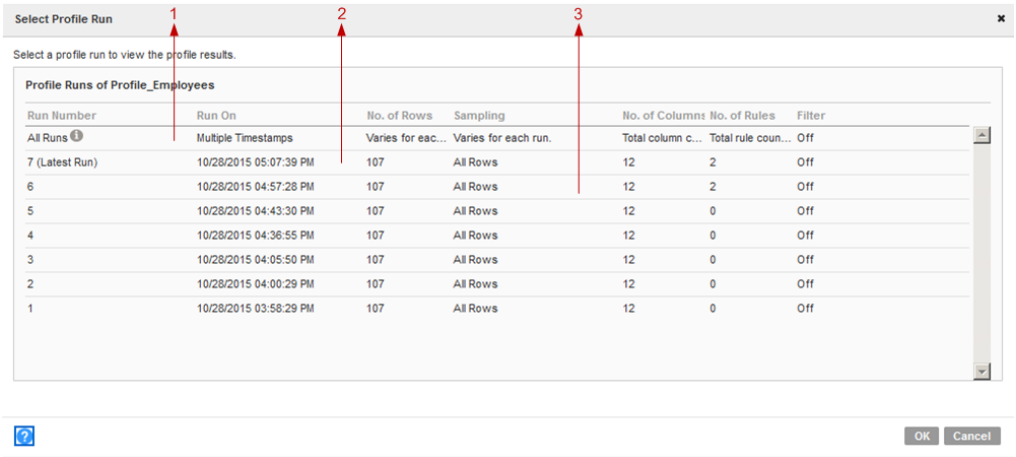
Selecionando uma execução de perfil

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.

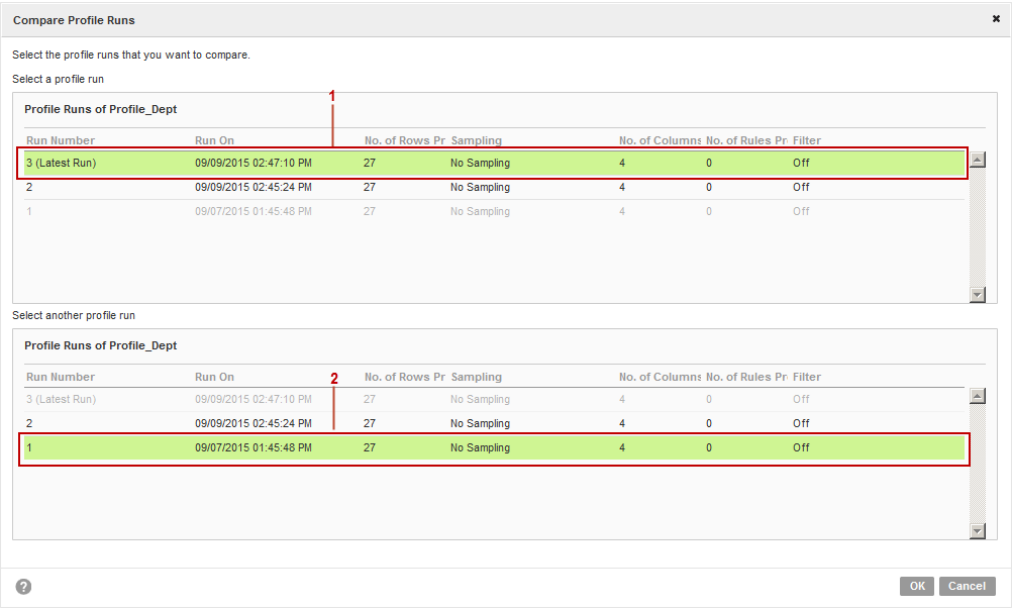
Na exibição resumida, você pode exibir os resultados de comparação de todas as colunas nas execuções de perfil.

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.

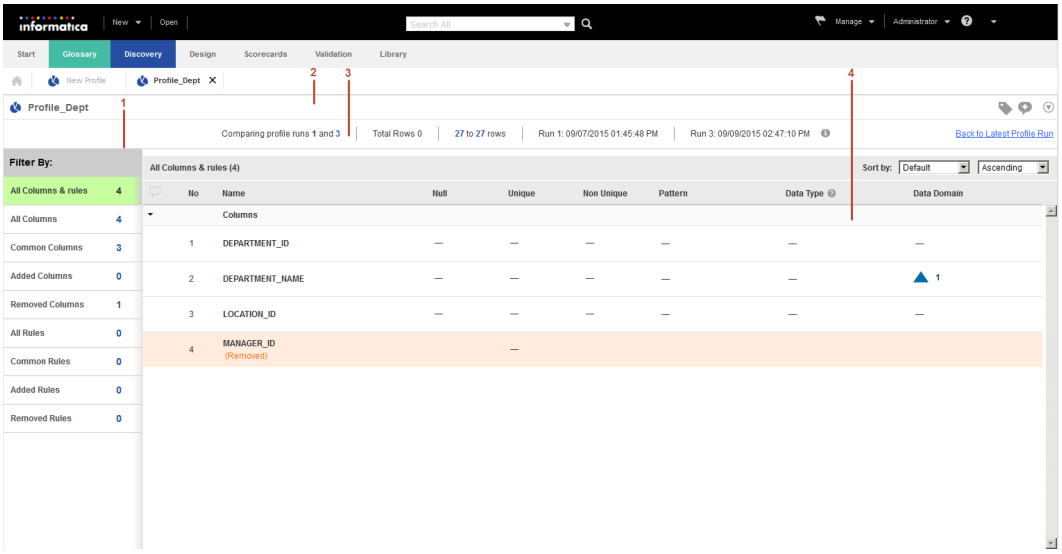
2. Selecione um perfil no painel **Execução A** e selecione outro perfil no painel **Execução B**.
3. Clique em **OK**.

   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 do perfil na exibição resumida

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

No modo de exibição de resumo, é possível visualizar de origem e virtuais. A saída para uma regra aparece como uma coluna virtual na exibição resumida. Quando você alterar a porta de saída para uma regra e compara a execução de perfil com uma execução histórica, a coluna de saída de regra histórica aparece no filtro **Regras Removidas** e a nova coluna de saída de regra aparece no filtro **Regras Adicionadas**. Se você alterar a lógica da regra para uma única regra de saída, ou se alterar as entradas para uma saída de várias regras em uma execução de perfil e a comparar com uma execução histórica, a saída dos filtros **Regras Adicionadas** e **Regras Removidas** não mudará. A saída do filtro não muda porque os filtros consideram apenas as alterações de nome nas colunas como entradas válidas para o filtro.

Você pode usar as seguintes opções de filtro padrão para exibir os resultados do perfil que atendem a condições específicas:

Opção de filtro padrão	Descrição
Todas as colunas e regras	Exibe os resultados do 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 do 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.

Opção de filtro padrão	Descrição
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 do 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.

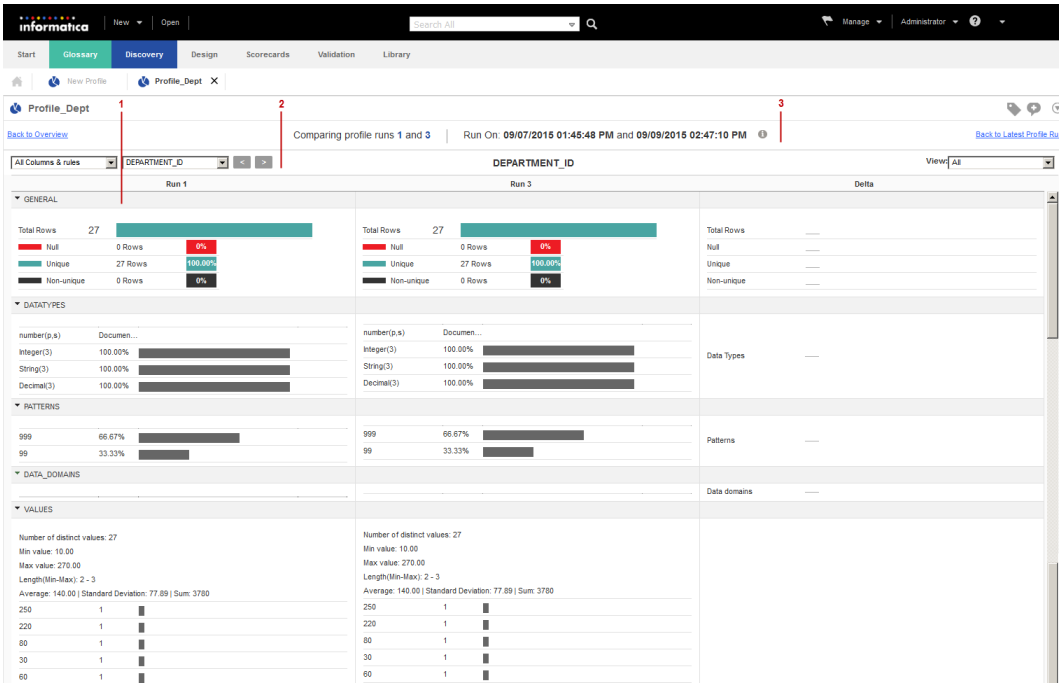
A exibição resumida mostra os resultados do perfil para todas as colunas de origem e colunas virtuais por padrão.

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 10

Termos comerciais, comentários e marcas no Informatica Analyst

Este capítulo inclui os seguintes tópicos:

- Termos comerciais, comentários e marcas na visão geral do Informatica Analyst, 78
- Termos Comerciais, 78
- Comentários, 79
- Marcas, 80

Termos comerciais, comentários e marcas na visão geral do Informatica Analyst

Você pode adicionar termos comerciais, comentários e marcas a um perfil ou a colunas no perfil. Você pode atribuir, exibir e editar termos comerciais, comentários e marcas nas exibições resumida e detalhada.

Termos Comerciais

Você pode atribuir termos comerciais a colunas em um perfil na ferramenta Analyst. Você pode editar um vínculo de ativo ou excluir um termo comercial de uma coluna. Um glossário comercial é um conjunto de termos que usam linguagem comercial para definir os conceitos para usuários comerciais. Um termo comercial fornece a definição e o uso comercial de um conceito.

Você pode atribuir, visualizar ou excluir termos comerciais em uma coluna na exibição resumida e na exibição detalhada. Para visualizar o termo comercial no espaço de trabalho **Glossário**, clique no termo comercial da exibição detalhada.

Você pode editar as propriedades de um vínculo de ativo para um termo comercial. Você pode exportar termos comerciais como uma coluna virtual para um arquivo CSV juntamente com os resultados do perfil.

Atribuindo termos comerciais a colunas

Atribua termos comerciais a colunas em um perfil nas exibições resumida e detalhada. Você pode excluir um termo comercial de uma coluna no painel **Termos Comerciais**. Você pode editar as propriedades de um vínculo de ativo de um termo comercial na caixa de diálogo **Editar Vínculo de Ativo**.

1. Na exibição resumida, clique com o botão direito do mouse no nome da coluna e selecione **Gerenciar Termos Comerciais**. Na exibição detalhada, selecione **Gerenciar Termos Comerciais** no menu **Ações** na guia **Termo Comercial**.

   O painel **Termos Comerciais** é exibido.
2. Clique no ícone de adição (+) ou no link **Atribuir termo comercial** para adicionar um termo comercial.

   O painel **Atribuir Termo Comercial** é exibido.
3. Selecione um termo comercial na lista de termos comerciais no painel **Atribuir Termo Comercial**. Insira um nome de ativo. Opcionalmente, você pode adicionar o contexto e a descrição para o ativo. Clique em **OK**.

   O termo comercial é exibido no painel **Termos Comerciais**.

Comentários

É possível adicionar comentários a um perfil e a colunas no perfil de forma que você possa fornecer informações adicionais para obter mais colaboração e análise.

No nível de perfil, você pode adicionar comentários sobre o perfil, definição de perfil ou metadados de perfil. Você pode visualizar os comentários de um perfil na exibição resumida.

Você pode adicionar e visualizar comentários de coluna nas exibições resumida e detalhada.

Você pode realizar as seguintes tarefas com comentários:

- Exporte comentários como uma coluna virtual para um arquivo CSV juntamente com os resultados do perfil. O arquivo CSV contém todos os comentários de um perfil e colunas no perfil.
- Pesquise os resultados do perfil usando as palavras-chave na coluna de comentários.
- Adicione comentários às colunas de origem e virtuais em um perfil.

Nota: Quando você não selecionar nenhuma coluna nem adicionar nenhum comentário de coluna, o painel de comentários na exibição resumida exibe os comentários do perfil.

Adicionando comentários a um perfil ou colunas

Você pode adicionar ou visualizar um comentário no painel **Comentários**.

1. Você pode adicionar comentários na exibição resumida ou detalhada.
   - Na exibição resumida, para adicionar comentários de perfil, clique em **Ações** > **Mostrar Comentários**.
   - Na exibição resumida, para adicionar comentários de coluna, clique com o botão direito do mouse em uma coluna e depois clique em **Mostrar Comentários**.
   - Na exibição detalhada, clique em **Adicionar Comentário** no painel **Geral**.

   O painel **Comentários** é exibido.
2. Clique em **Adicionar Comentário**.

Uma caixa de texto é exibida no painel **Comentários**.

3. Adicione um texto de comentário descritivo e clique em **Salvar**.

   O comentário é exibido no painel **Comentários** com o nome do usuário atual e a data e hora de criação.

Marcas

Você pode atribuir marcas a um perfil ou a colunas no perfil para agrupar objetos de acordo com o uso comercial.

Você pode visualizar ou atribuir marcas de perfil na exibição resumida. Você pode visualizar ou atribuir marcas de coluna nas exibições resumida e detalhada.

Você pode realizar as seguintes tarefas com marcas de coluna:

- Exporte marcas como uma coluna virtual para um arquivo CSV juntamente com os resultados do perfil. O arquivo CSV contém todas as marcas de um perfil e as colunas no perfil.
- Atribua marcas às colunas de origem e virtuais em um perfil.

Nota: Quando você não selecionar nenhuma coluna ou não adicionar nenhuma marca de coluna, o painel de marcas exibe as marcas de perfil na exibição resumida.

Atribuindo marcas a um perfil ou colunas

Adicione marcas a um perfil na exibição resumida. Adicione marcas a uma coluna nas exibições resumida e detalhada.

1. Você pode adicionar marcas na exibição resumida ou detalhada.
   - Na exibição resumida, para atribuir marcas a um perfil, clique em **Ações** > **Mostrar Marcas**.
   - Na exibição resumida, para adicionar marcas a uma coluna, clique com o botão direito do mouse em uma coluna e depois clique em **Mostrar Marcas**.
   - Na exibição detalhada, clique em **Adicionar Marca** no painel **Geral**.

   O painel **Marcas** é exibido.
2. Clique no ícone de adição (+) ou no link **Atribuir Marcas** para atribuir uma marca.

   A caixa de diálogo **Atribuir Marcas** é exibida.
3. Selecione uma ou mais marcas para atribuir a um perfil ou coluna. Clique em **OK** para abrir o painel **Marcas**.

   Nota: Para criar uma marca, clique em **Adicionar Nova Marca** no painel **Atribuir Marcas**.

CAPÍTULO 11

Scorecards no Informatica Analyst

Este capítulo inclui os seguintes tópicos:

- Scorecards na Visão Geral do Informatica Analyst, 81
- Processo de Scorecard no Informatica Analyst, 82
- Criando um scorecard no Informatica Analyst, 83
- Adicionando colunas a um scorecard existente, 84
- Executar um Scorecard, 85
- Exibir um Scorecard, 86
- Editar um Scorecard, 86
- Métricas, 87
- Grupos de Métrica, 88
- Buscas Detalhadas em Colunas, 90
- Gráficos de Tendência, 90
- Painel de scorecards no Informatica Analyst, 94
- Arquivos de Exportação de Scorecard no Informatica Analyst, 99
- Notificações de Scorecard, 99
- Linhagem de Scorecard, 102

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.

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 o painel de scorecards no espaço de trabalho **Scorecards**. No painel de scorecards, você pode visualizar os objetos de dados que têm scorecards, scorecards em um projeto, a tendência de execuções de scorecards nos últimos seis meses e a agregação das métricas boas, aceitáveis e inaceitáveis para todas as execuções de scorecards em um mês.

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.
16. Visualize informações consolidadas sobre os scorecards aos quais você tem acesso de leitura.

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

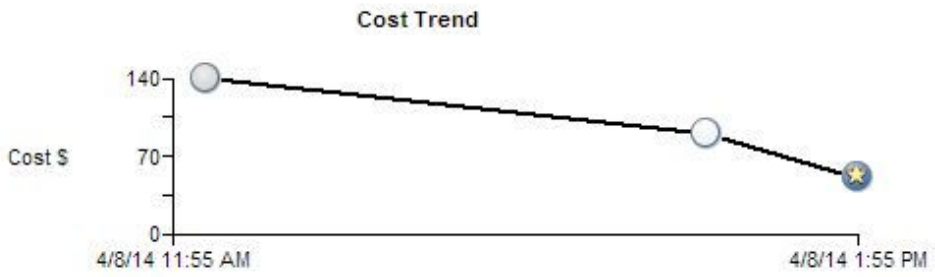
Como analista de dados, você pode monitorar a qualidade de dados para analisar se alterações em mapeamentos e em outros processos provocam aumentos na pontuação de qualidade de dados. Depois de medir a alteração na qualidade de dados, você pode informar essa alteração para que ela seja analisada e utilizada 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

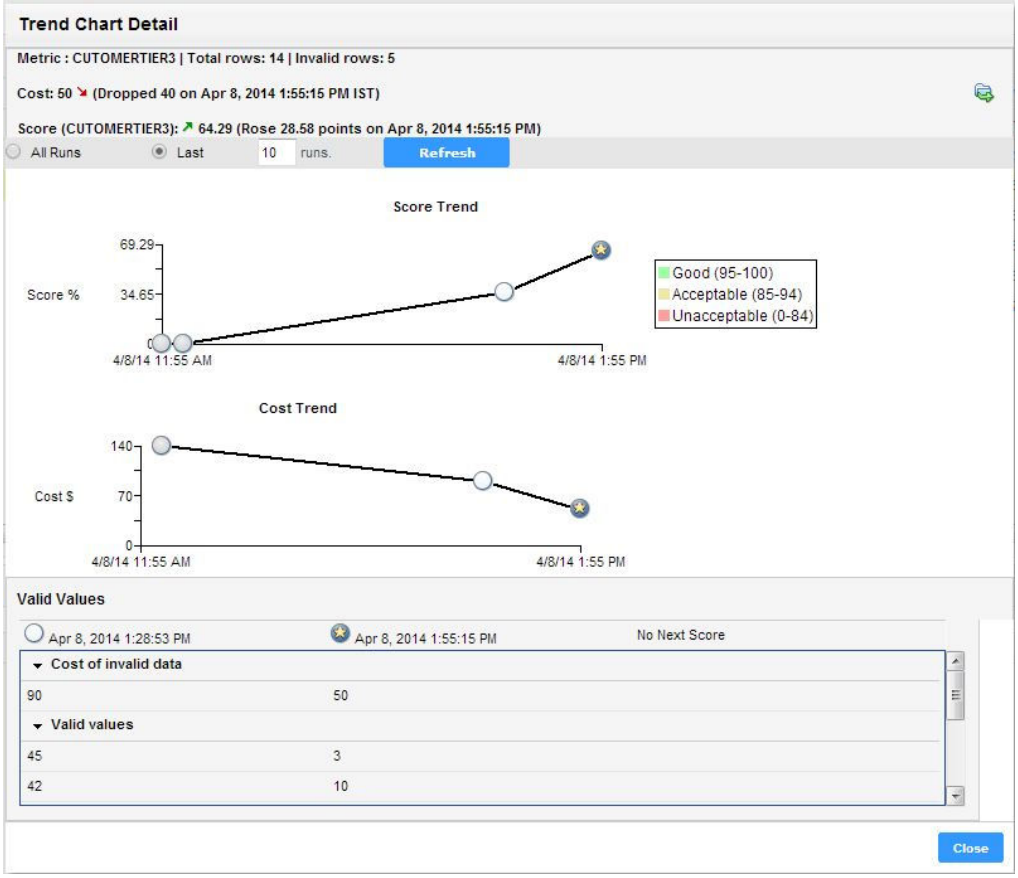
Você pode exibir gráficos de tendências para cada métrica para monitorar como a pontuação ou o custo de dados inválidos muda com o passar do tempo.

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**.

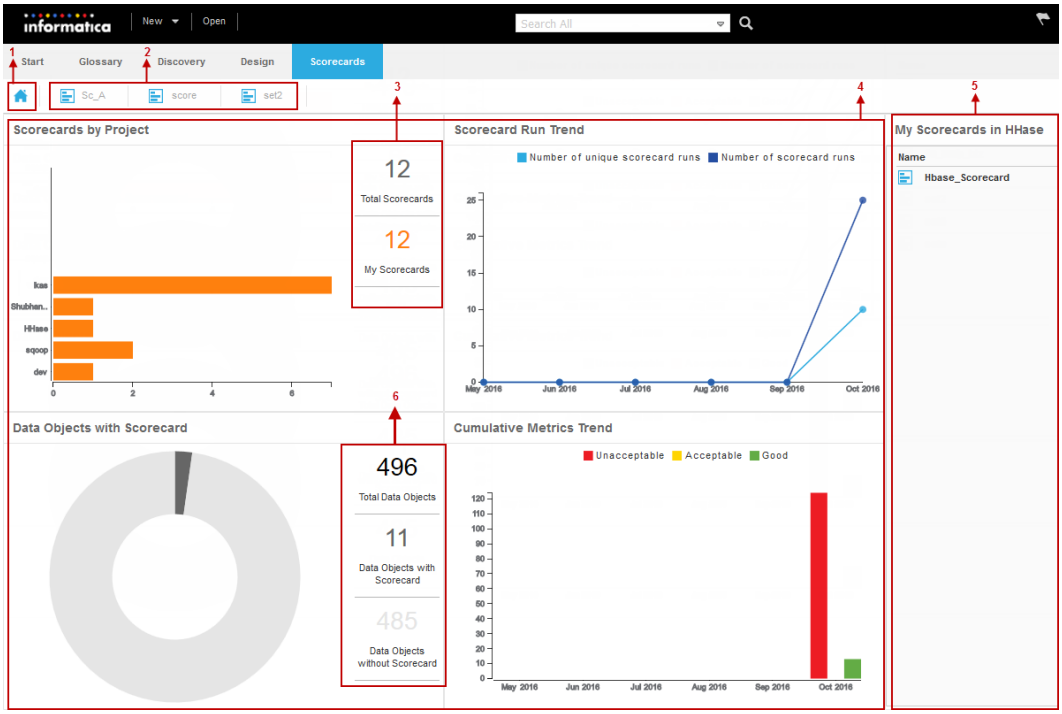
Painel de scorecards no Informatica Analyst

O espaço de trabalho **Scorecards** no Informatica Analyst exibe o painel de scorecards. No painel de scorecards, você pode visualizar os objetos de dados com scorecards, a tendência de execuções de scorecards nos últimos seis meses, scorecards em um projeto, um agregado de métricas boas, aceitáveis e inaceitáveis para todas as execuções de scorecards em um mês e o painel de lista de ativos.

O painel de scorecards na sua máquina não é atualizado automaticamente quando os scorecards são modificados por outros usuários. Use a tecla de função F5 ou alterne entre espaços de trabalho ou guias de resultados de scorecards para atualizar o painel de scorecards.

É possível visualizar os dados como uma série de dados ou como pontos de dados nos painéis. Os pontos de dados aparecem como pequenos círculos opacos, e as séries de dados aparecem como barras horizontais, barras verticais ou fatias nos gráficos.

A seguinte imagem mostra o painel de scorecards e o painel de ativos no espaço de trabalho **Scorecards**:



1. Ícone do painel de scorecards. Mostra o painel de scorecards.
2. Guias de resultados de scorecards. Mostra os resultados dos scorecards abertos.
3. Legenda no painel Scorecards por Projeto. Mostra o número total de scorecards em todos os projetos e o número total de scorecards aos quais você tem acesso de leitura em todos os projetos.
4. Painel de scorecards. Mostra os scorecards por projeto, a tendência de execuções de scorecards, objetos de dados com scorecard e painéis de tendências de métricas cumulativas no painel.
5. Painel de lista de ativos. Exibe a lista de scorecards ou objetos de dados associados a uma legenda, séries de dados ou ponto de dados no gráfico.
6. Legenda em Objetos de Dados com o painel Scorecard. Exibe o número total de objetos de dados, o número de objetos de dados com scorecards e o número de objetos de dados sem scorecards

Depois que você clicar em um ponto de dados ou em uma série de dados no painel de scorecards, os scorecards que são mapeados para o ponto de dados ou a série de dados aparecem no painel de lista de

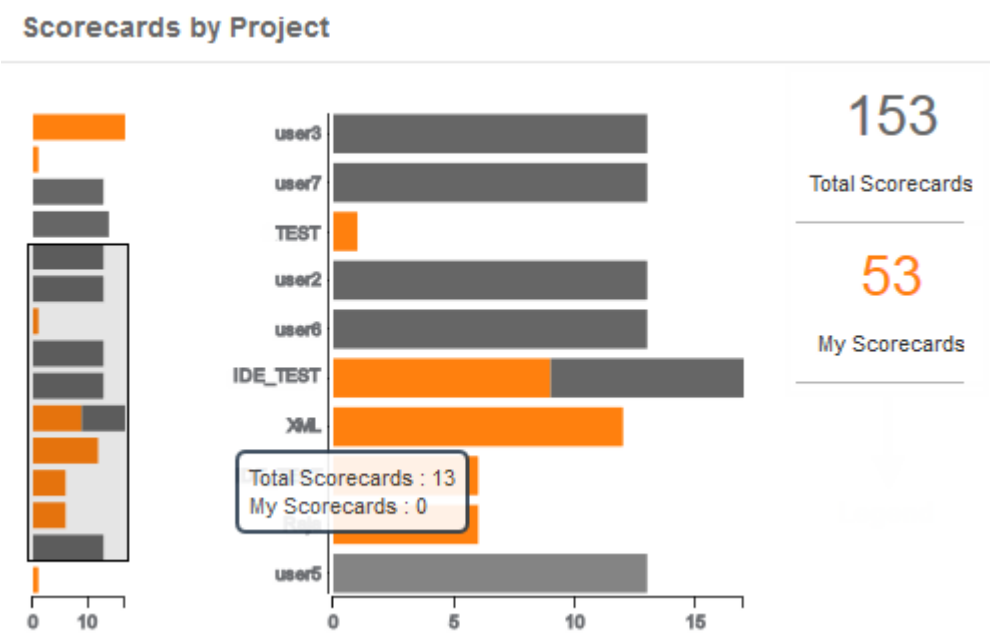
ativos. Depois que você clicar em um scorecard no painel de lista de ativos, os resultados do scorecard aparecem em uma guia no espaço de trabalho **Scorecards**. O painel de listas de ativos exibe os scorecards para os quais você tem acesso de leitura.

Scorecards por Projeto

O painel **Scorecards por Projeto** exibe os projetos no repositório do Modelo, juntamente com o número de scorecards em cada projeto como um gráfico de barras. Uma barra no gráfico de barras representa um projeto. O eixo X no gráfico exibe o número de scorecards, e o eixo Y exibe os projetos que têm scorecards.

Os scorecards em um projeto aparecem em cinza no gráfico de barras, e os scorecards aos quais você tem acesso de leitura aparecem em laranja no gráfico de barras. O seção **Total de Scorecards** na legenda mostra o número total de scorecards no repositório do Modelo. A seção **Meus Scorecards** na legenda mostra o número de scorecards aos quais você tem acesso de leitura no repositório do Modelo.

A seguinte imagem mostra o painel **Scorecards por Projeto** no painel de scorecards:



Você pode exibir os seguintes gráficos no painel:

- Gráfico detalhado. Mostra todos os projetos no repositório do Modelo com scorecards e o número de scorecards em cada projeto. Se o número de projetos for maior que 10, o painel **Scorecards por Projeto** mostrará um controle deslizante.
- Gráfico em miniatura. Mostra todos os projetos e o número de scorecards em cada um dentro de um controle deslizante no gráfico detalhado.

Quando você move o ponteiro sobre o gráfico em miniatura, o número total de scorecards e o número de scorecards aos quais você tem acesso em um projeto aparecem em um rótulo de dados.

Para visualizar os scorecards em um projeto ao qual você tem acesso de leitura, clique na parte laranja da barra horizontal. Para visualizar os scorecards aos quais você tem acesso de leitura no repositório do Modelo, clique em **Meus Scorecards** no gráfico de barras. Os scorecards aparecem no painel de lista de ativos. Clique em um scorecard no painel de ativos para visualizar os resultados do scorecard.

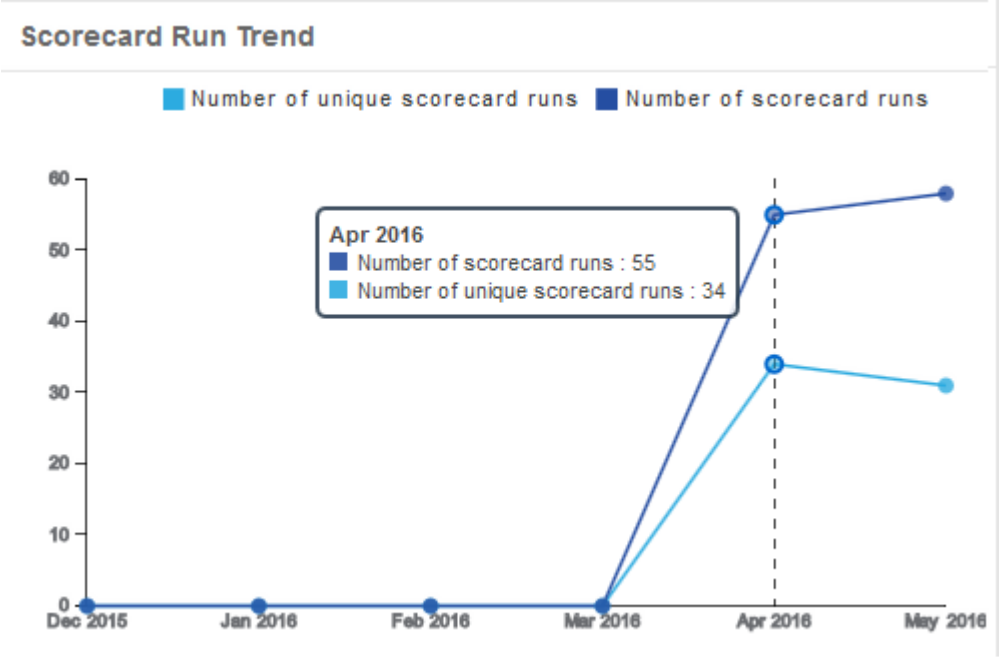
Tendência de Execução do Scorecard

O painel **Tendência de Execução do Scorecard** exibe a tendência de execução do scorecard para o mês atual e os cinco meses anteriores como gráficos de linha com marcadores. O eixo X do gráfico mostra o mês atual e os últimos cinco meses, enquanto o eixo Y mostra o número de scorecards. Um marcador é um ponto de dados no gráfico de linhas. Quando você move o ponteiro sobre um marcador no gráfico, o resumo de execuções do scorecard para o mês aparece em um rótulo de dados.

Você pode exibir as seguintes marcadores no painel:

- **Número de execuções de scorecard**. O marcador exibe o número total de execuções de scorecard no mês.
- **Número de execuções de scorecard exclusivas**. O marcador exibe o número total de execuções de scorecard exclusivas no mês.

A seguinte imagem mostra o painel **Tendência de Execução do Scorecard** no painel de scorecards:



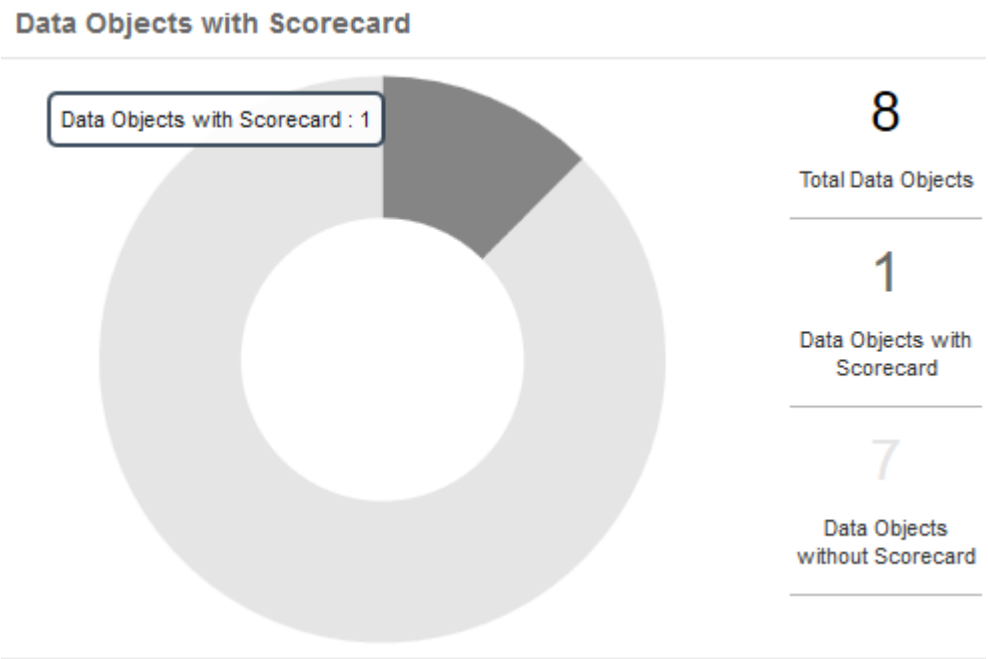
Quando você clicar em um marcador no painel, os scorecards que são mapeados para esse marcador aparecem no painel de lista de ativos. É possível visualizar os scorecards aos quais você tem acesso de leitura. Clique em um scorecard no painel de lista ativos para visualizar os resultados do scorecard no espaço de trabalho **Scorecards**.

Objetos de Dados com Scorecards

O painel **Objetos de Dados com Scorecards** exibe um gráfico de rosca. É possível exibir o número de objetos de dados com e sem scorecards como fatias no gráfico.

Quando você move o ponteiro sobre o gráfico de rosca, os dados que são mapeados para a fatia aparecem em um rótulo de dados.

A seguinte imagem mostra o painel **Objetos de Dados com Scorecards** no painel de scorecards:



A legenda no painel **Objetos de Dados com Scorecards** exibe as seguintes estatísticas de dados:

- Total de Objetos de Dados. Mostra o número total de objetos de dados na pasta **Objeto de Dados** do painel **Ativos** do espaço de trabalho **Biblioteca**. Os objetos de dados incluem objetos de dados lógicos e objetos de dados personalizados.
- Objetos de Dados com Scorecards. Mostra o número de objetos de dados com scorecards.
- Objetos de Dados sem Scorecards. Mostra o número de objetos de dados sem scorecards.

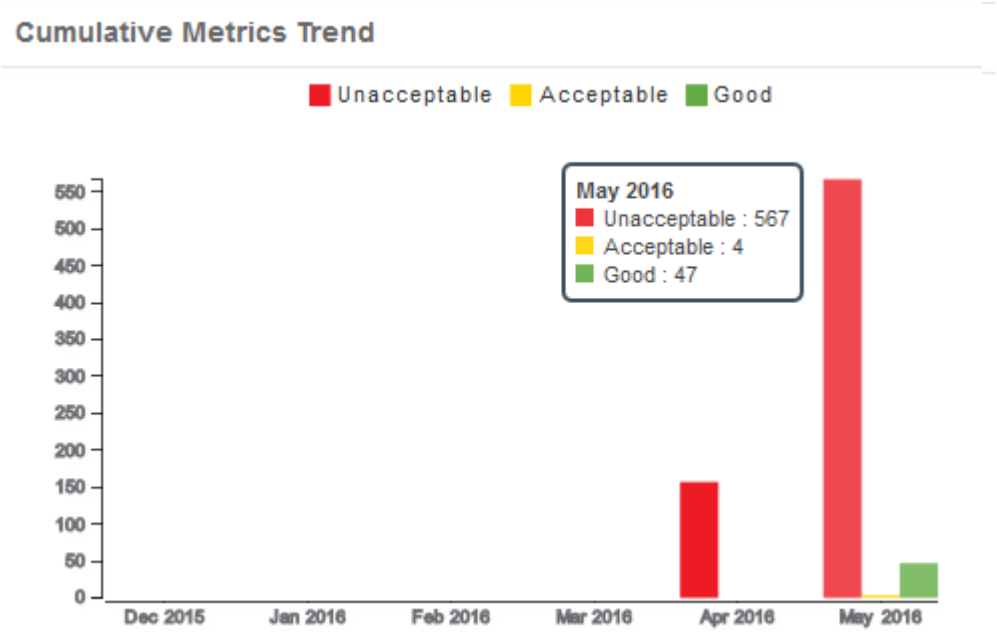
Depois de clicar nas fatias no gráfico de rosca ou na legenda de **Objetos de Dados com Scorecards** e **Objetos de Dados sem Scorecards**, os scorecards que são mapeados para a fatia no gráfico de rosca ou para a legenda aparecem no painel de lista de ativos.

Tendências de Métricas Cumulativas

O painel **Tendências de Métricas Cumulativas** exibe gráficos de colunas. Você pode visualizar a agregação de métricas boas, aceitáveis e inaceitáveis para todas as execuções de scorecard em um mês no gráfico como barras verticais. Você pode usar o painel **Métricas Cumulativas** para exibir e analisar a tendência de métricas para o mês atual e os últimos cinco meses.

Quando você move o ponteiro sobre uma barra vertical, o resumo de métricas para o mês é exibido em um rótulo de dados. Quando você clica em uma barra vertical no painel, os scorecards relevantes aparecem no painel de lista de ativos. É possível visualizar os scorecards aos quais você tem acesso de leitura. Clique em um scorecard no painel de lista de ativos para exibir os resultados do scorecard.

A seguinte imagem mostra o painel **Tendências de Métricas Cumulativas** no painel de scorecards:



O painel de lista de ativos pode exibir alguns scorecards na lista de métricas inaceitáveis e na lista de boas métricas quando a tendência do indicador muda com o tempo em um mês. Para analisar as métricas, abra o scorecard para exibir os resultados do scorecard.

Exemplo

Você é o gerente regional de uma cadeia de lojas de varejo no estado da California. Você cria o scorecard Sales_SC na tabela Sales.

É possível definir os seguintes limites para a métrica Sales_amt no scorecard Sales_SC:

- Inaceitável = 0% a 40%
- Aceitável = 41% a 89%
- Bom = 90% a 100%

Para capturar os dados de vendas em uma base diária, é possível usar o serviço de Agendador para executar os scorecards todas as noites na tabela Sales. Você prepara um relatório mensal de gerenciamento do mês de maio e usa o painel de scorecards para verificar seu relatório. Para o mês de maio, quando você clica nas barras verticais de Métricas inaceitáveis e Boas métricas no gráfico, o scorecard Sales_SC aparece no painel de lista de ativos.

Quando você analisa as vendas do mês, encontra a seguinte tendência:

1. De 1º de maio a 25 de maio, a métrica Sales_amt é calculada abaixo de 40% e está marcada como uma métrica inaceitável.
2. Na última semana de maio, devido a um aumento nas vendas, a métrica Sales_amt é calculada acima de 98% e marcada como uma boa métrica.

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.
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.

Marca	Descrição
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 172.

CAPÍTULO 12

Descoberta do Domínio de Dados no Informatica Analyst

Este capítulo inclui os seguintes tópicos:

- Descoberta do Domínio de Dados no Informatica Analyst - Visão Geral, 104
- Glossário do Domínio de Dados no Informatica Analyst, 104
- Opções de Descoberta do Domínio de Dados no Informatica Analyst, 107
- Criando um Perfil para Executar Descoberta do Domínio de Dados no Informatica Analyst, 110
- Editando um Perfil no Informatica Analyst, 111
- Executando um Perfil para Realizar a Descoberta do Domínio de Dados, 112
- Resultados da Descoberta do Domínio de Dados no Informatica Analyst, 112
- Arquivos de Exportação de Descoberta do Domínio de Dados no Informatica Analyst, 113

Descoberta do Domínio de Dados no Informatica Analyst - Visão Geral

Ao criar um perfil para realizar a descoberta do domínio de dados, você pode selecionar as colunas de origem, os domínios de dados com os quais deseja corresponder os dados e o nome das colunas, além de opções de amostragem. É possível escolher critérios de conformidade para a descoberta do domínio de dados e excluir valores nulos durante essa descoberta.

Você pode criar um perfil com uma opção de amostragem e filtros para realizar a descoberta do domínio de dados. Ao executar o perfil, você aplica a opção de amostragem e os filtros à fonte de dados e gera um conjunto de dados. O processo de descoberta do domínio de dados usa o conjunto de dados para descobrir domínios de dados.

Glossário do Domínio de Dados no Informatica Analyst

O glossário do domínio de dados lista domínios de dados e grupos de domínios de dados. É possível classificar a lista por domínios de dados ou grupos de domínios de dados. Use o glossário do domínio de

dados para pesquisar, adicionar, editar e remover domínios de dados e grupos de domínios de dados. Você pode exibir e fazer alterações na lógica da regra associada a domínios de dados na ferramenta Developer.

Criando um Grupo do Domínio de Dados no Informatica Analyst

Um grupo do domínio de dados organiza domínios de dados em grupos específicos, como Informações Pessoais de Saúde, Informações de Identificação Pessoal ou qualquer outro grupo conceitual que seja relevante para o projeto.

1. Clique em **Gerenciar** > **Glossário do Domínio de Dados**.

   O glossário do domínio de dados é aberto em uma guia que lista os domínios de dados e grupos de domínios de dados atuais.
2. No Navegador, clique em **Ações** > **Novo** > **Grupo do Domínio de Dados**.

   A caixa de diálogo **Criar Grupo de Domínio de Dados** é exibida.
3. Insira um nome e descrição para o grupo do domínio de dados.
4. Clique em **Avançar**.
5. No painel **Domínios de Dados Disponíveis**, selecione os domínios de dados que você deseja adicionar ao grupo do domínio de dados e clique em **Adicionar**.

   A ferramenta Analyst move os domínios de dados selecionados para o painel **Domínios de dados selecionados**.
6. Clique em **Concluir**.

   A ferramenta Analyst adiciona o grupo do domínio de dados ao glossário do domínio de dados.

Criando um Domínio de Dados no Informatica Analyst

Você pode criar domínios de dados, adicioná-los ao glossário do domínio de dados e agrupar os domínios de dados em um ou mais grupos. Para criar um domínio de dados, você pode usar regras de dados predefinidas e regras de nome de coluna ou gerar domínios de dados a partir dos valores e padrões nos resultados do perfil de coluna.

Quando você cria um domínio de dados, a ferramenta Analyst copia regras e outros objetos dependentes associados ao domínio de dados para o glossário do domínio de dados. Para editar uma regra associada a um domínio de dados, você deve ir para a regra original e fazer alterações nela. Você poderá então associar a regra modificada ao domínio de dados novamente.

1. Clique em **Gerenciar** > **Glossário do Domínio de Dados**.

   O glossário do domínio de dados é aberto em uma guia que lista os domínios de dados e grupos de domínios de dados atuais.
2. No Navegador, clique em **Ações** > **Novo** > **Domínio de Dados**.

   A caixa de diálogo **Criar Domínio de Dados** é exibida.
3. Insira um nome e descrição para o domínio de dados.
4. Clique na caixa de seleção **Regra de Dados** para descobrir domínios de dados com base nos dados da coluna. Você também pode marcar a caixa de seleção **Regra de Nome da Coluna** para descobrir domínios de dados com base nos títulos da coluna da fonte de dados.

   O botão **Selecionar** fica ativado.
5. Clique em **Selecionar** para abrir a caixa de diálogo **Selecionar Regra**.
6. Selecione uma regra apropriada e clique em **OK**.

   As regras que você selecionou aparecem nos campos **Regra de Dados** e **Regra de Nome da Coluna**.

7. Clique em **Avançar**.
8. No painel **Grupos de Domínio de Dados Disponíveis**, selecione os grupos de domínios de dados nos quais você deseja incluir o domínio de dados e clique em **Adicionar**.

   A ferramenta Analyst move os grupos de domínios de dados selecionados para o painel **Grupos de Domínio de Dados Selecionados**.
9. Clique em **Concluir**.

   A ferramenta Analyst adiciona o domínio de dados ao glossário do domínio de dados.

Criando um Domínio de Dados a Partir de Resultados do Perfil no Informatica Analyst

Execute um perfil de coluna para exibir os valores e padrões dos dados de origem. Você pode então verificar os resultados do perfil e criar um domínio de dados a partir deles.

1. Execute um perfil de coluna para exibir seus resultados.

   Os resultados do perfil aparecem na exibição resumida.
2. Na exibição resumida, clique em uma coluna para visualizar os resultados de coluna na exibição detalhada.
3. No painel **Valores** ou no **Padrões**, selecione o valor ou padrão com base no qual você deseja criar um domínio de dados.
4. Clique com o botão direito do mouse no valor ou padrão e selecione **Criar Domínio de Dados**.

   A caixa de diálogo **Criar Domínio de Dados** é exibida.
5. Digite o nome do domínio de dados e uma descrição opcional.
6. Clique em **Criar**.

   O domínio de dados é adicionado ao glossário do domínio de dados.

Localizar Domínios de Dados e Grupos de Domínios de Dados no Informatica Analyst

Você pode pesquisar domínios de dados e grupos de domínios de dados específicos no glossário do domínio de dados. Você pode escolher entre a exibição **Domínio de Dados** e a exibição **Grupo do Domínio de Dados** para exibir a lista de domínios de dados no glossário.

Por exemplo, pode haver um domínio de dados **Código Postal** que você adicionou ao grupo do domínio de dados **Informações de Identificação Pessoal**. Você pode encontrar mais informações sobre Código Postal e seu grupo do domínio de dados Informações de Identificação Pessoal das seguintes maneiras:

Pesquisar domínios de dados.

Digite parte do nome do domínio de dados como **Código** ou **Postal** no campo de texto na parte superior do Navegador. Se você estiver na exibição **Grupos de Domínios de Dados**, a ferramenta Analyst listará **Informações de Identificação Pessoal**, que é o grupo do domínio de dados que contém **Código Postal**. Se você estiver na **Exibição de Domínios de Dados**, a ferramenta Analyst listará todos os domínios de dados que contiverem a string de pesquisa, **Código** ou **Postal**.

Nota: A pesquisa não diferencia maiúsculas de minúsculas.

Exibir todos os grupos de domínios de dados e domínios de dados dentro deles.

No Navegador, clique em **Exibir Grupo do Domínio de Dados**.

Exibir todos os domínios de dados.

No Navegador, clique em **Exibir Domínio de Dados**.

Exibir propriedades de um domínio de dados.

Verifique se você está em **Exibir Domínio de Dados**. No Navegador, clique em **Código Postal** para exibir suas propriedades no painel direito. Você pode exibir o nome, tipo, descrição, regras associadas e grupos de domínio aos quais ele pertence, neste caso, **Informações de Identificação Pessoal**.

Exibir propriedades de um grupo do domínio de dados.

Verifique se você está em **Exibir Grupo do Domínio de Dados**. No Navegador, clique em **Informações de Identificação Pessoal** para exibir suas propriedades no painel direito. Você pode exibir o nome, tipo, descrição e a lista de domínios de dados, incluindo **Código Postal**, em **Informações de Identificação Pessoal**.

Atualizar o glossário do domínio de dados.

No Navegador, clique em **Ações** > **Atualizar**. O glossário do domínio de dados exibe a lista atual de domínios de dados ou grupos de domínios de dados com base na exibição em que você está.

Opções de Descoberta do Domínio de Dados no Informatica Analyst

Use as opções de descoberta do domínio de dados para escolher as colunas, domínios de dados e opções de inferência para descoberta do domínio de dados. As opções de inferência incluem escolher se você quer executar a descoberta do domínio de dados com base em uma regra nos dados da coluna, no nome da coluna ou em ambos.

Seleção de Colunas de Domínio de Dados no Informatica Analyst

Você pode clicar em **Editar** na tela **Especificar Configurações** para escolher as colunas que deseja executar como parte da descoberta do domínio de dados. É possível visualizar todas as colunas da fonte de dados na tela **Selecionar Origem** do assistente de perfil. Diferentes colunas podem ser escolhidas para o perfil de coluna e a descoberta de domínio de dados.

A seguinte tabela descreve as propriedades da caixa de diálogo **Editar** para a descoberta do domínio de dados:

Opção	Descrição
Nome	Exibe o nome da coluna.
Tipo	Exibe o tipo de dados documentado da coluna.
Precisão	Exibe a precisão máxima da coluna.
Escala	Exibe a escala da coluna.
Anulável	Indica uma coluna que pode ter valores nulos.
Chave	Indica se a coluna está documentada como uma chave primária ou uma chave externa.

Seleção de Domínio de Dados no Informatica Analyst

O painel **Domínios de Dados** na tela **Especificar Configurações** lista todos os domínios de dados do glossário do domínio de dados. Você pode escolher os domínios de dados que deseja executar como parte da descoberta do domínio de dados.

A seguinte tabela descreve as propriedades do **Domínios de Dados** para descoberta de domínio de dados:

Opção	Descrição
Nome	Exibe o nome do domínio de dados. Você pode escolher um ou mais domínios de dados ou um grupo de domínios de dados.
Descrição	Exibe a descrição do domínio de dados.
DomainGroups	Exibe o nome do grupo de domínios de dados ao qual o domínio de dados pertence.

Opções de Inferência de Domínio de Dados no Informatica Analyst

As opções de inferência determinam se a descoberta do domínio de dados deve ser executada nos dados da coluna, no nome da coluna ou em ambos. Você pode especificar o número máximo de linhas de origem que o perfil pode analisar. Você pode escolher critérios de conformidade para a descoberta do domínio de dados. Você pode excluir valores nulos da descoberta do domínio de dados. Você pode definir as opções de inferência de domínio de dados na tela **Especificar Configurações** do assistente de perfil.

A tabela a seguir descreve as opções de inferência para descoberta do domínio de dados:

Opção	Descrição
Dados	Executa o perfil nos dados das colunas.
Colunas	Executa o perfil nos títulos das colunas.
Data e Colunas	Executa o perfil tanto nos dados quanto nos títulos das colunas.
Porcentagem mínima de linhas	A porcentagem mínima de conformidade de linhas no conjunto de dados necessária para uma correspondência de domínio de dados.
Número mínimo de linhas	O número mínimo de linhas no conjunto de dados necessárias para uma correspondência de domínio de dados.
Excluir valores nulos para descoberta do domínio de dados	Exclui os valores nulos do conjunto de dados para a descoberta do domínio de dados.
Editar	Selecione as colunas para a descoberta do domínio de dados.
Todas as Linhas	Executa o perfil em todas as linhas a partir da origem.
Fazer amostragem das primeiras	Escolha o número máximo de linhas nas quais o perfil pode ser executado. A ferramenta Analyst escolhe as linhas começando com a primeira linha da origem.
Amostra aleatória	Escolha uma amostra aleatória de linhas na fonte de dados.

Opção	Descrição
Amostragem aleatória (automática)	A ferramenta Analyst escolhe uma amostra aleatória de linhas com base no tamanho da fonte 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.

Porcentagem Mínima de Conformidade

Você pode escolher uma porcentagem mínima de linhas no conjunto de dados como critérios de conformidade para a descoberta do domínio de dados.

A porcentagem de conformidade é a proporção entre o número de linhas correspondentes dividido pelo número total de linhas.

Nota: A ferramenta Analyst considera valores nulos como linhas sem correspondência. As colunas que contêm um número alto de valores nulos podem não resultar em inferência de domínio de dados, a menos que você especifique um valor baixo para a porcentagem mínima de conformidade.

Exemplo

Você tem uma fonte de dados com 10.000 linhas em que a coluna Comentários tem Números de Previdência Social em 2.500 linhas. Crie um perfil de coluna com descoberta do domínio de dados e defina uma porcentagem mínima de linhas de 30% como os critérios de conformidade. Quando você executa o perfil, os resultados do perfil não exibem os Números de Previdência Social como um domínio de dados inferido, pois os critérios mínimos de conformidade correspondem a 30% das linhas, ou 3.000 linhas, na fonte de dados.

Mínimo de Linhas em Conformidade

Você pode escolher um número mínimo de linhas no conjunto de dados como critérios de conformidade para a descoberta do domínio de dados.

Exemplo

Você tem uma fonte de dados com 10.000 linhas em que a coluna Comentários tem um endereço de e-mail em três linhas. Você cria um perfil de coluna com um perfil de descoberta do domínio de dados e define o número mínimo de linhas como 1 como os critérios de conformidade. Quando você executa o perfil, os resultados do perfil exibem o endereço de e-mail como um domínio de dados inferido com três linhas em conformidade, juntamente com os outros domínios de dados inferidos.

Excluir valores nulos na descoberta do domínio de dados

Você pode excluir valores nulos ao realizar uma descoberta do domínio de dados em uma fonte de dados. Quando você seleciona a porcentagem mínima de linhas com a opção de exclusão de valores nulos, a porcentagem de conformidade é a proporção do número de linhas correspondentes dividido pelo número total de linhas, menos os valores nulos na coluna.

O processo de descoberta do domínio de dados é diferente quando você escolhe a opção **Excluir valores nulos da descoberta do domínio de dados** e os vários filtros e opções de amostragem.

Os seguintes cenários explicam os resultados de uma descoberta do domínio de dados quando você escolhe a opção para excluir valores nulos junto com uma opção de amostragem e filtros:

- Com **Todas as linhas** como opção de amostragem e sem filtros. A descoberta do domínio de dados ignora todos os valores nulos na coluna.
- Com uma opção de amostragem e sem filtros. A descoberta do domínio de dados ignora todos os valores nulos nos dados amostrados e executa o restante desses dados.
- Com **Todas as linhas** como opção de amostragem e com filtros. A descoberta do domínio de dados ignora todos os valores nulos nos dados filtrados e executa o restante desses dados.
- Com uma opção de amostragem e filtros. A descoberta do domínio de dados ignora valores nulos nos dados filtrados da amostra e executa o restante desses dados.

Exemplo

Você tem uma fonte de dados com 10.000 linhas em que 3.000 linhas têm Números de Previdência Social na coluna Comentários. Você cria um perfil de coluna com descoberta do domínio de dados e escolhe as seguintes opções:

- Selecione a opção **Excluir valores nulos da descoberta do domínio de dados**.
- Selecione **Todas as linhas** como opção de amostragem.
- Selecione a opção **Porcentagem mínima de linhas** e configure a opção como 12%.

Quando você executa o perfil, ele é executado no conjunto de dados e ignora os valores nulos durante a descoberta do domínio de dados.

Criando um Perfil para Executar Descoberta do Domínio de Dados no Informatica Analyst

Você precisa criar pelo menos um domínio de dados antes de poder criar um perfil para executar descoberta do domínio de dados na ferramenta Analyst. O perfil pode descobrir o nome da coluna e dados da coluna que correspondam a domínios de dados predefinidos.

1. No espaço de trabalho **Descoberta**, clique em **Perfil** ou selecione **Novo** > **Perfil** em qualquer lugar na ferramenta Analyst.

   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**.
5. 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 descoberta de domínio de dados** para realizar a descoberta de domínio de dados. Selecione as opções de domínio de dados no painel **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 para o perfil de coluna também são aplicáveis para a descoberta de domínio de dados. Clique em **Editar** para marcar ou desmarcar as colunas para a descoberta de domínio de dados independentemente das colunas que você selecionar para o perfil de coluna.

- Escolha dados, colunas ou dados e colunas para executar a descoberta de domínio de dados.
- Escolha uma opção de amostragem no painel **Executar perfil em**.
- Escolha uma opção de 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 tipo de dados ou domínio de dados aprovado.
- Escolha critérios de conformidade, e você poderá selecionar a opção **Excluir valores nulos da descoberta do domínio de dados**.
- Escolha **Nativo**, **Hive** ou **Hadoop** como ambiente de tempo de execução. Se você escolher a opção Hive ou Hadoop, clique em **Escolher** para selecionar uma conexão do Hadoop na caixa de diálogo **Selecionar uma Conexão do Hadoop**.

6. Na tela **Especificar Regras e Filtros**, você pode adicionar, editar ou excluir regras e filtros para o perfil.
7. Clique em **Salvar e Concluir** para criar o perfil ou clique em **Salvar e Executar** para criar e executar o perfil.

Editando um Perfil no Informatica Analyst

Você pode alterar as propriedades de um perfil depois de executá-lo. Se você executou um perfil de coluna como parte da descoberta do domínio de dados, poderá alterar as configurações de perfil de coluna.

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. Edite as opções de domínio de dados, a opção de amostragem e as opção busca detalhada.
   - **Especificar Regras e Filtros**. Crie, edite ou exclua regras e filtros.
6. Clique em **Salvar e Concluir** para editar o 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.

Executando um Perfil para Realizar a Descoberta do Domínio de Dados

Execute um perfil como parte da descoberta do domínio de dados para exibir as colunas que correspondem aos padrões da regra de domínio de dados.

1. No **Navegador da 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 da coluna.

   A exibição detalhada é exibida.

Resultados da Descoberta do Domínio de Dados no Informatica Analyst

Você pode visualizar resultados de descobertas de domínio de dados nas exibições resumida e detalhada.

O campo de domínio de dados mostra estatísticas sobre colunas que correspondem a domínios de dados. Na exibição resumida, é possível exibir os domínios de dados inferidos junto com a porcentagem e o número de linhas em conformidade.

Na exibição detalhada, é possível realizar as seguintes tarefas:

- Exibir os domínios de dados inferidos com a porcentagem e o número de linhas em conformidade em um gráfico de barras horizontal.
- Fazer uma busca detalhada dos resultados até linhas em conformidade, linhas fora de conformidade e valores nulos.
- Aprovar, rejeitar ou redefinir o domínio de dados.
- Mostrar ou ocultar domínios de dados rejeitados.
- Executar a descoberta de domínio de dados em todas as linhas da fonte de dados para descobrir o domínio de dados inferido.

Aprovando Domínios de Dados

Você pode aprovar vários domínios de dados na ferramenta Analyst.

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. Clique na coluna cujo domínio de dados você deseja aprovar.

   Os resultados de coluna aparecem na exibição detalhada.
4. Na exibição detalhada, selecione o domínio de dados no painel **Domínio de Dados**. Clique em **Ações** > **Aprovar**.

   O status da coluna ou do domínio de dados muda para Aprovado.
5. Para restaurar o status inferido da coluna ou do domínio de dados, selecione este último e clique em **Ações** > **Redefinir**.

Rejeitando Domínios de Dados

Quando você abre os resultados do perfil, a ferramenta Analyst exibe domínios de dados aprovados por padrão. Você pode exibir ou ocultar os domínios de dados rejeitados.

1. No Navegador de 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. Clique na coluna cujo domínio de dados você deseja rejeitar.

   Os resultados de coluna aparecem na exibição detalhada.
4. Para rejeitar os domínios de dados inferidos, clique em **Ações** > **Rejeitar**.

   A ferramenta Analyst remove o domínio de dados rejeitado dos resultados da descoberta do domínio de dados.
5. Para exibir os domínios de dados rejeitados, clique em **Ações** > **Mostrar Rejeitados**.
6. Para ocultar os domínios de dados rejeitados, clique em **Ações** > **Ocultar Rejeitados**.

Arquivos de Exportação de Descoberta do Domínio de Dados no Informatica Analyst

Quando você exporta resultados da descoberta do domínio de dados na ferramenta Analyst, pode especificar o nome do arquivo e o valor da página de código. Você pode exportar os resultados da descoberta do domínio de dados para um arquivo do Microsoft Excel.

O arquivo do Microsoft Excel contém várias planilhas que separam os resultados da descoberta com base nas colunas, domínios de dados e grupos de domínio de dados. A planilha Propriedades exibe propriedades do perfil, como nome, descrição, tipo, localização, data e hora em que as últimas alterações foram feitas no perfil e um link para o perfil.

Resultados da Descoberta do Domínio de Dados no Microsoft Excel

Quando você exporta os resultados da descoberta do domínio de dados para o Microsoft Excel, a ferramenta Analyst salva os nomes de colunas, os nomes de domínios de dados correspondentes, os critérios de conformidade e os valores nulos. O arquivo do Excel também contém os nomes dos grupos de domínios de dados para cada domínio de dados e tipos de dados documentados para as colunas.

A tabela a seguir descreve cada planilha no arquivo de exportação:

Guia	Descrição
Exibir por Colunas	Resultados da descoberta do domínio de dados classificados por coluna de fonte de dados.
Exibir por Domínios de Dados	Resultados da descoberta do domínio de dados classificados por domínio de dados.
Exibir por Grupos de Domínios de Dados	Resultados da descoberta do domínio de dados classificados por grupo do domínio de dados.
Propriedades	Propriedades de perfil básico, como nome, descrição, tipo, localização, data e hora em que as últimas alterações foram feitas no perfil e um link para o perfil.

Exportando Resultados da Descoberta do Domínio de Dados do Informatica Analyst

Você pode exportar os resultados da descoberta do domínio de dados para um arquivo `.xlsx` para poder ver os dados em um arquivo e distribuí-lo na empresa para uso posterior.

1. Execute um perfil para realizar a descoberta do domínio de dados.
2. Na exibição resumida ou detalhada, clique em **Ações** > **Exportar Dados**

   A caixa de diálogo **Exportar dados para um arquivo** é exibida.
3. Digite o nome do arquivo. Como alternativa, use o nome de arquivo padrão.
4. Selecione a página de código do arquivo
5. Clique em **OK**.

CAPÍTULO 13

Descoberta Empresarial no Informatica Analyst

Este capítulo inclui os seguintes tópicos:

- Descoberta Empresarial no Informatica Analyst - Visão Geral, 115
- Processo de descoberta empresarial no Informatica Analyst, 116
- Opções de Configuração para Descoberta Empresarial, 116
- Criando um perfil de descoberta empresarial no Informatica Analyst, 118
- Editando Opções de Descoberta Empresarial, 119

Descoberta Empresarial no Informatica Analyst - Visão Geral

A descoberta empresarial é o processo de descobrir os metadados da coluna e os domínios de dados em várias fontes de dados de vários esquemas e conexões relacionais externas. Você pode realizar a descoberta empresarial nas fontes de dados que importou para o repositório do Modelo e nas fontes de dados de conexões relacionais externas.

Como um analista de dados, você pode realizar a descoberta empresarial na ferramenta Analyst para inferir características de metadados específicas em um grande número de fontes de dados. Você também pode exibir os dados de origem que correspondem aos domínios de dados predefinidos. Você pode fazer curadoria dos resultados inferidos da descoberta empresarial e tornar os dados prontos para as iniciativas de pesquisa de descoberta e de qualidade de dados. A descoberta empresarial na ferramenta Analyst gera um resumo consolidado dos resultados do perfil.

Os resultados da descoberta empresarial incluem estatísticas de perfil de coluna, como padrões, valores exclusivos e colunas com conflitos de tipo de dados. A descoberta do domínio de dados identifica as colunas de origem que correspondem aos domínios de dados predefinidos.

Você pode escolher um perfil do sistema operacional no Informatica Analyst. Depois que você escolhe um perfil do sistema operacional, o Serviço de Integração de Dados cria e executa os perfis de descoberta empresarial com base na permissão do usuário do sistema operacional que você define no perfil do sistema operacional.

Processo de descoberta empresarial no Informatica Analyst

Você pode criar, editar e excluir perfis de descoberta empresarial. Você pode executar um perfil de descoberta empresarial no espaço de trabalho de Descoberta. Você precisa configurar as opções de inferência para o perfil da coluna e a descoberta do domínio de dados antes de executar o perfil de descoberta empresarial.

Conclua as seguintes etapas para executar a descoberta empresarial na ferramenta Analyst:

1. Configure as propriedades gerais do perfil de descoberta empresarial.
2. Selecione os objetos de dados do repositório do Modelo que você deseja incluir no perfil de descoberta empresarial.
3. Importe as fontes de dados relacionais das conexões externas de banco de dados.
4. Configure as opções de inferência de dados e as opções de descoberta do perfil de descoberta empresarial.
5. Salve as alterações e execute o perfil de descoberta empresarial.
6. Monitore a execução do perfil e, se necessário, exiba os status das tarefas de perfil executadas pela ferramenta Analyst.
7. Examine o resumo de resultados da descoberta empresarial. Os resultados aparecem nos painéis **Resumo** e **Perfis**.

Opções de Configuração para Descoberta Empresarial

As opções de configuração para descoberta empresarial incluem opções de descoberta do domínio de dados, opções de amostragem de perfil de coluna e propriedades gerais de perfil, como nome e descrição.

Você pode optar por executar um perfil de coluna ou um perfil para executar a descoberta do domínio de dados. Você pode também optar por executar o perfil de coluna e um perfil para executar a descoberta do domínio de dados como parte da configuração.

Configurações da Descoberta do Domínio de Dados

As configurações de descoberta do domínio de dados incluem a opção de execução da descoberta do domínio de dados nos dados da coluna, no nome de coluna ou nos dados da coluna e nome de coluna. Você pode escolher domínios de dados e especificar se a descoberta do domínio de dados precisa processar

todas as linhas na fonte de dados. Você pode escolher critérios de conformidade para a descoberta do domínio de dados. Você pode excluir valores nulos da descoberta do domínio de dados.

A seguinte tabela descreve as configurações de descoberta do domínio de dados que você pode definir para a descoberta empresarial na ferramenta Analyst:

Opção	Descrição
Habilitar descoberta do domínio de dados	Execute a descoberta do domínio de dados como parte da descoberta empresarial.
Executar descoberta do domínio de dados nos dados	Execute a descoberta do domínio de dados nos dados da coluna.
Executar descoberta do domínio de dados no nome da coluna	Execute a descoberta do domínio de dados no nome de cada coluna.
Porcentagem mínima de conformidade	A porcentagem mínima de conformidade de linhas no conjunto de dados necessária para uma correspondência de domínio de dados. A porcentagem de conformidade é a proporção entre o número de linhas correspondentes dividido pelo número total de linhas. **Nota:** A ferramenta Analyst considera valores nulos como linhas sem correspondência.
Mínimo de linhas em conformidade	O número mínimo de linhas no conjunto de dados necessárias para uma correspondência de domínio de dados.
Excluir valores nulos da descoberta do domínio de dados	Exclui os valores nulos do conjunto de dados para a descoberta do domínio de dados.
Excluir colunas com domínios de dados aprovados	Exclui as colunas com domínios de dados aprovados da inferência de domínio de dados da execução de perfil.
Todas as linhas	Execute a descoberta do domínio de dados em todas as linhas de origem.
Primeiro	O número máximo de linhas nas quais o perfil pode ser executado. A ferramenta Analyst escolhe as linhas começando da primeira linha da origem.

Configurações de Perfil de Coluna

As opções de amostragem determinam se a ferramenta Analyst executa um perfil de coluna em todas as linhas das fontes de dados ou um número limitado de linhas.

A seguinte tabela descreve as configurações de perfil de coluna que você pode definir para um perfil de descoberta empresarial:

Opção	Descrição
Ativar criação de perfil de coluna	Executa um perfil de coluna como parte da descoberta empresarial.
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.

A seguinte tabela descreve as opções de amostragem que você pode definir para um perfil de descoberta empresarial:

Opção	Descrição
Todas as Linhas	Executa um perfil de coluna em todas as linhas da fonte de dados.
Primeiras <número> Linhas	O número de linhas nas quais você deseja executar o perfil de coluna. A ferramenta Analyst escolhe as linhas começando pela primeira na fonte de dados.

A seguinte tabela descreve as opções de ambiente de tempo de execução que você pode configurar para um perfil de descoberta empresarial:

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.
Hadoop	O Serviço de Integração de Dados envia a lógica de perfil por push para o mecanismo Blaze no cluster Hadoop para executar perfis.

Criando um perfil de descoberta empresarial no Informatica Analyst

Você pode executar o perfil de coluna e a descoberta do domínio de dados como parte da descoberta empresarial na ferramenta Analyst.

1. No espaço de trabalho **Descoberta**, selecione **Novo** > **Perfil**.

   O assistente de **Novo Perfil** é exibido.
2. Selecione **Descoberta Empresarial**. Clique em **Avançar**.

   A guia **Especificar Propriedades Gerais** é exibida.
3. Na guia **Especificar Propriedades Gerais**, insira um nome para o perfil de descoberta empresarial e uma descrição opcional. No campo Localização, selecione o projeto ou a pasta na qual você deseja criar o perfil. Clique em **Avançar**.

   A guia **Selecionar Objetos de Dados** é exibida.
4. Na guia **Selecionar Objeto de Dados**, clique em **Escolher**.

   A caixa de diálogo **Escolher Objetos de Dados** é exibida.
5. Na caixa de diálogo **Escolher Objetos de Dados**, escolha um ou mais objetos de dados a serem adicionados ao perfil. Clique em **Salvar**.

   Os objetos de dados aparecem no painel **Objetos de Dados**.
6. Clique em **Avançar**.

   A guia **Selecionar Recursos** é exibida.
7. Na guia **Selecionar Recursos**, clique em **Escolher** para abrir a guia **Selecionar Recursos**.

Você pode importar dados de várias fontes de dados relacionais.

8. Na guia **Selecionar Recursos**, selecione as conexões, os esquemas, as tabelas e as exibições que você deseja incluir no perfil. Clique em **Salvar**.

   O painel esquerdo na caixa de diálogo lista todas as conexões internas e externas, esquemas, tabelas e exibições sob o domínio Informatica.

   Os recursos aparecem no painel **Recursos**.

9. Clique em **Avançar**.

   A guia **Especificar Configurações** é exibida.

10. Na guia **Especificar Configurações**, você pode configurar as opções de perfil de coluna e as opções de descoberta do domínio de dados. Clique em **Salvar e Concluir** para salvar o perfil de descoberta empresarial ou clique em **Salvar e Executar** para executar o perfil.

    É possível realizar as seguintes tarefas na guia **Especificar Configurações**.

    - Ativar descoberta de domínio de dados. Clique em **Escolher** para selecionar os domínios de dados que você deseja descobrir na caixa de diálogo **Escolher Domínios de Dados**. Os domínios de dados selecionados aparecem no painel **Domínios de Dados para Descoberta de Domínio de Dados**.
    - Executar o domínio de dados nos dados, no nome da coluna ou em ambos.
    - Selecionar todas as linhas da fonte de dados ou escolher um número máximo de linhas na qual executar a descoberta de domínio.
    - Escolher uma porcentagem mínima de conformidade ou especificar o número mínimo de linhas em conformidade para a descoberta do domínio de dados.
    - Ativar configurações de perfil de coluna e selecionar todas as linhas ou as primeiras filas na fonte de dados para o perfil de coluna. Você pode excluir a inferência de tipo de dados para colunas com tipos de dados aprovados no perfil de coluna.
    - Escolha **Nativo** ou **Hadoop** como ambiente de tempo de execução.

    É possível exibir os resultados da descoberta empresarial nas guias **Resumo** e **Perfis**.

Editando Opções de Descoberta Empresarial

Você pode fazer alterações nas opções de descoberta empresarial depois de executar a descoberta empresarial. Você pode renomear o perfil e alterar a seleção de objeto de dados, a seleção de domínio de dados e as opções de inferência.

1. Abra um perfil que você executou para executar a descoberta empresarial.

   Os resultados do perfil são exibidos no espaço de trabalho de **Descoberta** .

2. Se o sistema de controle de versão estiver ativado, clique em **Ações** > **Fazer Check-Out** para fazer check-out do perfil.
3. Clique em **Editar Perfil**.
4. Na guia **Especificar Propriedades Gerais**, atualize as propriedades de perfil conforme necessário.
5. Para fazer alterações na seleção de objetos de dados, clique na guia **Selecionar Objetos de Dados**.
6. Para fazer alterações nas fontes de dados externas da descoberta empresarial, clique na guia **Selecionar Recursos**.
7. Para fazer alterações nas opções de inferência de domínio de dados e nas configurações de perfil de coluna, clique na guia **Especificar Configurações**.

8. Para aplicar as alterações de configuração a todas as tarefas de perfil do domínio de dados e as tarefas de perfil de coluna no perfil de descoberta empresarial, selecione a opção **Usar configurações globais para todos os perfis**. Se você não selecionar essa opção, as alterações feitas nas configurações de perfil só se aplicarão aos objetos de dados ou recursos que você tiver adicionado recentemente ao perfil.

   Por padrão, as alterações feitas por você são aplicadas aos objetos de dados recém-adicionados no perfil de descoberta de dados.
9. Para desfazer as alterações, clique em **Cancelar**.
10. Clique em **Salvar e Executar** para salvar as alterações e executar o perfil novamente.
11. 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.

CAPÍTULO 14

Resultados da Descoberta Empresarial no Informatica Analyst

Este capítulo inclui os seguintes tópicos:

- Visão geral dos resultados da descoberta empresarial no Informatica Analyst, 121
- Exibição Resumida, 121
- Conflito de Tipo de Dados, 124
- Exibição Perfis, 125

Visão geral dos resultados da descoberta empresarial no Informatica Analyst

Você pode exibir os resultados da descoberta empresarial nas exibições **Resumo** e **Perfis**.

A exibição **Resumo** exibe os resultados de perfil da coluna e os resultados da descoberta do domínio de dados. A seção **Descoberta do Domínio de Dados** lista os domínios de dados incluídos na execução de perfil e o número de colunas que têm uma correspondência de domínio de dados. A seção **Criação de Perfil de Coluna** exibe estatísticas sobre as colunas de origem. Você pode clicar em cada linha de resultados de perfil para exibir os detalhes no painel direito da exibição **Resumo**.

Exibição Resumida

A exibição **Resumo** exibe um resumo dos resultados de perfil da coluna e os resultados da descoberta do domínio de dados. Você pode exibir os nomes de domínios de dados que têm correspondências em colunas e número de colunas com correspondências de domínio de dados. As estatísticas de coluna incluem o número de colunas para as primeiras dez correspondências de padrão, todos os valores exclusivos e todos

os valores nulos. As estatísticas de coluna também incluem o número de colunas que apresentam conflitos de tipo de dados entre tipos de dados inferidos e tipos de dados documentados.

Resultados de Perfil da Exibição Resumo

A exibição Resumo exibe os resultados da descoberta empresarial nas seções Descoberta do Domínio de Dados e Criação de Perfil de Coluna.

Descoberta do Domínio de Dados

A seguinte tabela descreve as colunas nos resultados de descoberta de domínios de dados:

Nome da coluna	Descrição
Nome	Nome do domínio de dados.
Encontrado em colunas	O número total de colunas com uma correspondência de domínio de dados.
Perfil	Nome do perfil que contém a coluna correspondente.
Nome da coluna	Nome da coluna correspondente.
Porcentagem de conformidade de dados	A porcentagem mínima de conformidade de linhas necessária para uma correspondência de domínio de dados.
Nome da conexão	Nome da conexão de banco de dados relacional.
Nome da origem	Nome da fonte de dados.
Status de inferência	Status de inferência de domínio de dados. Os status são **Aceito**, **Rejeitado** e **Inferido**.
% Nulo	A porcentagem de valores nulos para a coluna.
Total de linhas	O número total de linhas.
Linhas em conformidade	O número mínimo de linhas necessárias para uma correspondência de domínio de dados.
Correspondência de nome de coluna	Indica se o nome da coluna corresponde a um nome de domínio de dados.
Tipo de dados documentados	Tipo de dados declarados para a coluna no objeto de perfil.
Verificado	Indica a validação da correspondência de domínio de dados em todas as linhas da fonte de dados.
Horário da última execução	Data e hora da última execução de perfil.

Criação de Perfil de Coluna

A tabela a seguir descreve as colunas nos resultados de perfil da coluna:

Nome da coluna	Descrição
Nome	Nome do tipo de resultados do perfil, como padrão, 100% nulos e 100% únicos.
Encontrado em colunas	Número total de colunas que têm a correspondência de tipo de resultado do perfil.
Perfil	Nome do perfil que contém a coluna correspondente.
Conexão	Nome da conexão de banco de dados relacional.
Fonte de dados	Fonte de dados do perfil.
Número de colunas	Número de colunas no perfil que têm a correspondência de tipo de resultado do perfil.

Exibindo Resultados da Descoberta do Domínio de Dados

Você pode clicar em um nome de domínio de dados para exibir os respectivos resultados da descoberta do domínio de dados. Você pode abrir perfis específicos nos resultados da descoberta do domínio de dados.

1. Execute um perfil para realizar a descoberta empresarial.
2. Verifique se você está na exibição **Resumo**.
3. Clique em um domínio de dados na seção **Descoberta do Domínio de Dados** para exibir os respectivos resultados de descoberta.

   Uma lista de perfis que contêm o domínio de dados é exibida no painel direito.
4. Selecione uma linha no painel direito, se necessário.

   Os hiperlinks para o perfil aparecem na cor Azul.
5. Clique no link de nome do perfil ou nome de coluna para abrir o perfil.

   O perfil abre e exibe os resultados da descoberta do domínio de dados. A ferramenta Analyst destaca a linha com o domínio de dados nos resultados. Se necessário, você pode fazer curadoria dos resultados do perfil para fazer uso mais eficaz, como a Pesquisa de Descoberta.
6. Para retornar à exibição **Resumida**, clique em **Retornar à Descoberta Empresarial**.

Exibindo Resultados do Perfil de Coluna

Você pode exibir os resultados do perfil da coluna da descoberta empresarial na exibição **Resumo**. Você pode abrir perfis específicos nos resultados da descoberta do domínio de dados.

1. Execute um perfil para realizar a descoberta empresarial.
2. Verifique se você está na exibição **Resumo**.
3. Para exibir os detalhes dos padrões inferidos, clique em um dos dez principais padrões na seção **Criação de Perfil de Coluna**.

   Uma lista de perfis que contêm os resultados de padrão inferidos é exibida no painel direito.

4. Para exibir informações, como todos os valores nulos, todos os valores exclusivos ou os conflitos de tipos de dados, clique em **100% de Nulos**, **100% de Exclusivos** ou **Tipos de dados inferidos versus documentados conflitantes**.

   A lista de correspondência de perfis é exibida no painel direito.

5. Clique no link de nome do perfil ou nome de coluna para abrir o perfil.

   O perfil abre e exibe os resultados do perfil da coluna.

6. Para retornar à exibição **Resumida**, clique em **Retornar à Descoberta Empresarial**.

Conflito de Tipo de Dados

A descoberta empresarial identifica conflitos de tipo de dados em colunas. Um conflito de tipo de dados é uma incompatibilidade de tipos de dados inferidos e documentados de uma coluna após a execução da descoberta empresarial. Um tipo de dados inferido é o tipo de dados que a ferramenta Analyst deriva para uma coluna de fonte de dados com base nos dados dessa coluna. Um tipo de dados documentado é o tipo de dados declarado para uma coluna no banco de dados de origem.

A descoberta empresarial pode inferir um tipo de dados diferente para uma coluna com base nos dados da coluna em comparação com o tipo de dados documentado de uma coluna. Por exemplo, a descoberta empresarial pode inferir uma coluna com um tipo de dados de cadeia documentado como um tipo de dados de data. Você pode revisar o conflito de tipo de dados, escolher o tipo de dados de data mais apropriado para a coluna e aprová-lo.

Exibindo Conflitos de Tipos de Dados

Quando você abre um perfil com conflitos de tipos de dados na exibição **Resumo**, a ferramenta Analyst destaca esses conflitos em Vermelho.

1. Execute um perfil para realizar a descoberta empresarial.
2. Verifique se você está na exibição **Resumo**.
3. Na seção **Criação de Perfil de Coluna**, clique em **Tipos de dados inferidos versus documentados conflitantes** para exibir os conflitos de tipos de dados nos resultados do perfil de coluna.

   Uma lista de perfis que contém colunas com conflitos de tipos de dados aparece no painel direito.

4. Selecione uma linha no painel direito, se necessário.

   Os hiperlinks para o perfil aparecem na cor Azul.

5. Clique no link de nome do perfil ou nome de coluna para abrir o perfil.

   O perfil se abre e exibe os conflitos de tipos de dados na cor Vermelha. Você pode optar por obter o controle dos tipos de dados inferidos para resolver o conflito de tipos de dados.

6. Para obter o controle dos tipos de dados, selecione uma linha com os tipos de dados conflitantes e clique na exibição **tipos de dados**.
7. Clique em **Ações** e selecione **Aprovar** ou **Rejeitar**.
8. Para retornar à exibição **Resumida**, clique em **Retornar à Descoberta Empresarial**.

Exibição Perfis

A exibição **Perfis** mostra uma lista de todos os perfis de objeto de dados únicos que a ferramenta Analyst executa como parte da descoberta empresarial. O perfil lista também exibe o status de execução de cada perfil. Você pode abrir cada perfil para exibir os resultados do perfil da coluna e os resultados da descoberta do domínio de dados.

Exibindo Propriedades do Perfil

Você pode exibir a lista de perfis que fazem parte da descoberta empresarial na exibição **Perfis**. Você pode abrir cada perfil e fazer curadoria dos resultados do perfil, se necessário.

1. Execute um perfil para realizar a descoberta empresarial.
2. Verifique se você está na exibição **Perfis**.
3. Para exibir as propriedades de um perfil, clique no nome do perfil.

   As propriedades do perfil aparecem no painel direito. As propriedades do perfil incluem o nome dos objetos de dados de origem, o nome da conexão e a contagem de linhas.
4. Para exibir os resultados do perfil, clique em **Abrir Perfil**.

   O perfil exibe os resultados do perfil da coluna.
5. Para reverter para a exibição **Perfis**, clique no link do nome de pasta ou de projeto no canto superior esquerdo do espaço de trabalho de **Descobertas**.

CAPÍTULO 15

Pesquisa de Descoberta no Informatica Analyst

Este capítulo inclui os seguintes tópicos:

- Pesquisa de Descoberta no Informatica Analyst - Visão Geral, 126
- Pré-requisitos de Pesquisa de Descoberta, 127
- Processo de Pesquisa de Descoberta no Informatica Analyst, 127
- Opções de Pesquisa de Descoberta, 128
- Resultados da Pesquisa de Descoberta no Informatica Analyst, 129
- Tipos de Correspondência, 131
- Ativos Relacionados, 133
- Perguntas Frequentes, 134

Pesquisa de Descoberta no Informatica Analyst - Visão Geral

A pesquisa de descoberta localiza ativos e identifica relacionamentos com outros ativos nos bancos de dados e esquemas da empresa. Usuários empresariais podem usar a pesquisa de descoberta para encontrar onde os dados e os metadados existem na empresa. Você pode procurar ativos específicos, como objetos de dados, regras e perfis.

Se você executar uma pesquisa global, a ferramenta Analyst executará uma pesquisa baseada em texto por objetos de dados, fontes de dados e pastas. Se você executar a pesquisa de descoberta, além das correspondências de texto, os resultados da pesquisa incluem objetos relacionados com os objetos que correspondem aos critérios de pesquisa. A pesquisa de descoberta também inclui correspondências com base nos metadados do perfil, como tipos de dados e padrões de dados. Por exemplo, você pode localizar objetos que contêm um determinado padrão de dados e têm nomes que contêm uma palavra-chave específica.

A pesquisa de descoberta inclui os seguintes tipos de informações nos resultados da pesquisa:

Objetos no repositório do Modelo

Localiza objetos primários relacionados aos objetos que correspondem aos critérios de pesquisa de descoberta. Por exemplo, quando você pesquisa um perfil, os resultados de perfil incluem o objeto de dados do perfil.

Resultados do depósito de perfil

Inclui resultados de inferência de perfis, como um domínio de dados ou um padrão de dados.

Termos do Business Glossary

Com base na licença, inclui metadados na pesquisa, como um termo comercial associado a uma regra.

Exemplo de Pesquisa de Descoberta

Você é um administrador de dados na empresa, responsável por garantir que dados empresariais confidenciais sejam mascarados de maneira adequada. Talvez você deseje identificar Informações de Identificação Pessoal (PII) em todos os esquemas e bancos de dados no qual você ou o arquiteto de dados executou a descoberta empresarial. Você pode ter criado domínios de dados para identificar dados importantes que permanecem não descobertos em fontes de dados. Execute a pesquisa na string "SSN". A ferramenta Analyst exibe o domínio de dados da Previdência Social domínio e todas as colunas correspondentes das fontes de dados. Além disso, a string de pesquisa de descoberta pode localizar colunas ou tabelas adicionais que têm "SSN" em suas descrições ou nomes. Para restringir a pesquisa, você pode filtrar as especificações de mapeamento para mostrar as especificações de mapeamento que fazem referência a objetos de dados correspondentes. Você pode aplicar filtros adicionais para filtrar as especificações de mapeamento com base em projetos ou usuários. Talvez você deseje abrir as especificações de mapeamento nos resultados para verificar se as especificações de mapeamento atendem às diretivas de privacidade da empresa.

Pré-requisitos de Pesquisa de Descoberta

Antes que seja possível realizar uma pesquisa eficiente nos bancos de dados da empresa, execute a descoberta empresarial nos bancos de dados e nos esquemas da empresa.

Depois de executar a descoberta empresarial, a ferramenta Analyst armazena todos os resultados de perfil no depósito de criação de perfil. Verifique se as fontes de dados estão no repositório do Modelo. Como alternativa, verifique se os ativos adequados no repositório do Modelo têm termos comerciais associados aos ativos. Quando você executa a pesquisa de descoberta, o Serviço de Pesquisa recupera as informações do índice de pesquisa com base nos resultados dos ativos do repositório do Modelo e do depósito de criação de perfil. Em seguida, o Serviço de Pesquisa usa as informações indexadas para exibir os resultados da pesquisa com base nos metadados de objeto e nos relacionamentos adequados.

Processo de Pesquisa de Descoberta no Informatica Analyst

Você pode procurar ativos com base em critérios, como texto, padrões e tipos de dados nos resultados do perfil. A pesquisa retornará uma lista de ativos relacionados à string de pesquisa.

Conclua as seguintes etapas para executar a pesquisa de descoberta na ferramenta Analyst:

1. Execute a descoberta empresarial e execute os perfis de objeto de dados únicos necessários nas fontes de dados da empresa. Quando você executa a pesquisa de descoberta, a ferramenta Analyst procura informações nos resultados do perfil e objetos de repositório do Modelo.

2. Escolha o tipo de informações que você deseja localizar. Por exemplo, talvez você deseje localizar todos os ativos associados a uma definição de domínio de dados de dados confidenciais ou um padrão específico de dados.
3. Execute a pesquisa.
4. Analise os resultados da pesquisa para identificar os ativos e seus relacionamentos com outros ativos.
5. Se necessário, verifique se os dados descobertos são compatíveis com os requisitos comerciais.

Opções de Pesquisa de Descoberta

Você pode realizar uma pesquisa global ou uma pesquisa de descoberta para localizar ativos e identificar seus relacionamentos com outros ativos. A pesquisa global recupera os resultados do repositório do Modelo e, como alternativa, do Business Glossary. A pesquisa de descoberta recupera os resultados do repositório do Modelo e do Business Glossary, além de perfis com base nos resultados do perfil no depósito de criação de perfil.

Você pode pesquisar ativos, como objetos de dados, perfis e especificações de mapeamento. Digite uma string de pesquisa para pesquisar ativos que correspondem à string de pesquisa e têm uma associação com a string de pesquisa. Você pode usar caracteres curinga quando pesquisar ativos.

Você pode usar os seguintes caracteres curinga quando pesquisar ativos:

*** (Asterisco)**

Adicione ao final da string de pesquisa para localizar todos os nomes de ativos que começam com a string de pesquisa. Por exemplo, para pesquisar todos os nomes de ativos que começam com a string "emp", você pode digitar "emp*" no campo de pesquisa.

? (Ponto de Interrogação)

Inclua na string de pesquisa para representar um caractere alfanumérico.

Nota: Você não pode iniciar uma string de pesquisa com um caractere curinga ao pesquisar ativos. A pesquisa não diferencia maiúsculas e minúsculas.

Para pesquisar duas ou mais palavras juntas como uma frase, inclua as palavras em aspas duplas. Use o caractere `+` para representar o operador AND e pesquisar um termo que deve aparecer nos resultados da pesquisa. Por exemplo, se a string de pesquisa for `+dados +confidenciais`, o Serviço de Pesquisa localizará os metadados que incluam ambos os termos. Use um espaço em branco para representar o operador OR. Por exemplo, se a string de pesquisa for `dados confidenciais`, o Serviço de Pesquisa localiza os metadados que contenham qualquer um dos termos.

Se a string de pesquisa contiver hífen (-), sublinhado (_) ou concatenação com maiúsculas e minúsculas, o Serviço de Pesquisa localiza a palavra inteira e parte de palavras separados pelo delimitador. Por exemplo, se você pesquisar `Profile_Customer`, o mecanismo de pesquisa localizará `Perfil`, `Cliente` e `Profile_Customer` nos repositórios. Para incluir caracteres especiais, como `*` e `?`, na string de pesquisa, inclua a string de pesquisa que contém caracteres especiais em aspas duplas.

Você pode realizar uma pesquisa de descoberta que inclua uma pesquisa de palavra-chave e um filtro de descoberta. Por exemplo, talvez você deseje localizar as colunas de ID de funcionário que usam o formato `<InicialPrimeiroNome><InicialSobrenome>-<NumeroPrevidenciaSocial>` para que seja possível identificar riscos de segurança de dados. Para procurar as colunas de ID de funcionário, insira `ID do Funcionário` no painel **Pesquisar** do espaço de trabalho Biblioteca e defina o filtro de padrão como `XX-999999999 <= 100%`.

Critérios de Pesquisa de Descoberta

Use os critérios de pesquisa de descoberta para procurar informações com base em critérios, como padrões, tipos de dados, valores exclusivos e valores nulos. Você pode usar os operadores condicionais =, >= ou <=na pesquisa.

A tabela a seguir descreve os critérios de pesquisa de descoberta que você pode usar na pesquisa de descoberta:

Opção	Descrição
Pesquisa	Expressão de texto que você deseja procurar.
Limpar	Limpa a cadeia de pesquisa e todos os outros critérios de pesquisa que você tiver selecionado.
Padrão de	O padrão de coluna e a porcentagem que você deseja incluir na pesquisa. **Nota:** A opção não aceita caracteres de controle em um padrão.
Tipo de dados de	O tipo de dados da coluna e a porcentagem que você deseja incluir na pesquisa.
Valores exclusivos	Porcentagem de valores exclusivos em colunas que você deseja incluir na pesquisa.
Nulos	Porcentagem de valores nulos nas colunas que você deseja incluir na pesquisa.

Pesquisando um Ativo

Você pode pesquisar um ativo no espaço de trabalho **Biblioteca**. Os resultados da pesquisa incluem ativos criados na Developer tool e na ferramenta Analyst.

1. Abra o espaço de trabalho **Biblioteca**.
2. Verifique se você está na seção **Pesquisa de Descoberta**.
3. No campo **Pesquisa**, digite a string de pesquisa que você deseja pesquisar.
4. Configure os filtros de pesquisa para restringir a pesquisa.

   Os filtros incluem padrões, tipos de dados, valores exclusivos e valores nulos.
5. Clique no ícone **Filtro**.

Resultados da Pesquisa de Descoberta no Informatica Analyst

A pesquisa de descoberta localiza ativos em todos os repositórios licenciados para pesquisa de descoberta, como o repositório do Modelo e o depósito de criação de perfil.

Os resultados da pesquisa de descoberta incluem o número total de correspondências e a lista de correspondências. Você pode expandir cada correspondência para exibir as propriedades da correspondência, informações da correspondência direta, informações da correspondência indireta e o número total de ativos relacionados, se houver. Uma correspondência direta é uma correspondência com alguns ou todos os metadados do ativo que corresponde à consulta de pesquisa. Uma correspondência indireta é uma correspondência de ativo que é vinculada ao ativo que corresponde diretamente à consulta de pesquisa.

A ordem dos resultados da pesquisa que você vê depende dos seguintes fatores:

- A propriedade do objeto que corresponde aos critérios de pesquisa. O nome do objeto tem uma prioridade mais alta do que a descrição do objeto. A descrição do objeto tem uma prioridade mais alta do que outras propriedades do objeto.
- Tipo de objeto. Os domínios de dados e os grupos de domínio de dados têm uma prioridade menor do que outros objetos.
- Curadoria. Os resultados de perfil com curadoria têm uma prioridade mais alta do que os resultados de perfil sem curadoria.
- O número de vezes que os critérios de pesquisa correspondem aos objetos, incluindo correspondências diretas e indiretas.
- Frequência relativa da palavra-chave.

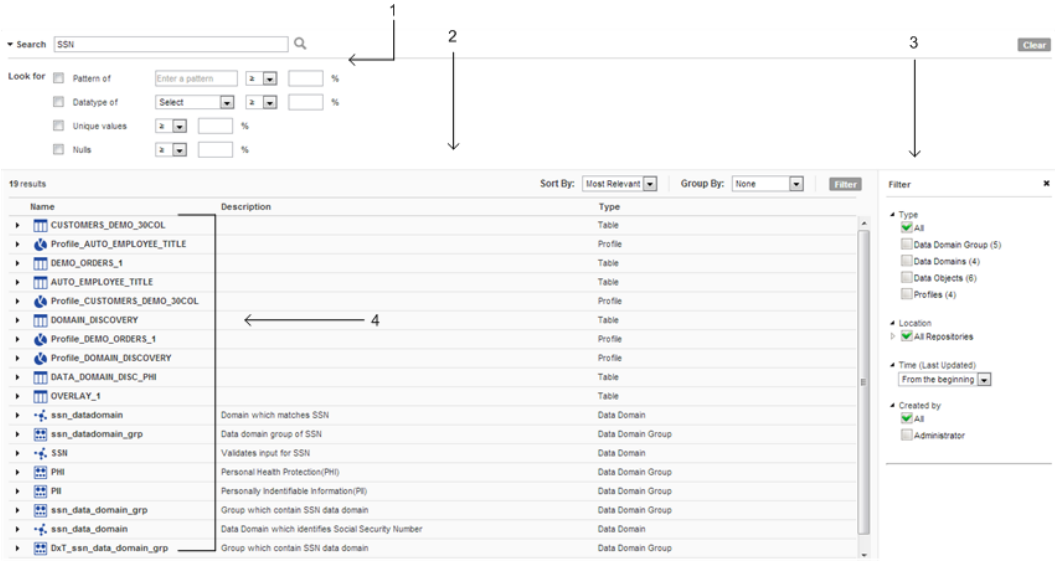
Para exibir os resultados da pesquisa, você precisa ter as permissões adequadas em projetos que contêm as correspondências diretas e indiretas.

Painel de Resultados da Pesquisa de Descoberta

Os resultados da pesquisa de descoberta incluem o nome do ativo, o tipo do ativo e a descrição do ativo. Use filtros para restringir os resultados da pesquisa.

Os resultados da pesquisa são exibidos na grade de resultados do espaço de trabalho **Biblioteca**. Você pode classificar os resultados com base em relevância. Você pode também agrupar os resultados com base no tipo de ativo, no local do repositório, na hora e no usuário que criou o ativos.

A seguinte imagem mostra a interface dos resultados da pesquisa de descoberta:



1. Critérios de pesquisa
2. Grade de resultados
3. Filtro
4. Resultados da pesquisa

O painel de resultados da pesquisa de descoberta exibe as seguintes seções por padrão:

Critérios de pesquisa

Exibe os campos de pesquisa, incluindo os filtros de pesquisa com base nos metadados de perfil que você pode definir para restringir a pesquisa. Os campos de Pesquisa aparecem na parte superior da ferramenta Analyst.

Grade de resultados

Exibe o número total de correspondência e a lista de correspondências com base nos critérios que você selecionou nos campos de critérios de pesquisa. A grade de resultados também contém uma descrição de objeto, o tipo de objeto e menus suspensos para classificar e agrupar os resultados da pesquisa.

Filtro

Exibe os filtros que você pode definir para filtrar os resultados da pesquisa. A seção **Filtro** aparece na área direita da ferramenta Analyst.

Resultados da pesquisa

Exibe os resultados da pesquisa correspondentes com base na string de pesquisa, incluindo os objetos que atendem aos critérios de pesquisa. A seção de resultados da pesquisa contém as propriedades de correspondência, a seção **Correspondências diretas** e a seção **Correspondências indiretas**, que aparece quando você expande uma correspondência. Você pode também exibir o número total de ativos relacionados à correspondência na grade de resultados.

Filtrando os Resultados da Pesquisa de Descoberta

Você pode filtrar os resultados da pesquisa com base no tipo de ativo, no local do repositório, na hora e no usuário que criou o ativo. Se você tiver o Glossário Comercial instalado, pode também usar filtros de ativos específicos para termos comerciais, categorias e diretivas no Glossário Comercial.

1. Execute a pesquisa global ou a pesquisa de descoberta na seção **Pesquisa de Descoberta** no espaço de trabalho **Biblioteca**.
2. Clique em **Filtro** na grade de resultados para abrir a seção **Filtro**.
3. Na seção **Filtro**, selecione os filtros necessários e as configurações relacionadas.
4. Os resultados da pesquisa revisados serão exibidos na grade de resultados com base nas configurações de filtro selecionadas.
5. Para limpar todas as configurações de filtro, clique em **Limpar Tudo** na parte superior da grade de resultados.

Tipos de Correspondência

Os resultados da pesquisa de descoberta incluem correspondências diretas e indiretas. Uma correspondência direta é uma correspondência com alguns ou todos os metadados do ativo que correspondem à consulta de pesquisa. Uma correspondência indireta é uma correspondência de ativo que está vinculada a outro ativo que corresponde diretamente à consulta de pesquisa.

Se uma consulta de pesquisa contém vários critérios de pesquisa, os resultados da pesquisa podem atender aos critérios de pesquisa diretamente, indiretamente ou ambos. Você pode abrir algumas das correspondências diretas e indiretas nos resultados da pesquisa no modo somente leitura ou editar.

Correspondência Direta

Uma correspondência direta é uma correspondência com alguns ou todos os metadados do ativo que correspondem à consulta de pesquisa. Por exemplo, se você pesquisar todos os ativos com o nome "Customer", a ferramenta Analyst pode listar objetos de dados e perfis com o nome "Customer" como correspondências diretas. Depois de executar uma pesquisa de descoberta, a lista de correspondências que você vê contém links para alguns dos objetos.

Você pode expandir um ativo nos resultados da pesquisa para exibir mais informações sobre a correspondência direta, como propriedades do ativo.

Correspondência Indireta

Uma correspondência indireta é uma correspondência vinculada à correspondência direta. Por exemplo, um scorecard usa uma regra que contém a palavra-chave da pesquisa. A pesquisa de descoberta retorna a regra como uma correspondência direta e o scorecard como uma correspondência indireta. O scorecard é uma correspondência indireta porque ele faz referência à regra.

Use as informações de correspondência indireta para identificar os relacionamentos ocultos entre objetos e entender melhor os relacionamentos do objeto. Você também pode os resultados de correspondência indireta para entender por que pesquisa de descoberta retorna um objeto.

Exibindo Informações de Correspondência

Depois de executar a pesquisa de descoberta, você pode exibir as informações de correspondência, incluindo correspondências diretas e indiretas. Você pode também exibir as propriedades do ativo, como tipo do ativo, descrição e ativos relacionados. Você pode abrir alguns dos ativos dos resultados da pesquisa e fazer alterações neles, se necessário.

1. Execute a pesquisa global ou a pesquisa de descoberta na seção **Pesquisa** do espaço de trabalho **Biblioteca**.
2. Na grade de resultados, clique no ícone "expandir" no início do nome de um ativo.

   As propriedades do ativo e as informações de correspondência serão exibidas em uma seção sob o nome do ativo.
3. Revise as informações de correspondência direta e indireta.

   Você pode ver os relacionamentos do ativo e outras informações, como o número total de ativos relacionados. Os relacionamentos do ativo incluem correspondências diretas e indiretas.
4. Se as informações do ativo contiverem hiperlinks, clique nos hiperlinks para abrir os ativos em outro espaço de trabalho.
5. Clique no ícone "expandir" novamente para fechar a seção de informações de correspondência.

Abrindo Ativos dos Resultados da Pesquisa de Descoberta

Você precisa ter o projeto necessário, o ativo e as permissões de licença para exibir ativos nos resultados da pesquisa de descoberta.

1. Execute a pesquisa global ou a pesquisa de descoberta na seção **Pesquisa** do espaço de trabalho **Biblioteca**.
2. Na grade de resultados, clique com o botão direito do mouse no nome de um ativo.

   Um menu de atalho é exibido.
3. Para exibir o ativo em um modo somente leitura em seu espaço de trabalho, selecione **Abrir**.

4. Para fazer alterações no ativo em seu espaço de trabalho, selecione **Editar**.
5. Para excluir o ativo dos resultados da pesquisa, selecione **Excluir**.

   Quando você exclui um ativo dos resultados da pesquisa, a ferramenta Analyst remove o ativo do repositório do Modelo.
6. Para navegar de volta ao espaço de trabalho **Biblioteca**, clique em **Biblioteca**.

Ativos Relacionados

Você pode exibir os ativos relacionados de um ativo dos resultados da pesquisa. Um ativo relacionado é um ativo no repositório do Modelo ou no Glossário Comercial associado a um ativo selecionado nos resultados da pesquisa. O ativo relacionado compartilha alguns metadados com o ativo nos resultados da pesquisa. Uma fonte de dados pode ter perfis, domínios de dados inferidos e mapeamentos como ativos relacionados.

Por exemplo, um perfil pode fazer parte dos resultados da pesquisa. Você pode exibir os ativos relacionados do perfil, como regras e fontes de dados do perfil. Você pode exibir os ativos relacionados no espaço de trabalho **Ativos Relacionados**. Os ativos relacionados que você vê diferem com base no tipo de ativo. Por exemplo, ao exibir os ativos relacionados de uma regra, você pode exibir ativos, como o termo comercial associado, a especificação de mapeamento e o perfil.

Ativos relacionados a Cada Tipo de Ativo

Os ativos relacionados que você vê em um ativo na ferramenta Analyst dependem do tipo de ativo que você pesquisar.

A seguinte tabela descreve os ativos relacionados de cada ativo:

Tipo de Ativo	Ativos Relacionados
Termo comercial	Domínio de dados, objeto de dados, mapplet e regra.
Domínio de dados	Termo comercial, grupo do domínio de dados, objeto de dados e perfil.
Grupo do domínio de dados	Domínio de dados, objeto de dados e perfil.
Objeto de dados	Termo comercial, domínio de dados, grupo do domínio de dados, especificação de mapeamento, perfil, scorecard, mapeamento e mapplet.
Perfil de descoberta empresarial	Objeto de dados e perfil.
Mapeamento **Nota:** Este objeto é aberto na ferramenta Developer.	Objeto de dados, especificação de mapeamento, mapplet e regra.
Especificação de mapeamento	Objeto de dados, perfil, scorecard, mapeamento, mapplet e regra.
Mapplet **Nota:** Este objeto é aberto na ferramenta Developer.	Termo comercial, objeto de dados, especificação de mapeamento, mapeamento, mapplet e regra.

Tipo de Ativo	Ativos Relacionados
Perfil	Domínio de dados, grupo do domínio de dados, objeto de dados, regra e especificação de mapeamento. **Nota:** Um scorecard não está incluído nos ativos relacionados de um perfil.
Regra	Objeto de dados, regra, termo comercial, especificação de mapeamento, perfil, scorecard e mapeamento.

Exibindo Ativos Relacionados

Você pode exibir o número total de ativos relacionados ao exibir as informações de correspondência do ativo nos resultados da pesquisa.

1. Execute a pesquisa de descoberta no espaço de trabalho **Biblioteca**.
2. Na grade de resultados, clique no ícone expandir e clique no link da contagem de ativos relacionados ou clique com o botão direito do mouse no nome do ativo e selecione **Mostrar Ativos Relacionados**.

   Uma lista de todos os ativos relacionados será exibida no espaço de trabalho **Ativos Relacionados**.
3. Para exibir os detalhes de um ativo, clique no nome do ativo ou clique duas vezes e selecione **Abrir**.
4. Para exibir os ativos relacionados de um ativo relacionado, clique com o botão direito do mouse no nome do ativo e selecione **Mostrar Ativos Relacionados**.

   As informações do ativos relacionados serão exibidas no espaço de trabalho.
5. Para navegar entre vários espaços de trabalho de ativos relacionados, do espaço de trabalho **Ativos Relacionados**, selecione um dos ativos abertos recentemente.

Perguntas Frequentes

Por que não consigo exibir alguns dos resultados da pesquisa que eu esperava ver?

Os resultados da pesquisa podem não ser exibidas devido a vários motivos. Verifique se os critérios de pesquisa atendem às seguintes diretrizes:

- Os ativos que são exibidos nos resultados da pesquisa dependem das permissões do projeto.
- Os resultados da pesquisa de descoberta não incluem as frequências de valor dos resultados do perfil da coluna.
- Os resultados da pesquisa não incluem os resultados de perfil que você rejeitar ao fazer curadoria dos resultados do perfil.
- Os resultados da pesquisa que você exibe dependem do intervalo de extração do índice de pesquisa e da disponibilidade dos ativos no índice de pesquisa.

Posso salvar os resultados de pesquisa de descoberta para uso futuro ou compartilhá-los com outro usuário?

Não. Não é possível salvar ou compartilhar os resultados de pesquisa de descoberta.

Por que vejo alguns dos resultados da pesquisa de descoberta na parte superior ou inferior dos resultados da pesquisa?

A ordem na qual a ferramenta Analyst exibe os resultados da pesquisa depende de vários fatores. Alguns dos fatores são o tipo de objeto, os resultados de perfil com curadoria, a propriedade do objeto que corresponde principalmente aos critérios de pesquisa e a classificação interna de pesquisa de cada objeto.

Posso exportar os resultados da pesquisa de descoberta?

Não. Não é possível exportar os resultados da pesquisa.

CAPÍTULO 16

Business Glossary Desktop no Informatica Analyst

Este capítulo inclui os seguintes tópicos:

- Termos Comerciais, 136
- Gerenciar Termos Comerciais no Glossário Comercial do Metadata Manager, 137
- Pesquisar um Termo Comercial no Glossário Comercial Desktop, 137

Termos Comerciais

Você pode procurar termos comerciais no Glossário Comercial Desktop. Também pode exibir termos comerciais e executar as tarefas de termo comercial com base na licença para o Metadata Manager.

Um glossário comercial é um conjunto de termos comerciais que usam o idioma para definir os conceitos para usuários comerciais. Um termo comercial fornece a definição e o uso comercial de um conceito.

O Glossário Comercial Desktop é um cliente que se conecta ao Serviço do Metadata Manager, que hospeda o glossário comercial. O Glossário Comercial Desktop deve estar aberto antes que um nome de objeto seja pesquisado na ferramenta Analyst. Você pode pesquisar o significado de um nome de objeto da ferramenta Analyst como termo comercial no Glossário Comercial Desktop para entender seus requisitos comerciais e implementação atual.

O Metadata Manager hospeda glossários comerciais. Você deve associar um Serviço do Metadata Manager ao Analyst Service para procurar em um glossário comercial do Metadata Manager na ferramenta Analyst. Também pode exibir os termos comerciais em um glossário comercial ou exibi-los agrupados por categoria. Você pode editar os termos comerciais do Metadata Manager.

Pode, ainda, pesquisar objetos do Metadata Manager no repositório do Metadata Manager por um termo comercial do Metadata Manager. Também é possível selecionar objetos do Metadata Manager dos resultados de pesquisa e importá-los como objetos de dados para a ferramenta Analyst. Não é possível adicionar um termo comercial do Metadata Manager ao glossário comercial do Metadata Manager.

Gerenciar Termos Comerciais no Glossário Comercial do Metadata Manager

Acesse o Glossário Comercial do Metadata Manager na ferramenta Analyst para gerenciar os termos comerciais do Metadata Manager.

1. No cabeçalho da ferramenta Analyst, clique em **Gerenciar** > **Gerenciar Termos**.

   O Metadata Manager e o Glossário Comercial do Metadata Manager são exibidos em outra guia. Os termos comerciais do Metadata Manager aparecem na exibição **Glossário** no Metadata Manager.
2. Para escolher um glossário comercial, selecione um glossário na lista Mostrar.
3. Para exibir os termos comerciais agrupados por uma categoria, clique em **Ações** > **Exibir** > **Categorias**.
4. Para exibir todos os termos comerciais em um glossário comercial em ordem alfabética, clique em **Ações** > **Exibir** > .
5. Para exibir todos os termos comerciais que começam com uma letra específica, clique na letra.
6. Para editar um termo comercial, selecione o termo comercial e clique em **Ações** > **Editar Propriedades**.

Pesquisar um Termo Comercial no Glossário Comercial Desktop

Pesquise o nome de um objeto da ferramenta Analyst como termo comercial no Glossário Comercial Desktop para entender seus requisitos comerciais e implementação atual.

O Glossário Comercial Desktop deve estar instalado no computador.

1. Destaque o nome de um objeto.
2. Use a combinação de tecla de acesso para pesquisar o nome de objeto como termo comercial no Glossário Comercial Desktop.

   A combinação padrão de tecla de acesso é `SHIFT+ALT+Q`.

Parte III: Descoberta de Dados com o Informatica Developer

Esta parte contém os seguintes capítulos:

- Perfis do Informatica Developer, 139
- Perfis de Objeto de Dados, 143
- Perfis de coluna em fontes de dados semiestruturadas, 156
- Coluna de Resultados de Perfil no Informatica Developer, 163
- Regras no Informatica Developer, 169
- Scorecards no Informatica Developer, 171
- Criação de Perfil de Mapplet e Mapeamento, 174
- Descoberta do Domínio de Dados no Informatica Developer, 177
- Descoberta Empresarial no Informatica Developer, 190
- Resultados da Descoberta Empresarial, 206
- Glossário Comercial Desktop no Informatica Developer, 217

CAPÍTULO 17

Perfis do Informatica Developer

Este capítulo inclui os seguintes tópicos:

- Perfis do Informatica Developer - Visão Geral, 139
- Exibições de Perfil do Informatica DeveloperExibições de Perfil, 141
- Bloqueios de Objetos de Repositório e Desenvolvimento Baseado em Equipe com Objetos com Controle de Versão, 142

Perfis do Informatica Developer - Visão Geral

Crie e execute perfis no Informatica Developer para descobrir problemas de qualidade de dados em um conjunto de dados e entender os relacionamentos entre as colunas de um conjunto de dados.

Você pode criar perfis para os seguintes tipos de análise de dados:

- Criação de perfil de coluna
- Criação de perfil de coluna em fontes de dados semiestruturadas
- Descoberta de chave primária
- Descoberta de dependência funcional
- Descoberta de chave externa
- Análise de associação
- Descoberta de sobreposição
- Descoberta do domínio de dados
- Descoberta empresarial

Você cria perfis na Developer tool por meio de um assistente. O assistente de criação de perfil oferece as opções **Perfil**, **Vários Perfis** e **Perfil de Descoberta Empresarial** para criar perfis.

Perfil

Crie um perfil para um único objeto de dados. Em um único perfil, você define filtros, regras e opções de busca detalhada para a criação de perfil de coluna. Você também pode escolher opções avançadas para criar um perfil de coluna, perfil de chave primária e perfil de dependência funcional, bem como para descoberta do domínio de dados. Os resultados mostram o perfil de coluna, a inferência de chave primária, a dependência funcional e a inferência de domínio de dados. Você pode criar um perfil de coluna para um objeto de dados de arquivo simples, um objeto de dados relacionais e objetos de dados semiestruturados.

Vários Perfis

Crie um conjunto de perfis para vários objetos. A Developer tool cria um perfil para cada objeto e executa os perfis simultaneamente. Quando você cria vários perfis de uma só vez, não pode analisar os dados de todos os objetos.

Perfil de Descoberta Empresarial

Crie um modelo de dados a partir de vários objetos de dados e crie um perfil que analisa dados de todos os objetos. Crie um perfil de descoberta empresarial e adicione a ele objetos de dados físicos junto com os quais você deseja criar um perfil. Você pode criar um perfil de objeto de dados, perfil de chave externa e perfil de associação. Para cada objeto de dados no perfil de descoberta empresarial, você pode configurar propriedades gerais, colunas para criação do perfil, chaves e relacionamentos. Você pode descobrir dados de sobreposição em uma ou várias fontes de dados.

Você pode também executar a descoberta empresarial que cria e executa tarefas de descoberta de dados, como perfil de coluna, descoberta do domínio de dados, perfil de chave primária e perfil de chave externa. A descoberta empresarial é executada em um grande número de fontes de dados em várias conexões.

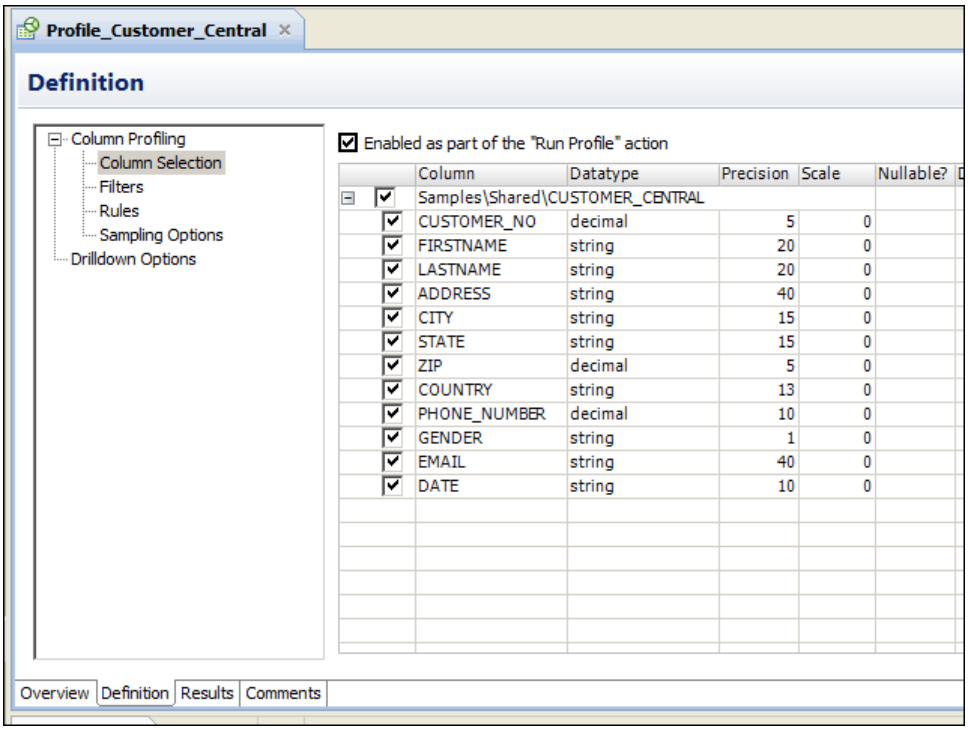
A tabela a seguir lista as operações que você pode executar com cada tipo de perfil:

Opção de Perfil	Operações de Perfil
Perfil	- Execute um perfil de coluna. - Localizar chaves primárias. - Localizar dependências funcionais. - Identificar domínios de dados
Vários Perfis	Criar e executar perfis de coluna em vários objetos ao mesmo tempo.
Perfil de Descoberta Empresarial	- Executar um perfil de coluna em um único conjunto de dados. - Localizar chaves primárias. - Localizar chaves externas. - Localizar dependências funcionais. - Executar análise de associação. - Descobrir sobreposição entre duas colunas. - Executar descoberta empresarial.

Exibições de Perfil do Informatica DeveloperExibições de Perfil

Você pode exibir e adicionar informações sobre um perfil no Informatica Developer usando as exibições **Visão Geral**,**Definição**,**Comentários** e **Resultados**.

A seguinte figura mostra as exibições de perfil no editor:



Quando você abre um perfil na exibição **Object Explorer**, o editor no painel direito mostra as informações do perfil nas seguintes exibições:

Visão Geral

Exiba e forneça informações gerais sobre o perfil, como nome, descrição e localização.

Definição

Exiba e configure a definição do perfil.

Essas informações incluem a lista de filtros e regras que você atribui ao perfil, opções de busca detalhada e funções de perfil habilitadas durante a execução do perfil.

Essas informações incluem a lista de filtros e regras atribuídos ao perfil.

Resultados

Mostra os resultados da execução do perfil. Você pode exportar os resultados após executar um perfil.

Comentários

Exiba e adicione comentários ao perfil.

Bloqueios de Objetos de Repositório e Desenvolvimento Baseado em Equipe com Objetos com Controle de Versão

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, desfazer check-outs e exibir os perfis que colocou em check-out.

O repositório do Modelo reterá bloqueios de objeto se a Developer tool for interrompida inesperadamente. Quando você se conectar ao repositório do Modelo novamente, poderá exibir os objetos que bloqueou. Você pode continuar a editar os objetos ou pode desbloqueá-los. Você pode visualizar e desbloquear objetos bloqueados usando a caixa de diálogo **Objetos Bloqueados**. Para exibir a caixa de diálogo **Objetos Bloqueados**, clique em **Exibir** > **Objetos Bloqueados**.

Se o repositório do Modelo estiver integrado a um sistema de controle de versão, você poderá gerenciar versões de objetos na Developer tool com o gerenciamento de objetos com controle de versão. É possível realizar ações como fazer check-in e check-out de objetos, exibir e recuperar versões históricas de objetos e desfazer um check-out.

O repositório do Modelo impede que os objetos sejam substituídos por outros membros da equipe de desenvolvimento. Se você abrir um objeto colocado em check-out por outro usuário, receberá uma notificação identificando o usuário responsável por esse check-out. Você pode abrir um objeto com check-out no modo somente leitura ou pode salvá-lo com um nome diferente.

O repositório do Modelo cria novas versões de objetos quando o objeto é modificado.

O repositório do Modelo incrementa o número de versão depois que você obtém o controle de um tipo de dados, uma chave primária, uma chave estrangeira ou um domínio de dados.

Quando você restaura uma versão, os resultados de perfil mais recentes aparecem na exibição Resultados, em vez dos resultados de perfil da versão restaurada. Isso acontece porque o sistema de controle de versão mantém as definições de perfil para todas as versões no repositório do Modelo, e os resultados de perfil são extraídos do warehouse de criação de perfil. Para obter mais informações sobre bloqueios de objetos de repositório e o gerenciamento de objetos com controle de versão, consulte o *Guia da Developer tool*.

CAPÍTULO 18

Perfis de Objeto de Dados

Este capítulo inclui os seguintes tópicos:

- Visão Geral dos Perfis de Objeto de Dados, 143
- Perfis de Coluna no Informatica Developer, 144
- Ambiente de tempo de execução, 146
- Descoberta de Chave Primária, 148
- Descoberta de Dependência Funcional, 149
- Perfis de sistema operacional no Informatica Developer, 151
- Criando um perfil de objeto de dados único no Developer Informatica, 152
- Criando Vários Perfis de Objetos de Dados no Informatica Developer, 153
- Editando um Perfil, 153
- Sincronizando um Objeto de Dados de Arquivo Simples, 153
- Sincronizar um Objeto de Dados Relacionais, 154
- Comentários, 154

Visão Geral dos Perfis de Objeto de Dados

Um perfil de objeto de dados descobre informações sobre os dados da coluna e os metadados em uma fonte de dados. Você pode executar um perfil em um único objeto de dados e em vários objetos de dados no Informatica Developer. Um perfil de objeto de dados único analisa uma única fonte de dados. Vários perfis de objeto de dados analisam mais de uma fonte de dados. Ao criar vários perfis de objeto de dados, você pode executar um perfil de coluna neles.

A seguinte tabela descreve as tarefas de descoberta de dados que você pode executar em um único perfil de objeto de dados:

Tarefa	Descrição
Criação de perfil de coluna	Descobre as características dos dados, como frequências, porcentagens e padrões. 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.
Descoberta de Chave Primária	Descobre colunas com valores que podem identificar exclusivamente as linhas de uma fonte de dados.

Tarefa	Descrição
Descoberta de Dependência Funcional	Descobre dependências entre pares de colunas em uma fonte de dados.
Descoberta do Domínio de Dados	Identifica todos os domínios de dados de uma coluna com base no valor ou no nome da coluna.

A seguinte tabela descreve as tarefas de descoberta de dados que você pode executar em vários objetos de dados ao criar um modelo de dados usando a opção **Perfil de Descoberta Empresarial**:

Tarefa	Descrição
Descoberta de Chave Externa	Descobre colunas que têm valores que correspondem aos valores de chave primária em outra fonte de dados.
Análise de Associação	Descobre o grau de possíveis associações entre os dados em duas colunas de uma fonte de dados ou entre duas fontes de dados.
Descoberta de Sobreposição	Descobre a porcentagem de dados de sobreposição entre pares de colunas de uma ou várias fontes de dados.
Descoberta Empresarial	Descobre estatísticas de perfil de coluna, domínios de dados, chaves primárias e chaves externas em um grande número de fontes de dados distribuídas por várias conexões ou esquemas.

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

É possível adicionar filtros avançados ou filtros SQL para determinar as linhas que um perfil de coluna utiliza quando você executa o perfil. O perfil não processa linhas que não atendam aos critérios de filtro.

Criando um filtro avançado

Você pode criar um filtro avançado com expressões, como AND, OR e NOT para fazer um subconjunto da fonte de dados original.

1. Crie ou abra um perfil de objeto de dados único.
2. Selecione a exibição **Filtro**.
3. Clique em **Adicionar**.

   A caixa de diálogo **Selecionar Assistente** é exibida.
4. Na caixa de diálogo **Selecionar Assistente**, clique em **Filtro Avançado**.

   A caixa de diálogo **Filtro** é exibida.
5. Insira um nome e uma descrição opcional para o filtro avançado.
6. Selecione **Definir como Ativo** para aplicar o filtro ao perfil. Clique em **Avançar**.
7. Selecione **Definição de Filtro** para definir um filtro.
8. Você pode criar um filtro avançado com os painéis **Funções** ou **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 um filtro avançado.
9. Para verificar o filtro avançado, clique em **Validar**.
10. Depois de criar ou editar o filtro, selecione **Visualização de Dados** para exibir os dados filtrados. Você pode definir a opção **Máximo de linhas para visualização**.
11. Clique em **Concluir**.

    O assistente de **Novo Perfil** aparece com o filtro na exibição **Filtros**.

Criando um filtro SQL

Você pode criar um filtro SQL com consultas SQL. Você pode criar um filtro SQL para fontes de dados relacionais.

1. Crie ou abra um perfil de objeto de dados único.
2. Selecione a exibição **Filtro**.
3. Clique em **Adicionar**.

   A caixa de diálogo **Selecionar Assistente** é exibida.
4. Na caixa de diálogo **Selecionar Assistente**, clique em **Filtro SQL**.

   A caixa de diálogo **Filtro** é exibida.
5. Insira um nome e uma descrição opcional para o filtro avançado.
6. Selecione **Definir como Ativo** para aplicar o filtro ao perfil. Clique em **Avançar**.
7. Selecione **Definição de Filtro** para definir um filtro.

8. Use as colunas no painel **Colunas** para criar um filtro SQL.
9. Para verificar o filtro, clique em **Validar**.
10. Depois de criar ou editar o filtro, selecione **Visualização de Dados** para exibir os dados filtrados. Você pode definir a opção **Máximo de linhas para visualização**.
11. Clique em **Concluir**.

    O assistente de **Novo Perfil** aparece com o filtro na exibição **Filtros**.

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.

Ambiente de tempo de execução

Você pode escolher Nativo ou Hadoop como ambiente de tempo de execução para um perfil de coluna. O Informatica Developer define o ambiente de tempo de execução na definição de perfil depois que você escolhe um ambiente de tempo de execução.

Ambiente Nativo

Quando você executa um perfil no ambiente de tempo de execução nativo, a Developer tool 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 na mesma máquina em que ele é executado e grava os resultados do perfil no depósito de criação de perfil. Por padrão, todos os perfis são executados no ambiente de tempo de execução nativo.

Você pode usar origens nativas para criar e executar perfis no ambiente nativo. Uma fonte de dados nativa é uma origem não Hadoop, como um arquivo simples, uma origem relacional ou uma origem de mainframe. Você também pode executar um perfil em uma especificação de mapeamento ou em uma fonte de dados lógica com uma fonte de dados Hive ou HDFS no ambiente nativo.

Ambiente Hadoop

Você pode escolher a opção Hive ou Hadoop para executar os perfis no ambiente de tempo de execução Hadoop. Depois que você escolher a opção Hive e selecionar uma conexão do Hadoop, o Serviço de Integração de Dados enviará a lógica do perfil ao mecanismo Hive no cluster Hadoop para executar perfis. Depois que você escolher a opção Hadoop e selecionar uma conexão do Hadoop, o Serviço de Integração de Dados enviará a lógica do perfil ao mecanismo Blaze no cluster Hadoop para executar perfis.

Quando você executa um perfil no ambiente Hadoop, a Developer tool 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 envia os mapeamentos ao ambiente Hadoop por meio da conexão do Hadoop. O mecanismo Blaze processa os mapeamentos, e o Serviço de Integração de Dados grava os resultados do perfil no depósito de criação de perfil.

Perfis de coluna para fontes de dados Sqoop

Você pode executar um perfil de coluna em objetos de dados que usam o Sqoop. Você pode selecionar o ambiente de tempo de execução Hive ou Hadoop para executar os perfis de coluna.

No mecanismo Hive, para executar um perfil de coluna em um objeto de dados relacional que usa o Sqoop, você deve definir o argumento Sqoop argument **m** como 1 na conexão JDBC. Utilize a seguinte sintaxe:

```
-m 1
```

Ao executar um perfil de coluna em um objeto de dados lógicos ou um objeto de dados personalizados, você pode configurar o argumento num-mappers para obter paralelismo e otimizar o desempenho. Você também deve configurar o argumento split-by para especificar a coluna com base no qual o programa Sqoop deve dividir as unidades de trabalho.

Utilize a seguinte sintaxe:

```
--split-by <nome_coluna>
```

Se a chave primária não tiver uma distribuição uniforme dos valores entre o intervalo mínimo e máximo, você poderá configurar o argumento split-by para especificar outra coluna que tenha uma distribuição equilibrada dos dados para dividir as unidades de trabalho.

Se você não definir a coluna split-by, o programa Sqoop dividirá as unidades de trabalho com base nos seguintes critérios:

- Se o objeto de dados contiver uma única chave primária, o programa Sqoop a usará como a coluna split-by.
- Se o objeto de dados contiver uma chave primária composta, o programa Sqoop assumirá como padrão o comportamento Sqoop de lidar com chaves primárias compostas sem o argumento split-by. Consulte a documentação do Sqoop para obter mais informações.

- Se um objeto de dados contiver duas tabelas com uma coluna idêntica, você deverá definir a coluna split-by com um nome de tabela qualificado. Por exemplo, se o nome da tabela for CUSTOMER e o nome da coluna for FULL_NAME, defina a coluna split-by da seguinte maneira:
  `--split-by CUSTOMER.FULL_NAME`
- Se o objeto de dados não contiver uma chave primária, o valor dos argumentos m e mappers-num assumirá 1 como padrão.

Quando você usa o Cloudera Connector Powered by Teradata ou o Hortonworks Connector for Teradata, e a tabela Teradata não contém uma chave primária, o argumento split-by é necessário.

Descoberta de Chave Primária

A descoberta de chave primária gera candidatos a chave primária a partir das colunas que você especifica.

Uma chave primária é uma coluna ou combinação de colunas que identifica exclusivamente uma linha em uma fonte de dados. A descoberta de chave primária identifica as colunas e combinações de colunas que atendem a um nível de confiança específico. Você pode editar o nível de confiança, bem como o número máximo de colunas a combinar para identificação de chave primária.

A descoberta de chave primária pode destacar possíveis problemas de qualidade de dados ao identificar as linhas não exclusivas em um candidato a chave primária. Isso é especialmente útil em casos onde a descoberta de chave primária combina muitas colunas, pois os registros sem conformidade provavelmente contêm informações duplicadas.

Propriedades de Inferência de Chave Primária

Quando você cria um perfil de objeto de dados único, pode usar a exibição **Criação de Perfil de Chave Primária** para configurar as propriedades de inferência de chave primária.

A tabela a seguir descreve as propriedades de inferência de chave primária na exibição **Criação de Perfil de Chave Primária**:

Propriedade	Descrição
Substituir as opções de inferência padrão	Permite que você defina configurações personalizadas para inferência de chave primária.
Máximo de Colunas de Chave	Número máximo de colunas que podem constituir uma chave primária.
Máximo de Linhas	Número de linhas para a criação de perfil.
Critérios de Conformidade	Porcentagem mínima ou número máximo de linhas de violações de chave que o perfil permite ao determinar chaves primárias.
Exclua objetos de dados com uma chave documentada definida pelo usuário	Exclui objetos de dados com chaves primárias documentadas ou chaves primárias definidas pelo usuário.
Exclui objetos de dados com uma chave aprovada	Exclui objetos de dados com chaves primárias aprovadas.

Propriedades de Chaves Primárias Inferidas

Depois de executar um perfil de objeto de dados único, você pode usar a exibição **Criação de Perfil de Chave Primária** para ver os detalhes das chaves primárias inferidas na fonte de dados.

A tabela a seguir descreve as propriedades das chaves primárias inferidas na exibição **Criação de Perfil de Chave Primária**:

Propriedade	Descrição
Coluna	Nome da coluna no perfil.
% em Conformidade	Porcentagem de valores exclusivos na coluna.
% de Duplicatas	Porcentagem de valores duplicados para a coluna.
% Nulo	Porcentagem de valores nulos da coluna.
Verificado	Determina se a coluna é uma coluna de chave primária.
Status de Inferência	O status de inferência da coluna.
Horário da Última Execução	A data e a hora em que o perfil de chave primária foi executado pela última vez.

Propriedades de Violações de Chave

Depois de executar um perfil de objeto de dados único, você pode usar a exibição **Criação de Perfil de Chave Primária** para ver os detalhes das violações de chave primária na fonte de dados.

A tabela a seguir descreve as propriedades das violações de chave na exibição **Criação de Perfil de Chave Primária**:

Propriedade	Descrição
Coluna(s)	Nome das colunas das quais o perfil infere uma chave primária candidata.
Número de Violações de Chave	Número de violações de chave no candidato a chave primária.

Descoberta de Dependência Funcional

A descoberta de dependência funcional fornece informações sobre as dependências entre pares de colunas em uma fonte de dados.

Um par de colunas são funcionalmente dependentes se os valores de uma coluna podem prever com confiabilidade os valores de outra coluna. Por exemplo, se um conjunto de dados contiver uma coluna ID do Empregador e uma coluna de data de nascimento, a data de nascimento deverá ser a mesma em todas as linhas que contiverem essa ID do Empregador.

Dependências funcionais podem realçar possíveis problemas de qualidade de dados ao identificarem registros sem conformidade com uma dependência funcional de coluna. Por exemplo, se 99,8% das linhas

em uma fonte de dados são funcionalmente dependentes, há uma grande probabilidade de que as linhas restantes contenham informações incorretas.

Propriedades de Inferência de Dependência Funcional

A exibição **Criação de Perfil de Dependência Funcional** fornece informações sobre as dependências funcionais entre as colunas.

A tabela a seguir descreve as propriedades de inferência de dependência funcional na exibição **Criação de Perfil de Dependência Funcional**:

Propriedade	Descrição
Substituir as opções de inferência padrão	Permite que você defina configurações personalizadas para inferência de dependência funcional.
Máximo de Colunas no Determinante	Número de colunas que o perfil pode combinar para localizar um determinante.
Máximo de Linhas	Número de linhas para a criação de perfil.
Dependências Retornadas	Número de dependências que o perfil exibe. O padrão é **Cobertura Mínima**, que exibe o menor conjunto de dependências onde cada coluna aparece pelo menos uma vez em um dependente ou determinante.
Máximo de Dependências Retornadas	Número máximo de dependências que o perfil exibe.
Critérios de Conformidade	A porcentagem mínima ou o número máximo de linhas para violações de dependência que o perfil permite ao determinar dependentes funcionais.

Propriedades de Dependências Funcionais Inferidas

Depois de executar um perfil de objeto de dados único, você pode usar a exibição **Inferência de Dependência Funcional** para ver os detalhes das dependências funcionais inferidas entre as colunas na fonte de dados.

A tabela a seguir descreve as propriedades das dependências funcionais inferidas propriedades na exibição **Inferência de Dependência Funcional**:

Propriedade	Descrição
Colunas de Determinantes	Nome da coluna analisada em busca de dependências funcionais.
Colunas de Dependentes	Nome da coluna dependente da coluna determinante.
% Nulo	Porcentagem de valores nulos da coluna.
% em Conformidade	Porcentagem de correspondência de dependência funcional.
Verificado	Determina se as colunas são ou não funcionalmente dependentes.
Horário da Última Execução	A data e a hora em que o perfil de dependência funcional foi executado pela última vez.

Propriedades de Violações de Dependência Funcional

A exibição fornece informações sobre as dependências funcionais entre colunas. Depois de executar um perfil de objeto de dados único, você pode usar a exibição **Inferência de Dependência Funcional** para ver os detalhes das violações de dependência funcional na fonte de dados.

A tabela a seguir descreve as propriedades das violações de dependência funcional na exibição **Inferência de Dependência Funcional**:

Propriedade	Descrição
Coluna determinante	Nome da coluna analisada em busca de dependências funcionais.
Dependentes Distintos	Número de dependências funcionais exclusivas.

Perfis de sistema operacional no Informatica Developer

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

Selecionando um perfil do sistema operacional

Você pode selecionar um perfil do sistema operacional no Informatica Developer. 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. No Informatica Developer, clique em **Windows** > **Preferências**.

   A caixa de diálogo **Preferências** é exibida.
2. Clique em **Informatica** > **Executar Configurações** > **Mapeamento**.

   A caixa de diálogo **Mapeamento** é exibida.
3. Na caixa de diálogo **Mapeamento**, desmarque a opção **Usar Serviço de Integração de Dados padrão**.
4. Clique em **Procurar** para selecionar um perfil do sistema operacional na lista.
5. Clique em **OK**.

Criando um perfil de objeto de dados único no Developer Informatica

Você pode criar um único objeto de dados de 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. 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.
10. 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.
11. Na seção **Configurações de Execução**, escolher um ambiente de tempo de execução. Escolha **Nativo**, **Hive** ou **Hadoop** como ambiente de tempo de execução. Quando você escolher a opção **Hive** ou **Hadoop**, selecione uma conexão do Hadoop.
12. Clique em **Concluir**.

Criando Vários Perfis de Objetos de Dados no Informatica Developer

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**.

Editando um Perfil

É possível editar um único perfil de objeto de dados ou vários perfis de objeto de dados. Se o sistema de controle de versão estiver ativado, o perfil será colocado em check-out por padrão.

1. Na exibição **Object Explorer**,clique com o botão direito do mouse no perfil e depois clique em,**Abrir**.

   A exibição **Resultados** aparece.
2. Na exibição **Definição**, atualize as propriedades conforme necessário.
3. Clique em **Equipe** > **Fazer Check-In** para fazer o check-in do perfil.
4. Clique com o botão direito do mouse no perfil e depois clique em **Executar Perfil** para executar esse perfil.

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

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.

Comentários

Você pode adicionar uma descrição como um comentário a um perfil. Você também pode adicionar comentários às colunas nos resultados de perfil de coluna.

Você pode adicionar vários comentários às colunas. Você pode usar a exibição **Comentários** na Developer tool para adicionar e exibir comentários.

Adicionando comentários no Informatica Developer

Adicione comentários em colunas nos resultados de perfil de coluna. Ao exportar os resultados de perfil, a Developer tool inclui os comentários.

1. Na exibição **Object Explorer**, abra um perfil.
2. Opcionalmente, execute o perfil para atualizar os resultados do perfil.
3. Selecione a exibição **Comentários**.
4. Clique em **Adicionar** para abrir a caixa de diálogo **Adicionar Comentário**.
5. Na lista, selecione o nome do perfil ou uma das colunas.

   Se você adicionou comentários previamente, poderá exibi-los na caixa de diálogo.
6. No campo **Comentário**, insira uma descrição.

7. Clique em **OK**.

   A Developer tool exibe o comentário na exibição **Comentários**.

8. Para excluir um comentário, selecione o comentário na exibição **Comentários** e clique em **Remover**.

CAPÍTULO 19

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, 156
- Perfil de coluna em uma fonte de dados JSON ou XML, 157
- Perfil de coluna em uma fonte de dados semiestruturada no HDFS, 158
- Criando uma conexão HDFS, 158
- Criando um objeto de dados de arquivo complexo em um arquivo JSON ou XML no HDFS, 159
- Criando um objeto de dados em uma fonte de dados Avro ou Parquet, 160
- Executando um perfil de coluna em fontes de dados semiestruturadas, 161

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

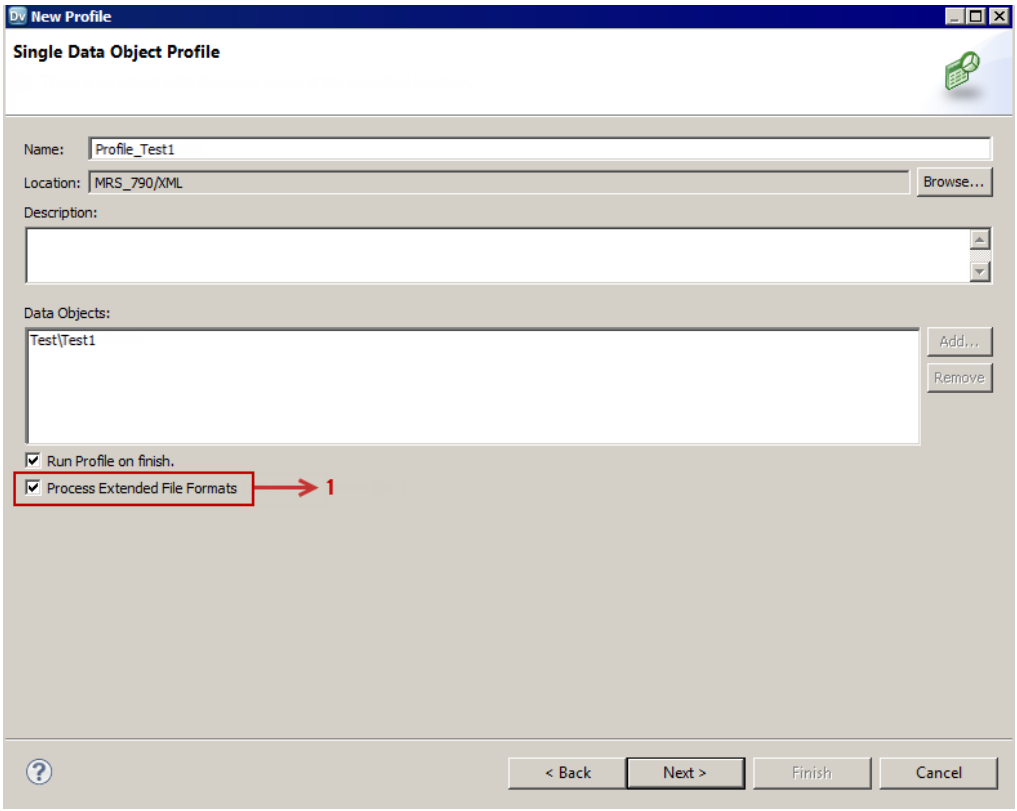
Depois de criar um objeto de dados de arquivo simples ou complexo a partir de fontes de dados Avro, JSON, Parquet ou XML, você pode criar e executar um perfil de coluna no objeto de dados.

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 semiestruturadas.

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 20

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, 163
- Propriedades de Valor de Coluna, 164
- Propriedades de Padrão de Coluna, 165
- Propriedades de Estatísticas de Coluna, 165
- Propriedades de tipo de dados da coluna, 166
- Curadoria no Informatica DeveloperCuradoria no Informatica Developer, 167
- Exportar Resultados de Perfil do Informatica Developer, 168

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.

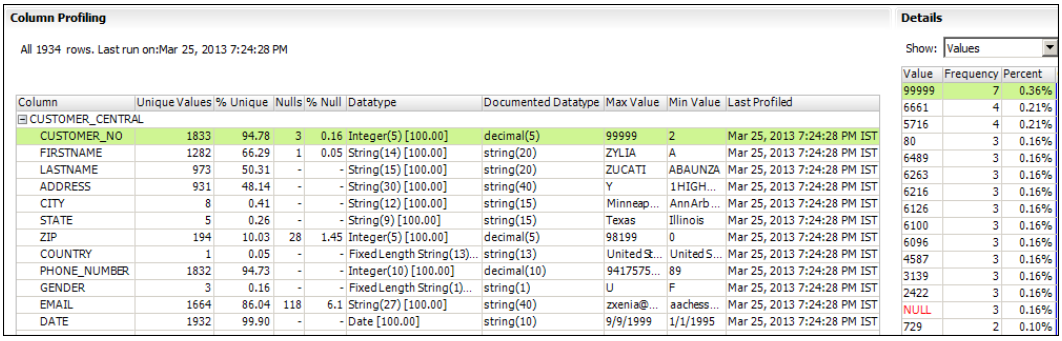
A tabela a seguir descreve os resultados do perfil para cada tipo de análise:

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:



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 21

Regras no Informatica Developer

Este capítulo inclui os seguintes tópicos:

- Visão Geral de Regras no Informatica DeveloperDiretrizes para Regras, 169
- Criar uma Regra no Informatica Developer, 170
- Aplicando uma Regra no Informatica DeveloperAplicando uma Regra, 170

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 22

Scorecards no Informatica Developer

Este capítulo inclui os seguintes tópicos:

- Visão Geral dos Scorecards no Informatica Developer, 171
- Criar um Scorecard, 171
- Exportar um Arquivo de Recursos para a Linhagem de Scorecard, 172
- Exibir a Linhagem de Scorecard do Informatica Developer, 172

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 23

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, 174
- Executar um Perfil em um Mapplet ou Objeto de Mapeamento, 174
- Comparar Perfis de Objetos de Mapplet ou de Mapeamento, 175
- Gerar um Mapeamento de Perfil, 175

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.

CAPÍTULO 24

Descoberta do Domínio de Dados no Informatica Developer

Este capítulo inclui os seguintes tópicos:

- Descoberta do Domínio de Dados no Informatica Developer - Visão Geral, 177
- Glossário do Domínio de Dados no Informatica Developer, 178
- Opções de Descoberta do Domínio de Dados no Informatica Developer, 182
- Criando um Perfil para Executar Descoberta do Domínio de Dados no Informatica Developer, 185
- Editando um Perfil no Informatica Developer, 185
- Executando um Perfil para Realizar a Descoberta do Domínio de Dados no Informatica Developer, 186
- Resultados da Descoberta do Domínio de Dados no Informatica Developer, 186

Descoberta do Domínio de Dados no Informatica Developer - Visão Geral

Use o glossário do domínio de dados para gerenciar domínios de dados. Para criar um domínio de dados, você pode usar regras de dados e regras de nome de coluna predefinidas. Você também pode gerar o domínio de dados com base nos valores ou padrões específicos dos resultados do perfil da coluna.

Você pode selecionar as colunas de origem, os domínios de dados com os quais deseja corresponder os dados e os nomes das colunas, além de opções de amostragem, opções de busca detalhada e o ambiente de tempo de execução. Você pode escolher o número máximo de linhas nas quais deseja executar a descoberta do domínio de dados. Você pode escolher critérios de conformidade para a descoberta do domínio de dados. Você pode excluir valores nulos da descoberta do domínio de dados. Depois de executar um perfil, você pode verificar, fazer curadoria e uma busca detalhada nos resultados. Você pode também adicionar os resultados a um modelo de dados do editor na ferramenta Developer.

Você pode criar um perfil com uma opção de amostragem e filtros para realizar a descoberta do domínio de dados. Ao executar o perfil, você aplica a opção de amostragem e os filtros à fonte de dados e gera um conjunto de dados. O processo de descoberta do domínio de dados usa o conjunto de dados para descobrir domínios de dados.

Glossário do Domínio de Dados no Informatica Developer

Você gerencia domínios de dados e grupos de domínios de dados no glossário do domínio de dados. É possível adicionar, editar e remover domínios de dados e grupos de domínios de dados. Você pode também pesquisar domínios de dados e grupos de domínios de dados específicos.

Você pode exportar domínios de dados do glossário do domínio de dados para um arquivo XML. Pode também importar domínios de dados de um arquivo XML para o glossário do domínio de dados. Você cria um grupo do domínio de dados para organizar domínios de dados em grupos específicos, como Informações Pessoais de Saúde, Informações de Identificação Pessoal ou qualquer outro grupo conceitual que seja relevante para o projeto. Você pode ter um domínio de dados em vários grupos de domínios de dados. Por exemplo, número de telefone pode pertencer aos grupos de domínios de dados Informações de Identificação Pessoal e Informações Pessoais de Saúde.

Criando um Grupo do Domínio de Dados no Informatica Developer

Você pode adicionar domínios de dados a grupos de domínios de dados para uma análise eficaz dos dados da coluna.

1. Clique em **Janela** > **Preferências**.

   A caixa de diálogo **Preferências** é exibida.
2. Na caixa de diálogo **Preferências**, selecione **Informatica** > **Glossário do Domínio de Dados**.

   A ferramenta Developer exibe uma lista de todos os domínios de dados no painel **Glossário do Domínio de Dados**.
3. No campo **Mostrar**, selecione **Grupos de Domínios de Dados**.

   O painel **Glossário do Domínio de Dados** reorganiza a lista de domínios de dados com base nos grupos de domínios de dados.
4. No painel **Glossário do Domínio de Dados**, selecione **Grupos de Domínios de Dados**.
5. Clique em **Adicionar**.

   A caixa de diálogo **Grupo do Domínio de Dados** é exibida.
6. Insira um nome e uma descrição.
7. Clique em **Avançar**.
8. Clique em **Escolher** para abrir a caixa de diálogo **Selecionar Domínio de Dados**.
9. Selecione os domínios de dados que você deseja adicionar ao grupo do domínio de dados e clique em **OK**.

   A ferramenta Developer lista os domínios de dados selecionados no painel **Selecionar Domínio de Dados**.
10. Clique em **Concluir**.

    A ferramenta Developer adiciona o grupo do domínio de dados ao glossário do domínio de dados.

Criando um Domínio de Dados no Informatica Developer

Você pode criar domínios de dados e adicioná-los ao glossário do domínio de dados. Pode também adicionar domínios de dados a um ou mais grupos de domínios de dados.

1. Clique em **Janela** > **Preferências**.

   A caixa de diálogo **Preferências** é exibida.
2. Na caixa de diálogo **Preferências**, selecione **Informatica** > **Glossário do Domínio de Dados**.

   A ferramenta Developer exibe uma lista de todos os domínios de dados no painel **Glossário do Domínio de Dados**.
3. No painel **Glossário do Domínio de Dados**, selecione **Domínios de Dados**.
4. Clique em **Adicionar**.

   A caixa de diálogo **Domínio de Dados** é exibida.
5. Insira um nome e uma descrição.
6. Clique em **Usar regra de dados** para descobrir domínios de dados com base nos dados da coluna. Você também pode selecionar **Usar regra de nome da coluna** para descobrir domínios de dados com base nos nomes da coluna na fonte de dados.

   O botão **Procurar** fica ativado.
7. Clique em **Procurar** para abrir a caixa de diálogo **Selecionar Localização**.
8. Selecione as regras apropriadas e clique em **OK**.

   Quando você cria um domínio de dados, a ferramenta Developer copia regras e outros objetos dependentes associados ao domínio de dados para o glossário do domínio de dados. Para editar uma regra associada a um domínio de dados, você deve ir para a regra original e fazer alterações nela. Você poderá então associar a regra modificada ao domínio de dados novamente.

   As regras que você selecionou aparecem nos campos **Regra de dados** e **Regra de nome da coluna**.
9. Clique em **Avançar**.
10. Clique em **Escolher** para abrir a caixa de diálogo **Selecionar Grupos de Domínios de Dados**.
11. Selecione os grupos de domínios de dados nos quais você deseja incluir o domínio de dados e clique em **OK**.

    A ferramenta Developer exibe os grupos de domínios de dados selecionados no painel **Atribuir a Grupos de Domínios de Dados**.
12. Clique em **Concluir**.

    A ferramenta Developer adiciona o domínio de dados ao glossário do domínio de dados.

Criando um Domínio de Dados a Partir de Resultados do Perfil no Informatica Developer

Após executar um perfil de coluna, você pode exibir os valores e padrões dos dados de origem. Você pode então criar um domínio de dados para eles.

1. Execute um perfil de coluna para exibir seus resultados.
2. Selecione os valores ou padrões com base nos quais você deseja criar um domínio de dados.

   Os valores, padrões e estatísticas aparecem na exibição **Resultados**.
3. Clique com o botão direito do mouse nos valores ou padrões e selecione **Enviar para** > **Novo Domínio de Dados**.

A caixa de diálogo **Domínio de Dados** é exibida.

4. Digite o nome do domínio de dados e uma descrição opcional.

   A localização está definida como o glossário do domínio de dados por padrão.

5. Clique em **Concluir**.

   O domínio de dados é adicionado ao glossário do domínio de dados.

Localizar Domínios de Dados no Informatica Developer

O glossário do domínio de dados exibe todos os domínios de dados por padrão. Você pode pesquisar domínios de dados e grupos de domínios de dados específicos.

O painel **Glossário do Domínio de Dados** na caixa de diálogo **Preferências** exibe todos os domínios de dados e grupos de domínios de dados. Você pode pesquisar e exibir mais informações sobre domínios de dados e grupos de domínios de dados das seguintes maneiras:

Pesquisar domínios de dados e grupos de domínios de dados.

Digite parte do nome do domínio de dados ou do grupo do domínio de dados no campo na parte superior do painel **Glossário do Domínio de Dados**. Se você estiver na **Exibição Grupos de Domínios de Dados**, a ferramenta Developer listará os domínios de dados que contiverem a string de pesquisa em seus nomes e os grupos de domínios de dados associados a eles. Se você estiver na **Exibição Domínios de Dados**, a ferramenta Developer listará todos os domínios de dados que contiverem a string de pesquisa em seus nomes.

Exibir grupos de domínios de dados e domínios de dados dentro deles.

No campo **Mostrar**, selecione **Grupos de Domínios de Dados**.

Exibir todos os domínios de dados.

No campo **Mostrar**, selecione **Domínios de Dados**.

Exibir propriedades de um domínio de dados.

Clique em um nome de domínio de dados para exibir suas propriedades abaixo do painel **Glossário do Domínio de Dados**. Você pode exibir a descrição e as regras associadas. Para exibir os grupos aos quais um domínio de dados pertence, clique em **Mostrar nos grupos de domínios de dados**.

Exibir propriedades de um grupo do domínio de dados.

Clique em um nome de grupo do domínio de dados para exibir sua descrição abaixo do painel **Glossário do Domínio de Dados**.

Importando Domínios de Dados

Você pode importar domínios de dados de um arquivo XML de origem para o glossário do domínio de dados na ferramenta Developer. Você deve verificar se o arquivo contém informações sobre os domínios de dados que é necessário importar.

1. Abra o glossário do domínio de dados.
2. Verifique se **Domínios de Dados** ou **Grupos de Domínios de Dados** está selecionado.
3. Clique em **Importar**.

   A caixa de diálogo **Importar** é exibida.

4. No campo **Nome do Arquivo**, digite o nome do arquivo XML do qual você deseja importar domínios de dados.

   Clique em **Procurar** para selecionar o arquivo.

5. Clique em **Avançar**.

   O painel **Selecionar Objetos para Importação** é exibido, onde você pode especificar a origem e o destino.

6. No painel **Origem**, selecione os domínios de dados que você deseja importar.

   Nota: Para selecionar vários domínios de dados, mantenha a tecla Shift pressionada.

7. Clique em **Corresponder Automaticamente ao Destino** para mover os domínios de dados para o painel **Destino**.

   A ferramenta Developer tenta corresponder os descendentes da seleção de origem atual individualmente pelo nome, tipo e hierarquia pai na seleção de destino e adiciona os objetos com correspondência.

8. Clique em **Resolução** para especificar como lidar com objetos duplicados.

   Você pode renomear o objeto importado, substituir o objeto existente pelo objeto importado ou reutilizar o objeto existente. A ferramenta Developer renomeia todos os objetos duplicados por padrão.

9. Clique em **Avançar**.

   A ferramenta Developer resume as configurações de importação para sua revisão. Você pode especificar configurações de importação adicionais no painel **Configurações de Importação Adicionais**.

10. Clique em **Concluir**.

Exportando Domínios de Dados

Você pode exportar domínios de dados e regras de domínio de dados do glossário do domínio de dados na ferramenta Developer para um arquivo XML.

1. Abra o glossário do domínio de dados.
2. Verifique se **Domínios de Dados** ou **Grupos de Domínios de Dados** está selecionado.
3. Clique em **Exportar**.

   A caixa de diálogo **Exportar** é exibida.

4. Para exportar domínios de dados, selecione **Exportar Domínios de Dados**. Selecione **Exportar Regras de Domínio de Dados** para exportar regras de domínio de dados.
5. Clique em **Avançar**.
6. No painel **Exportar para arquivo**, selecione os domínios de dados ou regras de domínio de dados que você deseja exportar.
7. Para exportar domínios de dados, clique em **Procurar** para escolher o arquivo de exportação e sua localização. Para exportar regras de domínio de dados para outro projeto no Model Repository Service, selecione **Copiar para o projeto** e escolha o projeto para o qual você deseja copiar as regras de domínio de dados.
8. Clique em **Avançar**.

   O painel **Dependências** exibe uma lista dos objetos dependentes.

9. Clique em **Avançar**.

   O painel **Configurações de Exportação de Conteúdo** é exibido. Você pode selecionar tabelas de referência associada para exportar.

10. Clique em **Concluir**.

    Se você associar uma regra que usa tabelas de referência com um domínio de dados, poderá não ser possível exportar as tabelas de referência na mesma sessão da ferramenta Developer que você usa para criar o domínio de dados. Após clicar em **Exportar** no glossário do domínio de dados, desconecte-se do Model Repository Service e se conecte novamente antes de importar a regra que usa tabelas de referência.

Opções de Descoberta do Domínio de Dados no Informatica Developer

Você pode selecionar as colunas de origem, domínios de dados e opções de inferência quando cria um perfil para executar a descoberta do domínio de dados. Também é possível optar por omitir colunas da descoberta de domínio de dados com base nos tipos e tamanho dos dados.

Seleção de Domínio de Dados no Informatica Developer

As opções de **Seleção de Domínio de Dados** listam todos os domínios do glossário do domínio de dados. Você pode pesquisar domínios de dados específicos e selecioná-los antes de executá-los como parte da descoberta do domínio de dados.

A tabela a seguir descreve as opções de **Seleção de Domínio de Dados** para descoberta do domínio de dados:

Opção	Descrição
Ativado como parte da ação "Executar Perfil"	Inclui as opções de descoberta do domínio de dados quando você executa o perfil.
Nome	Nome do domínio de dados.
Descrição	Descrição para o domínio de dados.
Grupo do Domínio de Dados	Nome dos grupos de domínios de dados aos quais o domínio de dados pertence.
Mostrar grupo do domínio de dados na hierarquia	Lista todos os grupos de domínios de dados com os domínios agrupados sob cada grupo.

Seleção de Colunas de Domínio de Dados no Informatica Developer

Você pode usar as opções de **Seleção de Coluna** para escolher as colunas que deseja executar como parte da descoberta do domínio de dados.

A tabela a seguir descreve as opções de **Seleção de Coluna** para descoberta do domínio de dados:

Opção	Descrição
Coluna	Nome da coluna.
Tipo de dados	Tipo de dados da coluna.
Precisão	Precisão máxima para a coluna.
Escala	Escala da coluna.
Anulável	Indica uma coluna que pode ter valores nulos.
Descrição	Descrição para a coluna.

Opções de Inferência de Domínio de Dados no Informatica Developer

As opções de inferência determinam se a descoberta de domínio deve ser executada nos dados da coluna, nome da coluna ou em ambos. Você poderá especificar se o perfil precisar processar todas as linhas da fonte de dados. Você pode escolher critérios de conformidade para correspondência de domínio de dados e optar por excluir nulos da descoberta do domínio de dados.

A tabela a seguir descreve as opções de **Inferência** para a descoberta do domínio de dados:

Opção	Descrição
Substituir as opções de inferência padrão	Permite que você altere as opções de inferência predefinidas.
Dados	O perfil é executado nos dados da coluna.
Nome da coluna	O perfil é executado nos títulos da coluna.
Nome da coluna e dados	O perfil é executado tanto nos dados quanto nos títulos da coluna.
Máximo de linhas para o perfil	O número máximo de linhas nas quais o perfil pode ser executado. A Developer tool escolhe as linhas começando da primeira linha da origem.
Porcentagem mínima de linhas	A porcentagem mínima de conformidade de linhas no conjunto de dados necessária para uma correspondência de domínio de dados.
Número mínimo de linhas	O número mínimo de linhas no conjunto de dados necessárias para uma correspondência de domínio de dados.
Excluir valores nulos da descoberta do domínio de dados	Exclui os valores nulos do conjunto de dados para a descoberta do domínio de dados.

Porcentagem Mínima de Conformidade

Você pode escolher uma porcentagem mínima de linhas no conjunto de dados como critérios de conformidade para a descoberta do domínio de dados.

A porcentagem de conformidade é a proporção entre o número de linhas correspondentes dividido pelo número total de linhas.

Nota: A Developer tool considera valores nulos como linhas sem correspondência. As colunas que contêm um número alto de valores nulos podem não resultar em inferência de domínio de dados, a menos que você especifique um valor baixo para a porcentagem mínima de conformidade.

Exemplo

Você tem uma fonte de dados com 10.000 linhas em que a coluna Comentários tem Números de Previdência Social em 2.500 linhas. Você cria um perfil da coluna com descoberta do domínio de dados e define uma porcentagem mínima de linhas como 30% como os critérios de conformidade. Quando você executa o perfil, os resultados do perfil não exibem os Números de Previdência Social como um domínio de dados inferido, pois os critérios mínimos de conformidade correspondem a 30% das linhas, ou 3.000 linhas, na fonte de dados.

Mínimo de Linhas em Conformidade

Você pode escolher um número mínimo de linhas no conjunto de dados como critérios de conformidade para a descoberta do domínio de dados.

Exemplo

Você tem uma fonte de dados com 10.000 linhas em que a coluna Comentários tem um endereço de e-mail em três linhas. Você cria um perfil de coluna com um perfil de descoberta do domínio de dados e define o número mínimo de linhas como 1 como os critérios de conformidade. Quando você executa o perfil, os resultados do perfil exibem o endereço de e-mail como um domínio de dados inferido com três linhas em conformidade, juntamente com os outros domínios de dados inferidos.

Excluir Valores Nulos

Você pode excluir valores nulos ao realizar uma descoberta do domínio de dados em uma fonte de dados. Quando você seleciona a porcentagem mínima de linhas com a opção de exclusão de valores nulos, a porcentagem de conformidade é a proporção do número de linhas correspondentes dividido pelo número total de linhas, menos os valores nulos na coluna.

O processo de descoberta do domínio de dados é diferente quando você escolhe a opção **Excluir valores nulos da descoberta do domínio de dados** e os vários filtros e opções de amostragem.

Os seguintes cenários explicam os resultados de uma descoberta do domínio de dados quando você escolhe a opção para excluir valores nulos junto com uma opção de amostragem e filtros:

- Com **Todas as linhas** como opção de amostragem e sem filtros. A descoberta do domínio de dados ignora todos os valores nulos na coluna.
- Com uma opção de amostragem e sem filtros. A descoberta do domínio de dados ignora todos os valores nulos nos dados amostrados e executa o restante desses dados.
- Com **Todas as linhas** como opção de amostragem e com filtros. A descoberta do domínio de dados ignora todos os valores nulos nos dados filtrados e executa o restante desses dados.
- Com uma opção de amostragem e filtros. A descoberta do domínio de dados ignora valores nulos nos dados filtrados da amostra e executa o restante desses dados.

Exemplo

Você tem uma fonte de dados com 10.000 linhas em que 3.000 linhas têm Números de Previdência Social na coluna Comentários. Você cria um perfil de coluna com descoberta do domínio de dados e escolhe as seguintes opções:

- Selecione a opção **Excluir valores nulos da descoberta do domínio de dados**.
- Selecione **Todas as linhas** como opção de amostragem.
- Selecione a opção **Porcentagem mínima de linhas** e configure a opção como 12%.

Quando você executa o perfil, ele é executado no conjunto de dados e ignora os valores nulos para a descoberta do domínio de dados.

Criando um Perfil para Executar Descoberta do Domínio de Dados no Informatica Developer

Você pode descobrir domínios de dados em uma fonte de dados como parte de um único perfil de objeto de dados ou um perfil de descoberta empresarial. Depois de executar a descoberta de domínio de dados, você pode verificar, fazer uma busca detalhada e adicionar os resultados a um modelo de dados no editor na ferramenta Developer.

1. Na exibição **Object Explorer**, selecione o projeto que contém o objeto de dados para o perfil.
2. Clique com o botão direito do mouse no objeto de dados e selecione **Perfil**.

   O assistente **Novo** é exibido.
3. Selecione **Perfil**.
4. Clique em **Avançar**.

   A ferramenta Developer exibe outro painel onde você pode configurar as propriedades gerais do perfil.
5. Altere o nome do perfil e a descrição, se necessário. Você também pode adicionar ou remover objetos de dados.
6. Clique em **Avançar**.
7. Selecione as colunas nas quais você deseja executar a descoberta do domínio de dados e os domínios de dados com os quais você deseja corresponder as colunas.
8. Altere as opções de inferência padrão conforme necessário.
9. Clique em **Concluir** para criar o perfil.

Editando um Perfil no Informatica Developer

Você pode fazer alterações em um perfil após executá-lo para realizar a descoberta do domínio de dados. Você pode excluir colunas com tipos de dados específicos, alterar a seleção de coluna, a seleção de domínio de dados e as opções de inferência.

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

   A definição do perfil é exibida em uma guia.

3. Faça as alterações na seleção de coluna, seleção de domínio de dados e opções de inferência conforme necessário.
4. Na seção **Seleção de Coluna**, você pode clicar em **Excluir Colunas** para configurar opções de exclusão com base em tipos de dados.

   A caixa de diálogo **Excluir Colunas** é exibida.
5. Salve as alterações.

Executando um Perfil para Realizar a Descoberta do Domínio de Dados no Informatica Developer

Você pode optar por executar o perfil imediatamente após criá-lo. Você também pode executar um perfil manualmente após criá-lo.

1. Na exibição **Object Explorer**, selecione o projeto ou pasta que contém o perfil que você deseja executar.

   Para executar um perfil automaticamente, selecione **Executar perfil ao terminar** no assistente de **Novo Perfil** quando você criar o perfil.
2. Clique duas vezes no perfil para abri-lo.

   A definição do perfil é exibida em uma guia.
3. Clique com o botão direito do mouse no perfil e selecione **Executar Perfil**.

   A caixa de diálogo **Executar Perfil** é exibida com o status da execução do perfil.

Resultados da Descoberta do Domínio de Dados no Informatica Developer

Resultados da descoberta do domínio de dados exibem estatísticas sobre colunas que correspondem a domínios de dados, incluindo critérios de conformidade para a correspondência de domínio de dados e se os nomes de colunas correspondem aos domínios de dados.

Você pode fazer uma busca detalhada nos resultados para análise. Pode também verificar os resultados em todas as linhas da fonte de dados e adicionar os resultados a um modelo de dados na Developer tool. Você pode classificar os resultados com base em domínios de dados, grupos de domínios de dados e colunas. Você pode exportar os resultados da descoberta do domínio de dados para um arquivo do Microsoft Excel.

A tabela a seguir descreve os resultados da descoberta do domínio de dados:

Nome da coluna	Descrição
Nome	Nome do domínio de dados, grupo de domínio de dados ou coluna, dependendo se você selecionar a exibição **Domínio de Dados**, **Grupo de Domínio de Dados** ou **Colunas**.
Conexão	Nome da conexão.

Nome da coluna	Descrição
Status	O status de inferência da coluna.
% de conformidade dos dados	A porcentagem mínima de conformidade de linhas necessária para uma correspondência de domínio de dados.
Número de linhas em conformidade	O número mínimo de linhas necessárias para uma correspondência de domínio de dados.
% nulo	A porcentagem de valores nulos para a coluna.
Total de linhas	O número total de linhas.
Correspondência de nome de coluna	Indica se o nome da coluna corresponde a um nome de domínio de dados.
Grupos de domínio de dados	O grupo de domínio de dados ao qual o domínio de dados pertence.
Tipo de dados documentado	Tipo de dados declarado para a coluna no objeto de perfil.
Busca detalhada	Se selecionado, faz uma busca detalhada em linhas.
Verificado	Indica a validação da correspondência de domínio de dados em todas as linhas da fonte de dados.
Horário da última execução	Data e hora da última execução de perfil.

Exibindo por grupos de domínios de dados

Você pode exibir resultados da descoberta do domínio de dados classificados por grupos de domínios de dados.

1. Execute o perfil para exibir seus resultados.
2. Clique em **Resultados**.
3. Clique em **Descoberta do Domínio de Dados**.

   É possível exibir os resultados da descoberta do domínio de dados no painel direito.
4. Verifique se a opção **Domínio de Dados** está selecionada no campo **Mostrar**.
5. Selecione **Mostrar grupo do domínio de dados na hierarquia** para exibir os resultados classificados pelos grupos de domínios de dados.

Exibindo por colunas

Você pode exibir resultados de descoberta do domínio de dados classificados por colunas de origem que correspondem aos domínios de dados.

1. Execute o perfil para exibir seus resultados.
2. Clique em **Resultados**.
3. Clique em **Descoberta do Domínio de Dados**.

   É possível exibir os resultados da descoberta do domínio de dados no painel direito.

4. Selecione **Colunas** para exibir os resultados classificados por colunas de origem que correspondem aos domínios de dados.

Verificando os resultados

Quando você executa um perfil, ele analisa uma amostra da fonte de dados para inferir os resultados do perfil. Você pode executar o perfil em todas as linhas dos dados de origem para verificar os resultados de inferência.

1. Execute o perfil para exibir seus resultados.
2. Clique em **Resultados**.
3. Clique em **Descoberta do Domínio de Dados**.

   É possível exibir os resultados da descoberta do domínio de dados no painel direito.
4. Selecione uma coluna que você queira verificar no painel direito.
5. Clique com o botão direito do mouse na coluna e selecione **Verificar** para executar o perfil em todas as linhas da fonte de dados.

   Você poderá ver uma alteração no valor **% de Conformidade dos Dados** ou **Contagem de linhas em conformidade** depois de verificar os resultados.
6. Para verificar os resultados de inferência de várias colunas, selecione várias colunas. Em seguida, você pode clicar com o botão direito do mouse e selecionar **Verificar Tudo**.

Aprovando Domínios de Dados

Se você executar a descoberta do domínio de dados em um único perfil de objeto de dados, poderá aprovar domínios de dados inferidos para várias colunas de uma vez. Se você executar a descoberta do domínio de dados como parte da descoberta empresarial, poderá aprovar o domínio de dados de uma coluna de origem por vez. Para aprovar os domínios de dados de várias colunas após a descoberta empresarial, você pode abrir as tarefas de perfil de objeto de dados individuais e aprovar os domínios de dados.

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. Se você tiver executado um único perfil de objeto de dados, selecione **Descoberta do Domínio de Dados** e selecione uma linha. A linha contém os resultados da descoberta do domínio de dados para cada coluna.
4. Clique com o botão direito do mouse na linha e selecione **Aceitar**.

   O status de inferência do domínio de dados muda para **Aceito**.
5. Se você executou a descoberta empresarial, selecione a exibição **Domínios de Dados** e, em seguida, selecione um domínio de dados.

   As colunas que correspondem ao domínio de dados aparecem no painel direito.
6. Clique com o botão direito do mouse na coluna que você deseja aprovar e selecione **Aceitar**. Você também pode selecionar várias colunas rejeitadas e aprová-las, conforme necessário.

   O status de inferência do domínio de dados muda para **Aceito**.
7. Para restaurar o status inferido do domínio de dados, clique com o botão direito do mouse na linha e clique em **Redefinir**.

Rejeitando Domínios de Dados

O Informatica Developer exibe domínios de dados inferidos nos resultados do perfil por padrão. Você pode rejeitar domínios de dados inferidos ou aprovados. Você pode optar por exibir ou ocultar os domínios 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 **Descoberta do Domínio de Dados** ou **Domínios de Dados**, selecione uma linha.
4. Para rejeitar os domínios de dados inferidos, clique com o botão direito do mouse na linha e selecione **Rejeitar**.

   O Informatica Developer esmaece o domínio de dados rejeitado dos resultados da descoberta do domínio de dados.
5. Para ocultar os domínios de dados rejeitados, clique com o botão direito do mouse na linha e selecione **Ocultar Rejeitados**.
6. Para exibir os domínios de dados rejeitados, clique com o botão direito do mouse em uma das linhas e selecione **Mostrar Rejeitados**.

Exportando Resultados da Descoberta do Domínio de Dados do Informatica Developer

Ao exportar os resultados da descoberta do domínio de dados para um arquivo `.xlsx` a partir do Informatica Developer, você pode salvar esse arquivo no servidor ou em uma localização específica da máquina cliente.

1. Execute um perfil para realizar a descoberta do domínio de dados.
2. Clique na exibição **Resultados**.
3. Clique no ícone **Exportar Resultados para Arquivo**.

   A caixa de diálogo **Exportar dados para um arquivo** é exibida.
4. Digite o nome do arquivo. Como alternativa, use o nome de arquivo padrão.
5. Em **Salvar**, escolha **Salvar no Cliente** e clique em **Procurar** para selecionar uma localização e salvar o arquivo localmente no seu computador. Por padrão, o Informatica Developer grava o arquivo em uma localização no servidor definida nas propriedades do Data Integration Service do Informatica Administrator.
6. Clique em **OK**.

CAPÍTULO 25

Descoberta Empresarial no Informatica Developer

Este capítulo inclui os seguintes tópicos:

- Descoberta Empresarial no Informatica Developer - Visão Geral, 190
- Processo de Descoberta Empresarial, 191
- Opções de Perfil para Descoberta Empresarial, 192
- Criando um Perfil de Descoberta Empresarial no Informatica Developer, 197
- Editando um Perfil, 198
- Executando um Perfil de Descoberta Empresarial, 199
- Descoberta de Chave Externa, 200
- Análise de Associação, 201
- Descoberta de Sobreposição, 203
- Arquivos de Script DDL, 205

Descoberta Empresarial no Informatica Developer - Visão Geral

A descoberta empresarial é o processo para detectar estatísticas de perfil de coluna, domínios de dados, chaves primárias e chaves externas em um grande número de fontes de dados. Você pode realizar a descoberta empresarial em várias conexões ou esquemas.

Sendo um analista de dados empresariais, você pode querer encontrar características de dados importantes em um grande número de fontes de dados. Os requisitos pode incluir a identificação de ativos de dados relacionais, execuções de perfil de coluna nos ativos de dados descobertos, descoberta de características de dados críticas dentro da empresa, chaves primárias e chaves candidatas. Você também pode querer exibir os relacionamentos de chave externa que existem entre as fontes de dados, para poder derivar um modelo de dados com base nos relacionamentos descobertos.

A descoberta empresarial encontra problemas, padrões, tendências e características de dados críticas nos ativos de informações da sua empresa. Você pode escolher tanto fontes de dados que importou para o repositório do Modelo quanto fontes de dados de conexões relacionais externas. O processo de descoberta de dados inclui descobrir estatísticas de perfil de coluna, análise de domínio de dados, estruturas de objeto de dados, incluindo chaves candidatas, e relacionamentos de objeto de dados que incluem chaves externas.

Você executa a descoberta empresarial na ferramenta Developer e ela realiza as seguintes tarefas em cada fonte de dados:

- Executar um perfil de coluna.
- Descobrir domínios de dados.
- Inferir chaves primárias.

Após a execução de perfis de coluna, descoberta do domínio de dados e perfis de chave primária, a ferramenta Developer executa um perfil de chave externa em todas as fontes de dados. Depois que a ferramenta Developer conclui a criação de perfil e as tarefas de descoberta, gera um resumo dos resultados consolidados nos formatos gráfico e tabular.

Você pode escolher um perfil do sistema operacional no Informatica Developer. Depois que você escolhe um perfil do sistema operacional, o Serviço de Integração de Dados cria e executa os perfis de descoberta empresarial com base na permissão do usuário do sistema operacional que você define no perfil do sistema operacional.

Processo de Descoberta Empresarial

Você pode executar o perfil de descoberta empresarial para realizar a descoberta empresarial na ferramenta Developer. Você precisa configurar as opções de descoberta de dados de diferentes tipos de perfil antes de executar o perfil.

A ferramenta Developer cria objetos de dados para as fontes de dados e as tarefas de perfil selecionadas para cada objeto de dados. A ferramenta executa as tarefas de perfil para gerar resultados de perfil.

Conclua as etapas a seguir para realizar a descoberta empresarial.

1. Crie um perfil de descoberta empresarial selecionando vários objetos de dados importados no repositório do Modelo e fontes de dados em várias conexões relacionais externas.
2. Defina as configurações para descoberta do domínio de dados, perfil de coluna, perfil de chave primária e perfil de chave externa.
3. Execute o perfil de descoberta empresarial.
4. Atualize o Serviço de Repositório do Modelo.

   Nota: Você precisa executar essa ação à medida que a importação de metadados para conexões externas ocorre no repositório do Modelo. Você precisa atualizar o Serviço de Repositório do Modelo de forma que a ferramenta Developer reflita as alterações no repositório do Modelo.
5. Monitore a execução do perfil e, se necessário, exiba os status das tarefas de perfil executadas pela ferramenta Developer.
6. Examine o resumo de resultados da descoberta empresarial. O resumo inclui uma exibição de interface gráfica do usuário interativa e uma exibição tabular.

Opções de Perfil para Descoberta Empresarial

Configure opções de perfil antes de executar um perfil para realizar a descoberta empresarial. As opções de perfil incluem opções de descoberta do domínio de dados, opções de amostragem de perfil de coluna e opções de inferência para chaves primárias e externas.

Você pode optar por executar o perfil de descoberta empresarial após configurar as opções de perfil. Pode também escolher criar tarefas de perfil após a instalação sem executar o perfil.

Seleção de Domínio de Dados para Descoberta Empresarial

As opções de inferência determinam se a descoberta do domínio de dados deve ser executada nos dados da coluna, no nome da coluna ou em ambos. Você poderá especificar se o perfil precisar processar todas as linhas na fonte de dados e escolher critérios de conformidade para a correspondência de domínio de dados.

A tabela a seguir descreve as opções de inferência de domínio de dados que você configura para descoberta empresarial:

Opção	Descrição
Substituir as opções de inferência padrão	Altera as opções de inferência predefinidas.
Dados	O perfil é executado nos dados da coluna.
Nome da coluna	O perfil é executado nos títulos da coluna.
Nome da coluna e dados	O perfil é executado tanto nos dados quanto nos títulos da coluna.
Porcentagem mínima de linhas	A porcentagem mínima de conformidade de linhas no conjunto de dados necessária para uma correspondência de domínio de dados. A porcentagem de conformidade é a proporção entre o número de linhas correspondentes dividido pelo número total de linhas. **Nota:** A Developer tool considera valores nulos como linhas sem correspondência.
Número mínimo de linhas	O número mínimo de linhas no conjunto de dados necessárias para uma correspondência de domínio de dados.
Excluir valores nulos da descoberta do domínio de dados	Exclui os valores nulos do conjunto de dados para a descoberta do domínio de dados.
Todas as linhas	O perfil é executado em todas as linhas da fonte de dados.
Fazer amostragem das primeiras	O número máximo de linhas nas quais o perfil pode ser executado. A Developer tool escolhe as linhas começando da primeira linha na origem.
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 amostragem de perfil de coluna para descoberta empresarial

As opções de amostragem determinam se a ferramenta Developer executa um perfil de coluna em todas as linhas das fontes de dados ou um número limitado de linhas.

A tabela a seguir descreve as opções de amostragem de perfil de coluna que você configura para descoberta empresarial:

Opção	Descrição
Todas as Linhas	Seleciona todas as linhas na fonte de dados.
Primeiras <número> Linhas	O número de linhas nas quais você deseja executar o perfil de coluna. A Developer tool escolhe as linhas começando da primeira linha da fonte de dados.
Excluir a inferência de tipo de dados para colunas com um tipo de dados aprovado	Exclui colunas com um tipo de dados aprovado na inferência de tipo de dados da execução do perfil de coluna.

Opções de ambientes de tempo de execução

Você pode escolher Nativo ou Hadoop como ambiente de tempo de execução para um perfil de coluna. O Informatica Developer 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 descoberta empresarial:

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.
Hadoop	O Serviço de Integração de Dados envia a lógica de perfil por push para o mecanismo Blaze no cluster Hadoop para executar perfis.

Opções de Inferência de Chave Primária para Descoberta Empresarial

Você pode substituir as opções de inferência de chave primária padrão para a descoberta empresarial. As opções incluem o número máximo de linhas nas quais você pode executar o perfil e a porcentagem mínima de conformidade.

A tabela a seguir descreve as opções de inferência de chave primária que você configura para descoberta empresarial:

Opções	Descrição
Substituir as opções de inferência padrão	Permite que você defina configurações personalizadas para inferência de chave primária.
Máximo de Colunas de Chave	Número máximo de colunas que podem constituir uma chave primária.
Máximo de Linhas	Número máximo de linhas nas quais você pode executar o perfil.
Porcentagem Mínima	A porcentagem mínima de conformidade dos dados de colunas necessária para a correspondência de chaves primárias.
Máximo de Linhas de Violação	O número máximo de linhas com violações de chave que o perfil permite ao determinar as chaves primárias.

Opções de Inferência de Chave Estrangeira para Descoberta Empresarial

Configure as opções de inferência de chave externa para definir as configurações de coluna, a fim de descobrir relacionamentos de chave externa entre objetos de dados. Os resultados da inferência de chave externa dependem das opções de inferência de chave primária de opções que você configura para descoberta empresarial, chaves primárias documentadas e chaves primárias definidas pelo usuário.

Você pode inferir as chaves externas no Informatica Developer com um dos seguintes métodos:

- Use valores padrão.
- Configure as opções de inferência de chave externa.
- Use o arquivo de configuração de chaves externas para configurar os parâmetros automáticos de curadoria.

A tabela a seguir descreve as opções de inferência de chave externa que você configura para descoberta empresarial:

Opções	Descrição
Substituir as opções de inferência padrão	Altera as opções de inferência predefinidas.
Tipos de dados usados em comparações	O tipo de dados usados em comparações de chave primária e chave externa. **Nota:** Essa opção se aplica se você executa um perfil de coluna na fonte de dados antes da inferência de chave externa.
Distinção entre maiúsculas e minúsculas na comparação	Inclui a distinção entre maiúsculas e minúsculas ao comparar dados da coluna.

Opções	Descrição
Ajustar valores antes da comparação	Determina se a Developer tool inclui espaços à direita ou à esquerda em dados da coluna durante o processamento.
Chaves primárias inferidas usadas em comparações Usar as chaves com classificação superior	O número de chaves primárias com classificação superior usadas na inferência de chave externa quando a Developer tool executa um perfil de chave externa em todas as fontes de dados. A Developer tool usa o método de classificação superior nas chaves primárias documentadas e definidas pelo usuário para inferir os relacionamentos de chaves externas. A classificação superior das chaves inferidas tem base em uma porcentagem de conformidade decrescente arredondada para a precisão de uma única casa decimal. Por exemplo, a Developer tool considera uma porcentagem de conformidade de 99,75 como 99,8 e 99,74 como 99,7. O valor padrão é 1. Defina o valor como -1 se desejar que a Developer tool use todas as chaves inferidas na inferência de chave externa. **Nota:** Se as fontes de dados de chave primária tiverem chaves primárias aprovadas, a Developer tool não usará chaves primárias inferidas para a inferência de chaves externas.
Número máximo de chaves externas entre objetos de dados	O número máximo de colunas inferidas retornadas pela Developer tool após a execução de perfil e que são necessárias para a descoberta de chaves externas.
Porcentagem mínima de conformidade	O valor mínimo de elegibilidade em porcentagem para incluir colunas nos resultados da chave externa.
Recriar assinatura	Recarrega assinaturas de coluna se os dados de origem mudam.

Parâmetros de curadoria automáticos para inferência de chaves externas

Você pode configurar parâmetros automáticos de curadoria para inferir relacionamentos de chaves primárias e chaves externas sem intervenção manual. Os parâmetros automáticos de curadoria são atributos personalizados definidos pelo usuário que você pode configurar para identificar os relacionamentos de dados com base em determinadas condições.

Quando os resultados da descoberta incluem muitos relacionamentos de chaves primárias e chaves externas, você pode encontrar dificuldades para identificar os relacionamentos de dados críticos entre centenas de relacionamentos de dados. Você também pode encontrar dificuldades para fazer a curadoria dos relacionamentos com base em determinadas condições, como correspondência de dados ou tipo de dados. Para resolver esse problema, você pode configurar os parâmetros de curadoria automático e executar o perfil de descoberta comercial.

Se as fontes de dados tiverem várias chaves externas candidatas e você quiser fornecer regras para escolher uma chave externa candidata, será possível realizar as seguintes ações:

- Configure as opções **Máximo de chaves externas entre objetos de dados** e **Porcentagem de conformidade mínima** no assistente de perfil para descoberta empresarial.
- Configure os pesos e as pontuações para os parâmetros de curadoria automática no arquivo `ForeignKeyConfig.xml`.

Um administrador pode editar e salvar o arquivo de configuração de chave externa. Configure os parâmetros de curadoria automática no arquivo de configuração de chave externa. O algoritmo infere os relacionamentos de chave primária e chave externa entre vários objetos de dados com base nos parâmetros automáticos de curadoria.

O arquivo de configuração de chave externa, `ForeignKeyConfig.xml`, está disponível no seguinte diretório:

```
<diretório de instalação do Informatica>\services\DataIntegrationService\modules
\ProfilingService
```

Os parâmetros automáticos de curadoria são correspondência de dados, correspondência de nomes de colunas, correspondência de tipo de relacionamento e correspondência de tipo de dados.

Correspondência de sobreposição de dados

A correspondência de sobreposição de dados é a sobreposição estimada de valores entre a chave primária e as chaves estrangeiras. Você pode definir a correspondência de sobreposição no assistente de perfil de descoberta corporativa com a opção **Porcentagem mínima de conformidade**. Por padrão, a opção **Porcentagem mínima de conformidade** está definida como 90.

Se a correspondência de sobreposição de dados não atender à porcentagem mínima de conformidade, a chave externa não será levada em consideração para curadoria automática. Quando a conformidade mínima para essa correspondência de dados for atendida, os parâmetros restantes serão usados para calcular a pontuação ajustada.

Correspondência de nomes

O parâmetro de correspondência de nome é um parâmetro opcional. Ele usa o algoritmo Editar Distância para determinar como estreitamente os nomes das colunas de chave primárias e externas correspondem e define a pontuação entre 0 e 1. Defina a ponderação da correspondência de nome como 0 se não quiser usar esse parâmetro para determinar o relacionamento entre chaves primárias e externas.

Correspondência de tipos de relacionamento

A correspondência de tipo de relacionamento determina o tipo de relacionamento entre as colunas de chave primária e chave estrangeira e atribui uma pontuação fixa entre 0 e 1. A correspondência de tipo de relacionamento é calculada com base no tipo de coluna da coluna de chave externa.

As seguinte correspondências de tipos de relacionamento podem ser definidas no arquivo `ForeignKeyConfig.xml`:

- Relação de chave primária/chave externa em que a coluna de chave externa é uma coluna do tipo não chave. O padrão para essa correspondência de tipo de relacionamento é 1. Você pode encontrar esse relacionamento em muitas fontes de dados.
- Relação de chave primária/chave primária em que a coluna de chave externa é uma coluna do tipo chave primária. O padrão para esse tipo de relacionamento é 0,25. Você raramente encontra esse tipo de relacionamento, pois ele representa uma tabela que foi particionada verticalmente.
- Relacionamento de sequência de chave primária/chave primária em que a coluna de chave externa é uma coluna de chave primária e o tipo de dados da coluna é um tipo de dados de sequência. Por exemplo, a coluna OrderID em uma tabela de Order tem um tipo de dados de sequência. O padrão para esse tipo de relacionamento é zero, pois as chaves de sequência podem causar várias chaves externas falso positivas, que o algoritmo de chave primária/chave primária tenta evitar. Você pode definir a correspondência de tipos de relacionamento como uma maior pontuação se a fonte de dados for conhecida por conter alguns tipos de dados de sequência.

Correspondência de tipos de dados

A correspondência de tipos de dados compara os tipos de dados das colunas de chave primária com as colunas de chave externa e atribui uma pontuação de conformidade fixa que se baseia na proximidade de correspondência dos tipos de dados das colunas.

A seguinte tabela lista as pontuações de correspondência de tipos de dados fixos para diferentes combinações de chaves primárias e chaves externas:

	Chave externa numérica	**Chave externa de data**	**Chave externa de cadeia**
Chave primária numérica	1.0	0.5	0.0
Chave primária de data	0.5	1.0	0.5
Chave primária de cadeia	0.0	0.0	1.0

Se necessário, você pode alterar as pontuações de correspondência de tipos de dados padrão.

Criando um Perfil de Descoberta Empresarial no Informatica Developer

Você pode criar um perfil em várias fontes de dados com várias conexões. A ferramenta Developer cria tarefas de perfil individuais para cada origem.

1. Na exibição **Object Explorer**, selecione vários objetos de dados nos quais você queira executar um perfil.
2. Clique em **Arquivo > Novo > Perfil** para abrir o assistente de perfil.
3. Selecione **Perfil de Descoberta Empresarial** e clique em **Próxima**.
4. Insira um nome para o perfil e verifique a localização do projeto. Se necessário, navegue para uma nova localização.
5. Verifique se o nome dos objetos de dados selecionados aparece na seção **Objetos de Dados**. Clique em **Escolher** para selecionar mais objetos de dados, se necessário.
6. Clique em **Avançar**.

   O painel **Adicionar Recursos à Definição de Perfil** é exibido. Você pode selecionar várias conexões relacionais externas e fontes de dados nesse painel.
7. Clique em **Escolher** para abrir a caixa de diálogo **Selecionar Recursos**.

   O painel **Recursos** lista todas as conexões internas e externas e objetos de dados sob o domínio Informatica.
8. Clique em **OK** para fechar a caixa de diálogo.
9. Clique em **Avançar**.
10. Configure os tipos de perfil que você deseja executar. Você pode configurar os seguintes tipos de perfil:
    - Descoberta do domínio de dados
    - Perfil de coluna
    - Perfil de chave primária

- Perfil de chave externa

Nota: Selecione **Ativado como parte da ação "Executar Perfil de Descoberta Empresarial"** para os tipos de perfil que você deseja executar como parte do perfil de descoberta empresarial. A criação de perfil da coluna é habilitada por padrão.

11. Examine as opções para o perfil.

    Você pode editar as opções de amostragem dos perfis de coluna. Você pode também editar as opções de inferência do domínio de dados, da chave primária e dos perfis de chave externa.

12. Selecione **Criar Perfis**.

    A ferramenta Developer cria perfis para cada fonte de dados individual.

13. Selecione **Executar perfil de descoberta empresarial ao concluir** para executar o perfil quando você concluir a configuração do perfil. Se você ativou todas as operações de criação de perfil, a ferramenta Developer executa perfis de coluna, domínio de dados e chave primária em todas as fontes de dados selecionadas. Em seguida, a ferramenta Developer executa um perfil de chave externa em todas as fontes de dados.

14. Clique em **Concluir**.

    Depois de executar um perfil de descoberta empresarial, você precisa atualizar o Serviço de Repositório do Modelo antes de exibir os resultados. Esta etapa é necessária, pois a importação dos metadados para conexões externas acontece no repositório do Modelo. Você precisa atualizar o Serviço de Repositório do Modelo de forma que a ferramenta Developer reflita as alterações no repositório do Modelo.

Editando um Perfil

Você pode fazer alterações em um perfil de descoberta Empresarial depois de configurá-lo. Você pode excluir colunas com tipos de dados específicos, alterar a seleção de coluna, a seleção de domínio de dados e as opções de inferência.

1. Na exibição **Object Explorer**, selecione o projeto ou pasta que contém o perfil que você deseja editar.
2. Clique em **Equipe** > **Fazer Check-out** para fazer check-out do perfil.
3. Clique duas vezes no perfil para abri-lo.
4. Clique na exibição **Propriedades**.

   A exibição Propriedades está sob a exibição Padrão.

5. Clique em **Perfis** para exibir as tarefas do perfil.
6. Selecione um perfil de tarefa que você queira editar no painel direito e clique em **Abrir** .

   A definição do perfil é exibida em uma guia.

7. Para fazer alterações nas configurações globais do perfil de descoberta empresarial, selecione o perfil na parte superior da **Lista de tarefas de criação de perfil** e clique em **Configurar**.
8. Faça as alterações necessárias nas opções de definição de perfil .
9. Salve as alterações.
10. Clique em **Equipe** > **Fazer Check-In** para fazer o check-in do perfil.

Executando um Perfil de Descoberta Empresarial

Você pode executar um perfil de descoberta empresarial de várias maneiras. Você pode executar o perfil a partir da exibição **Object Explorer** ou da guia **Perfis** na janela **Propriedades**. Você pode optar por executar uma ou várias tarefas de perfil que compõem uma parte da descoberta de perfil empresarial.

1. Na exibição **Object Explorer**, selecione o projeto ou pasta que contém o perfil que você deseja executar.

   Para executar um perfil automaticamente, selecione **Executar o perfil de descoberta empresarial ao concluir** no assistente **Nova Descoberta Empresarial** ao criar o perfil.

2. Clique duas vezes no perfil para abri-lo.

   O perfil é aberto em uma guia.

3. Na exibição **Object Explorer**, clique com o botão direito do mouse no perfil e selecione **Executar Perfil de Descoberta Empresarial**.

   Como alternativa, você pode selecionar **Perfis** na janela **Propriedades**, selecionar o nome do perfil na **Lista de tarefas de criação de perfil** e clicar em **Executar**.

   Nota: Ao executar um perfil de descoberta empresarial, você deve atualizar o Serviço de Repositório do Modelo antes de poder exibir os resultados. Esta etapa é necessária, pois a importação dos metadados para conexões externas acontece no repositório do Modelo. Você precisa atualizar o Serviço de Repositório do Modelo de forma que a ferramenta Developer reflita as alterações no repositório do Modelo.

4. A caixa de diálogo **Executar** é exibida. Você pode alterar as configurações globais do perfil nesta caixa de diálogo.

   Por padrão, as alterações feitas por você são aplicadas aos objetos de dados recém-adicionados no perfil de descoberta de dados.

5. Para aplicar as alterações a todas as tarefas de perfil do objeto de dados e à tarefa de perfil da chave externa gerada pelo sistema no perfil de descoberta empresarial, selecione **Usar configurações globais para perfis atuais**.

   A ferramenta Developer atualiza todas as tarefas de perfil do objeto de dados e a tarefa de perfil da chave externa com base nas configurações atualizadas.

6. Para executar tarefas de perfil individuais, selecione uma tarefa e clique em **Executar**.

7. Para executar várias tarefas de perfil, clique em **Executar Vários**.

   A caixa de diálogo **Executar Vários** é exibida.

   Sugestão: Se os resultados da descoberta empresarial demorarem muito para carregar, convém atualizar as estatísticas do banco de dados do depósito de criação de perfil. Várias execuções de perfil de descoberta empresarial podem resultar em alterações significativas nos valores de volume de dados e coluna. Quando você atualiza as estatísticas, o banco de dados realiza um plano de execução para as consultas SQL com base nas últimas estatísticas e otimiza as operações de banco de dados.

8. Todas as tarefas estão selecionadas por padrão. Limpe as tarefas que você não deseja executar e clique em **OK**.

Descoberta de Chave Externa

Uma coluna é uma chave externa se seus valores de dados correspondem aos valores da coluna de chave primária em outro objeto de dados.

Você pode executar a descoberta de chave externa em vários objetos de dados na ferramenta Developer. Crie um perfil de descoberta empresarial para selecionar objetos de dados e definir o perfil.

Antes de executar a descoberta de chave externa, você deve identificar os objetos de dados pai e filho no perfil de descoberta empresarial. O perfil usa uma ou mais chaves no objeto pai, incluindo sua chave primária, para descobrir chaves externas no objeto filho. Após definir os objetos pai e filho e identificar as chaves no objeto pai, você cria e executa o perfil.

Definindo Relacionamentos de Objeto Pai e Filho

Para localizar relacionamentos de chave externa entre dois objetos de dados, você deve selecionar um objeto de dados pai e especificar a chave primária nesse objeto.

1. Abra um perfil de descoberta empresarial que contenha os objetos de dados que você deseja analisar.
2. Selecione o objeto pai.
3. Selecione a chave primária no objeto pai:
   - Clique na guia **Propriedades** e clique em **Chaves**.
   - Clique em **Adicionar** e selecione a coluna de chave primária na caixa de diálogo Nova Chave.
   - Clique em **OK** na caixa de diálogo **Nova Chave**. Verifique se a chave primária é exibida no painel **Campos selecionados** e se a opção **Chave Primária** está marcada.

Crie um perfil de chave externa para analisar o objeto filho para chaves externas.

Descobrindo Relacionamentos de Chave Externa entre Objetos de Dados

Use um perfil de descoberta empresarial na ferramenta Developer para localizar relacionamentos de chave entre dois objetos de dados

O objeto de dados que contém a chave primária é o objeto pai, e o objeto de dados que contém a chave externa é o objeto filho.

1. Abra um perfil de descoberta empresarial que contenha os objetos de dados que você deseja analisar.
2. Clique com o botão direito do mouse no nome de um objeto de dados e selecione **Perfil de Chave Externa**.
3. Insira um nome para o perfil e verifique a localização do projeto. Se necessário, navegue para uma nova localização. Opcionalmente, insira um texto de descrição do perfil.
4. Selecione as chaves no objeto pai que o perfil usará para localizar chaves externas no objeto filho.
5. Salve e execute o perfil.

Resultados de Análise de Chave Externa

Depois de executar um perfil de chave externa, clique no nome do perfil abaixo do editor de modelagem para ver os resultados da análise.

A exibição de resultados lista as colunas que atendem aos critérios de inferência de chave primária-externa definidos por você. Clique no botão **Opções** para editar as configurações de inferência. Clique em um nome de coluna e selecione **Validar** para verificar se uma chave inferida é válida para os objetos de dados.

A tabela a seguir descreve as propriedades da análise de chave externa:

Propriedade	Descrição
Chave Primária Pai	Uma coluna de chave primária no objeto de dados pai que o perfil usa para localizar as chaves externas em um objeto filho.
Chave Externa Filho	Uma coluna que o perfil infere como uma chave externa para a chave primária pai na linha atual.
Inclusão %	A quantidade de valores de dados com correspondência entre a chave primária e a externa, expressa como uma porcentagem. **Nota:** Você pode ver uma variação no valor da % de Inclusão para uma coluna inferida nos resultados da chave externa e depois de validá-la. Para uma coluna inferida, % de Inclusão é o número de valores exclusivos de coluna de chave externa de um objeto filho que correspondem aos valores exclusivos de coluna de chave primária do objeto pai. Depois que você valida uma coluna inferida, ela é o número de valores de coluna de chave externa valores de um objeto filho que correspondem aos valores de coluna de chave primária do objeto pai.
Tipo de Relacionamento	O tipo de relacionamento definido para as colunas de chave primária e externa antes de o perfil ser executado. Se você definir um relacionamento antes de o perfil ser executado, o perfil retornará dados para o relacionamento mesmo se o número de porcentagem de inclusão não atingir o limite de confiança definido para o perfil.
Verificado	Indica que um usuário validou o relacionamento de chave primária-externa.
Horário da última execução	Data e hora da última execução do perfil.
Tipo de Relacionamento (no Modelo)	Indica que o perfil verificou o relacionamento entre as colunas.

Análise de Associação

A análise de associação descreve o grau de possíveis associações entre duas colunas de dados. Use um perfil de associação para analisar associações de colunas em uma fonte de dados ou entre várias fontes de dados.

Um perfil de associação exibe os resultados como um diagrama de Venn e como valores numéricos e percentuais. Você pode criar e executar um perfil de associação de um perfil de descoberta empresarial.

Criando um Perfil de Associação

Você pode analisar possíveis associações entre objetos de dados em um perfil de descoberta empresarial. O perfil de associação armazena a análise no repositório do Modelo.

1. Crie ou abra um perfil de descoberta empresarial.
2. Verifique se o perfil de descoberta empresarial contém os objetos de dados que você precisa.

   Para adicionar um objeto de dados ao perfil de associação, arraste-o da exibição **Object Explorer** para o editor de modelagem.
3. Selecione os objetos de dados para o perfil.
4. Clique com o botão direito do mouse nos objetos e selecione **Perfil de Junção**.

   O assistente de perfil é aberto.
5. Insira um nome para o perfil. Opcionalmente, insira um texto de descrição do perfil.
6. Verifique se os nomes dos objetos de dados são exibidos sob **Objetos de Dados** no assistente.
7. Selecione ou limpe a opção **Executar perfil ao terminar**.
8. Clique em **Avançar**.
9. Selecione as colunas de dados a serem incluídas e clique em **Avançar**.

   Se necessário, role os objetos de dados para baixo para exibir todas as colunas disponíveis. O perfil é executado em todas as colunas por padrão.
10. Clique em **Adicionar**.

    A caixa de diálogo **Condição de Associação** aparecerá.
11. Clique em **Novo** para ativar os campos de seleção de coluna.
12. Selecione os objetos de dados e as colunas para validar.

    Você define uma condição de junção entre duas colunas. Você pode definir várias condições de junção em um ou mais objetos de dados.
13. Clique em **OK** para criar a condição de junção.

    Opcionalmente, clique em **Adicionar** para definir condições adicionais.
14. Verifique se as colunas de junção Esquerda e Direita estão prefixadas com os nomes dos objetos de dados corretos.
15. Clique em **Concluir**.

Resultados da Análise de Associação

A guia **Resultados** da análise de associação fornece informações sobre o número e a porcentagem de linhas órfãs pai, linhas órfãs filho e linhas de associação. Os resultados da análise de associação também incluem diagramas de Venn que mostram os relacionamentos entre colunas.

A seguinte tabela descreve as propriedades mostradas na guia **Resultados**:

Propriedade	Descrição
Tabela Esquerda	Nome da tabela esquerda e das colunas usadas na análise de associação.
Tabela Direita	Nome da tabela direita e das colunas usadas na análise de associação.
Linhas Somente à Esquerda	Número de linhas na tabela esquerda que não podem ser associadas.

Propriedade	Descrição
Linhas Somente à Direita	Número de linhas na tabela direita que não podem ser associadas.
Linhas de Junção	Número de linhas incluídas na junção.

Selecione uma condição de junção para exibir um diagrama de Venn que mostra os relacionamentos entre colunas. A área abaixo do diagrama de Venn também exibe o número e a porcentagem de valores órfãos, nulos e associados nas colunas.

Clique duas vezes em uma seção no diagrama de Venn para exibir os registros que a seção representa. Esses registros são abertos na exibição Visualizador de Dados.

Nota: Você pode exportar a lista de registros da exibição Visualizador de Dados para um arquivo simples.

Exportando Resultados do Perfil de Associação para um Arquivo

Você pode exportar as linhas de dados retornadas para uma condição de associação para um arquivo delimitado. Exporte as linhas de sobreposição entre as origens esquerda e direita ou as linhas órfãs de uma origem.

1. Na exibição **Object Explorer**, abra o perfil de descoberta empresarial que contém a análise de associação.
2. Execute o perfil de associação.
3. Selecione a exibição **Resultados da Associação**.
4. Na guia **Visualizador de Dados**, clique no ícone **Exportar Resultados de Busca Detalhada para Arquivo**.

   A caixa de diálogo **Exportar Dados** é exibida.
5. Insira um nome de arquivo e clique em **Salvar**.

Descoberta de Sobreposição

A descoberta de sobreposição fornece informações sobre dados de sobreposição em pares de colunas dentro de uma ou várias fontes de dados. Você pode localizar dados de sobreposição em um perfil de descoberta empresarial. Pode também validar os resultados do perfil e exibi-los em um diagrama de Venn.

A descoberta de sobreposição identifica dados de sobreposição com base nas configurações padrão ou nas configurações que você especificar. Você pode substituir as configurações padrão e especificar opções de inferência, incluindo o número máximo de pares superiores retornados pela descoberta de sobreposição com base na porcentagem de sobreposição. Você também pode especificar um nível de segurança que determine a elegibilidade à descoberta de sobreposição.

Resultados da descoberta de sobreposição

A guia **Descoberta de Sobreposição** exibe informações sobre as colunas participantes e o valor da porcentagem de sobreposição. Os resultados da descoberta de sobreposição incluem diagramas de Venn que representam os dados de sobreposição em pares de colunas e a data e hora quando você executou a descoberta de sobreposição pela última vez.

Você pode clicar em uma coluna e selecionar **Verificar** para exibir os resultados como um diagrama de Venn.

A tabela a seguir descreve as propriedades da descoberta de sobreposição:

Propriedade	Descrição
Coluna Esquerda	A coluna principal em relação à qual as colunas restantes são comparadas para análise de sobreposição.
Coluna Direita	A coluna que é comparada à coluna principal.
% de Sobreposição	A porcentagem de sobreposição entre duas colunas.
Verificado	Indica que você validou a linha de resultados de sobreposição.
Horário da Última Execução	A data e a hora em que a descoberta de sobreposição foi executada pela última vez.

O Informatica Developer exibe cada par de sobreposição duas vezes nos resultados da descoberta de sobreposição. Considere as fontes de dados Itens e Pedidos. Os itens têm colunas "m" e "n". Os pedidos tem colunas "p" e "q".

A seguinte tabela mostra os resultados da descoberta de sobreposição para Itens e Pedidos:

Coluna Esquerda	Coluna Direita
Itens	-
m	Pedidos.p
m	Pedidos.q
n	Pedidos.p
n	Pedidos.q
Pedidos	-
p	Itens.m
p	Itens.n
q	Itens.m
q	Itens.m

Descobrindo Dados de Sobreposição

Você pode determinar dados de sobreposição entre pares de colunas em um perfil de descoberta empresarial. A análise de sobreposição é baseada em valores exclusivos nas colunas e não considera valores nulos.

1. Crie ou abra um perfil de descoberta empresarial que contenha os objetos de dados.
2. Selecione os objetos de dados nos quais você deseja localizar dados de sobreposição.

   É possível selecionar um único objeto de dados para localizar dados de sobreposição dentro de pares de colunas ou vários objetos de dados.

3. Clique com o botão direito do mouse nos objetos e selecione **Descoberta de Sobreposição**.

   A caixa de diálogo **Nova Descoberta de Sobreposição** é exibida.
4. Digite um nome.
5. Opcionalmente, insira um texto de descrição da análise de sobreposição.
6. Verifique se os nomes dos objetos de dados aparecem sob **Objetos de Dados** no assistente.
7. Opcionalmente, selecione **Executar perfil ao terminar** para executar o perfil quando você concluir a definição das configurações.
8. Clique em **Avançar**.
9. Selecione as colunas para a descoberta de sobreposição.
10. Clique em **Avançar**.

    As opções de inferência padrão são exibidas na caixa de diálogo.
11. Opcionalmente, especifique as opções de inferência da descoberta de sobreposição para substituir as configurações padrão.
12. Clique em **Concluir**.

Arquivos de Script DDL

Os arquivos de script DDL (Data Definition Language) contêm as instruções SQL `Create`, `Alter` e `Drop`.

Você pode especificar um nome de arquivo, localização e tipo de banco de dados de destino quando gera os arquivos de script. A ferramenta Developer acrescenta os nomes de arquivo de script com os rótulos "_create" e "_drop". Colunas virtuais não fazem parte dos arquivos de script DDL.

Criando Scripts DDL de um Perfil de Descoberta Empresarial

Ao gerar arquivos de script DDL de um perfil de descoberta empresarial, você pode escolher a localização onde deseja salvar os arquivos de script. Você também pode escolher o tipo de banco de dados no qual deseja executar os scripts. Certifique-se de verificar e confirmar todas as alterações necessárias no perfil de descoberta empresarial antes de gerar scripts DDL.

1. Na exibição **Object Explorer**, selecione um perfil de descoberta empresarial.
2. Clique com o botão direito do mouse no perfil e selecione **Gerar DDL**.

   A caixa de diálogo **Gerar DDL** é exibida.
3. Clique em **Procurar** para abrir a caixa de diálogo **Salvar Como**.

   A extensão de arquivo padrão é `.sql`.
4. Escolha uma localização para o arquivo e digite um nome.
5. Selecione o tipo de banco de dados de destino.
6. Clique em **OK**.

   A ferramenta Developer gera os arquivos de script DDL no local especificado.

CAPÍTULO 26

Resultados da Descoberta Empresarial

Este capítulo inclui os seguintes tópicos:

- Resultados da Descoberta Empresarial - Visão Geral, 206
- Exibição Relacionamentos, 207
- Exibição Criação de Perfil de chave Externa, 208
- Exibição Tabular, 211
- Exibição Domínios de Dados, 213
- Exibição Perfil de Coluna, 214
- Exibindo Resultados do Perfil de Coluna Durante a Execução da Descoberta Empresarial, 214
- Exibindo Resultados da Descoberta do Domínio de Dados Durante a Execução da Descoberta Empresarial, 215
- Exibindo o Status da Descoberta Empresarial em Tempo de Execução, 215
- Arquivos de Exportação de Descoberta Empresarial, 216

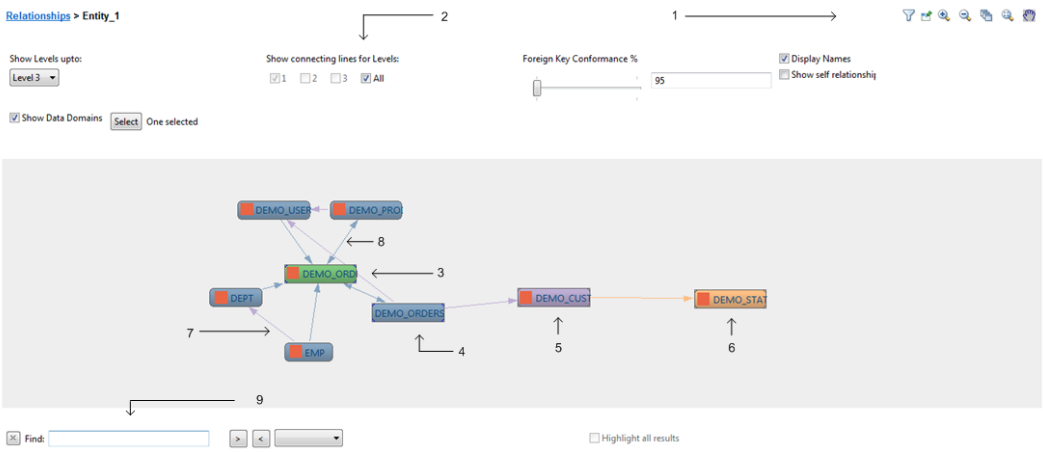
Resultados da Descoberta Empresarial - Visão Geral

É possível visualizar os resultados da descoberta empresarial em várias exibições. As exibições são **Relacionamentos**, **Domínios de Dados**, **Perfil de Coluna**, **Perfil de Associação** e **Descoberta de Sobreposição**.

A exibição **Relacionamentos** mostra grupos de objetos de dados como círculos. Você pode iniciar os resultados do perfil de chave externa a partir dessa exibição. É possível visualizar os resultados do perfil de chave externa em exibições gráfica e tabular. A exibição **Domínios de Dados** mostra os resultados da descoberta do domínio de dados. A exibição **Perfil de Coluna** mostra resultados do perfil de coluna para cada objeto de dados. A exibição **Perfil de Junção** mostra o número de linhas órfãs pai, linhas órfãs filho e linhas incluídas na junção. A exibição **Descoberta de Sobreposição** fornece informações sobre as colunas participantes e o valor da porcentagem de sobreposição.

Objetos de dados podem ter vários relacionamentos entre eles. A exibição gráfica dos resultados da chave externa mostra os relacionamentos de objeto de dados que têm a mais alta porcentagem de conformidade.

A imagem a seguir mostra a exibição gráfica de alguns objetos de dados de amostra dos resultados da descoberta empresarial:



1. Ícones da barra de ferramentas que incluem filtro, fixar objeto de dados, ampliar e reduzir o zoom, organizar tudo, ajustar à janela e panorâmica.
2. Opções de filtro, como exibir níveis diferentes de relacionamentos de objeto de dados, mostrar linhas de conexão e mostrar objetos de dados auto-relacionados.
3. O objeto de dados selecionado com base no qual o editor visual exibe o restante dos relacionamentos entre objetos de dados.
4. Primeiro nível de relacionamentos de objeto de dados.
5. Segundo nível de relacionamentos de objeto de dados.
6. Terceiro nível de relacionamentos de objeto de dados.
7. Conector entre objetos de dados. Uma cabeça de seta única indica um relacionamento entre chave primária e chave externa entre dois objetos de dados. A cabeça de seta aponta para o objeto de dados com a chave primária.
8. Conector entre objetos de dados. Um conector de cabeça de seta dupla indica um relacionamento de chave primária para chave primária entre dois objetos de dados. Passe o cursor do mouse sobre o conector para exibir as colunas de associação com conformidade máxima para o relacionamento inferido.
9. Pressione CTRL+F para exibir o campo Localizar e use um asterisco (*) como um caractere curinga para localizar objetos de dados na exibição gráfica.

Exibição Relacionamentos

Você pode ver um resumo dos resultados da descoberta empresarial na exibição **Relacionamentos**, incluindo entidades. Entidades são grupos de objetos de dados representados como círculos. As entidades incluem objetos de dados relacionados, auto-relacionados e não relacionados das várias conexões e esquemas dos bancos de dados de origem.

Um objeto de dados auto-relacionado tem colunas dentro do objeto de dados que têm relacionamentos. Um objeto de dados não relacionado não tem relacionamentos com outros objetos de dados no banco de dados de origem, nem tem um relacionamento entre colunas dentro do objeto de dados. O diagrama de relacionamentos de entidades dos objetos de dados nos resultados de descoberta empresarial baseia-se nos relacionamentos inferidos e não nos relacionamentos documentados nas fontes de dados.

Pesquisando um Objeto de Dados

Você pode pesquisar um objeto de dados na exibição **Relacionamentos** ou **Criação de Perfil de Chave Externa**. Você pode usar um asterisco (*) como caractere curinga para localizar objetos de dados.

1. Verifique se você está na exibição **Relacionamentos** ou **Criação de Perfil de Chave Externa**.

2. Digite uma parte do nome do objeto de dados que você está pesquisando e adicione o caractere curinga ***** ao início ou fim da string de pesquisa com base no requisito da pesquisa. Por exemplo, para pesquisar todos os objetos de dados começando com a string "CA", digite "CA*" e pressione a tecla **Enter**. Para pesquisar todos os objetos de dados que tenham a string "ZIP" em seus nomes, digite "*ZIP*".

   A pesquisa diferencia maiúsculas de minúsculas.

Navegando até a Exibição Criação de Perfil de Chave Externa

A exibição **Criação de Perfil de Chave Externa** mostra uma exibição consolidada dos relacionamentos de chave externa entre vários objetos de dados nos quais você executou o perfil. Os círculos na exibição representam entidades, objetos auto-relacionados e objetos não referenciados.

1. Verifique se você está na guia **Relacionamentos**.

   Você pode exibir o link do perfil de chave externa link no painel direito.

2. Clique em **Perfil de Chave Externa** para abrir a exibição.

   A exibição aparece em uma nova guia. A exibição exibe objetos de dados em diferentes conjuntos de círculos com base em seu tipo de relacionamento. Você também pode ver o número total de objetos de dados que são uma parte da exibição de chave externa consolidada.

3. Opcionalmente, você pode clicar no link **Relacionamentos** para voltar à exibição **Relacionamentos**.

Exibição Criação de Perfil de chave Externa

Você pode exibir um resumo dos resultados da descoberta empresarial em um formato gráfico na exibição **Criação de Perfil de Chave Externa**. Você pode abrir os resultados do perfil e os relacionamentos em nível de coluna de um objeto de dados em um formato tabular da exibição.

Os resultados do perfil de um objeto de dados incluem o perfil de coluna, a inferência de chave primária, a inferência de dependência funcional e os resultados da descoberta do domínio de dados. Depois de abrir os relacionamentos em nível de coluna de um objeto de dados, você pode verificar e fazer curadoria dos relacionamentos de dados. Quando você verifica um relacionamento de dados, a ferramenta Developer executa o perfil em todas as linhas dos dados de origem para verificar os resultados de inferência. Você pode aprovar, rejeitar e redefinir os relacionamentos de dados na exibição **Criação de Perfil de Chave Externa**.

Nota: A exibição **Criação de Perfil de Chave Externa** não exibe objetos de dados quando você usar fontes de dados Hive para criar um perfil de descoberta empresarial.

Exibindo Relacionamentos de Objeto de Dados

Você pode exibir os relacionamentos entre objetos de dados em um formato gráfico. Clique duas vezes no círculo de uma entidade para exibir as tabelas e seus relacionamentos.

1. Verifique se você está na exibição **Criação de Perfil de Chave Externa**.
2. Para incluir domínios de dados na exibição gráfica consolidada, selecione **Mostrar Domínios de Dados**.

   O botão **Selecionar** será ativado.

3. Clique em **Selecionar** para escolher os domínios de dados que você deseja incluir na exibição gráfica.

   A caixa de diálogo **Criar Domínio de Dados** será exibida.

4. Selecione os domínios de dados necessários e clique em **OK**.

   A ferramenta Developer destaca os círculos de entidade que incluem os domínios de dados que você selecionou.

5. Clique duas vezes no círculo de uma entidade para exibir a representação visual dos relacionamentos de tabela dentro da entidade. A ferramenta Developer exibe as tabelas em um formato gráfico que representa o relacionamento que cada objeto de dados tem com outros objetos de dados dentro da entidade.

   O objeto de dados com o número máximo de relacionamentos com outros objetos de dados ou o objeto de dados de onde você iniciou a navegação é destacado na cor Verde. Se você incluiu domínios de dados, a ferramenta Developer destaca a seleção de domínios de dados à esquerda da representação visual de cada objeto de dados.

6. Verifique as informações de relacionamento direto e as informações de domínio de dados no painel direito da exibição **Criação de Perfil de Chave Externa**.

7. Como alternativa, você pode clicar no link **Relacionamentos** para voltar à exibição **Relacionamentos**.

Ampliando e Reduzindo o Zoom da Exibição

Você pode ampliar o zoom à representação gráfica de relacionamentos de objetos de dados na exibição **Criação de Perfil de Chave Externa** para obter mais clareza visual. Quando você amplia o zoom, a ferramenta Developer aumenta o nível de ampliação da imagem. Reduza o zoom para reduzir o nível de ampliação.

1. Verifique se você está na exibição **Criação de Perfil de Chave Externa**.

2. Clique com o botão direito do mouse na exibição e selecione **Ampliar Zoom** para aumentar o nível de ampliação de imagem.

3. Para reduzir o nível de ampliação do layout gráfico, clique com o botão direito do mouse na exibição e selecione **Reduzir Zoom**.

Localizando um Objeto de Dados

Você pode pesquisar e localizar um objeto de dados na exibição gráfica dos resultados de chave externa. Use um asterisco (*) como um caractere curinga para localizar objetos de dados.

1. Verifique se você está na exibição **Criação de Perfil de Chave Externa**.

   Lembre-se de abrir a exibição gráfica dos resultados de chave externa na exibição.

2. Pressione Ctrl+F para exibir o campo **Localizar**.

3. No campo **Localizar**, digite uma parte do nome do objeto de dados que você está pesquisando e adicione o caractere curinga ***** ao início ou fim da string de pesquisa com base no requisito da pesquisa. Por exemplo, para pesquisar todos os objetos de dados começando com a string "EMP", digite "EMP*" e pressione a tecla **Enter**. Para pesquisar todos os objetos de dados que tenham a string "ZIP" em seus nomes, digite "*ZIP*".

4. Clique no botão **Próxima Correspondência** para ir para a próxima correspondência de objeto de dados.

   Use o botão **Correspondência Anterior** para ir para a correspondência de objeto de dados anterior.

5. Selecione **Realçar todos os resultados** para realçar todas as correspondências de objeto de dados.

6. Para limpar a string de pesquisa no campo **Localizar**, clique no botão **Limpar** ao lado do campo.

Exibindo Relacionamentos de Colunas

Você pode exibir o relacionamento de cada coluna em um objeto de dados com colunas em objetos de dados relacionados. Você também pode verificar e confirmar o relacionamento do objeto de dados com o modelo de dados.

1. Verifique se você está na exibição **Criação de Perfil de Chave Externa**.
2. Clique com o botão direito do mouse em um objeto de dados e selecione **Exibir Relacionamentos de Colunas**.

   Os relacionamentos das colunas aparecem em uma exibição tabular. A exibição exibe informações de relacionamento, como objeto de dados de origem, objeto de dados relacionado e colunas no objeto de dados relacionado.
3. Verifique o status de inferência, o status de verificação e o status da curadoria.
4. Selecione **Todos os objetos de dados no grupo** para exibir todos os objetos de dados na entidade pai e suas informações de relacionamento de coluna.

   Por padrão, a exibição exibe informações de relacionamento do objeto de dados selecionado na exibição.
5. Como alternativa, você pode clicar no link da **entidade** na parte superior da exibição para voltar à representação gráfica dos objetos de dados.

Salvando o Diagrama de Relacionamento de Entidades como uma Imagem

Você pode salvar o diagrama de relacionamento de entidades dos objetos de dados dos resultados da descoberta empresarial como um arquivo ".png".

1. Execute um perfil para realizar a descoberta empresarial.
2. Verifique se você está na exibição **Criação de Perfil de Chave Externa**.
3. Alterne para o formato gráfico dos relacionamentos de objeto de dados na exibição.
4. Clique com o botão direito do mouse e selecione **Salvar como Imagem**.

   A caixa de diálogo **Salvar Como** é exibida. A imagem é salva como um arquivo ".png" por padrão.
5. Escolha a localização para o arquivo e digite um nome.
6. Clique em **Salvar**.

Exibindo Resultados do Perfil de Objeto de Dados na Exibição Criação de Perfil de Chave Externa

Você pode exibir o perfil de coluna, a chave primária e os resultados da descoberta do domínio de dados de um objeto de dados selecionado na exibição **Criação de Perfil de Chave Externa**. Lembre-se de escolher o objeto de dados fixando-o como a tabela selecionada na tela.

1. Verifique se você está na exibição **Criação de Perfil de Chave Externa**.
2. Clique com o botão direito do mouse em um objeto de dados e selecione **Fixar Objeto de Dados como Foco** para selecionar a tabela.

   Alternativamente, use o ícone de **Pino** para selecionar um objeto de dados.
3. Clique com o botão direito do mouse em qualquer lugar da tela e selecione **Exibir Perfil de Objeto de Dados**.

   Os resultados do perfil de objeto de dados são exibidos em uma guia.

Exibição Tabular

Quando você abre a exibição **Criação de Perfil de Chave Externa**, a ferramenta Developer mostra uma exibição gráfica dos resultados por padrão. Alterne para uma exibição tabular para ver as tabelas e seus detalhes de relacionamento em formato tabular.

Você pode exibir o número de objetos de dados na entidade, nomes das tabelas relacionadas, suas informações de conexão e o número de relacionamentos entre os dois objetos de dados. Você também pode verificar os relacionamentos de coluna e adicioná-los a um modelo de dados.

Painel Detalhes da Tabela

Você pode ver detalhes do objeto de dados nos resultados da descoberta empresarial em formato gráfico e tabular. Na exibição gráfica, o painel de detalhes da tabela exibe o número de objetos de dados que têm um relacionamento direto com o objeto de dados selecionado e os nomes de objetos de dados.

A tabela a seguir descreve as colunas do painel de detalhes da tabela na exibição tabular:

Nome da Coluna	Descrição
Nome da Tabela	Nome do objeto de dados que tem um relacionamento direto com o objeto de dados selecionado no painel esquerdo.
Conexão	Nome da conexão para o objeto de dados relacionado.
Relacionamentos	O número de relacionamentos entre o objeto de dados selecionado no painel esquerdo e o objeto de dados relacionado no painel de detalhes da tabela.

Verificando os Resultados da Descoberta Empresarial

Quando você verifica os resultados da descoberta empresarial, a ferramenta Developer executa o perfil em todas as linhas da fonte de dados. O valor de porcentagem de conformidade pode variar após a verificação, dependendo dos valores da coluna em todas as linhas da fonte de dados.

1. Abra um perfil depois de executá-lo.
2. Verifique se você está na exibição **Criação de Perfil de Chave Externa**.
3. Clique no ícone **Exibição tabular** na parte superior da exibição.

   A exibição tabular mostra entidades no painel esquerdo.
4. Clique com o botão direito do mouse em um objeto de dados no painel esquerdo e selecione **Exibir Relacionamentos de Colunas**.

   Você pode exibir os relacionamentos de colunas no objeto de dados selecionado com colunas em outros objetos de dados. Role a tela para a direita para exibir detalhes como tipo de relacionamento, porcentagem de conformidade, status de verificação e status de confirmação.
5. Clique com o botão direito do mouse em uma linha e selecione **Verificar**.

   A caixa de diálogo **Executar Perfil** é exibida. Após a verificação ser concluída, selecione a linha para exibir a sobreposição dos relacionamentos de chave primária e chave externa em um diagrama de Venn.

Fazendo a curadoria de relacionamentos de coluna

Você pode aprovar, rejeitar e redefinir os relacionamentos de dados na exibição **Criação de Perfil de Chave Externa**.

1. Abra um perfil depois de executá-lo.
2. Verifique se você está na exibição **Criação de Perfil de Chave Externa**.
3. Na exibição gráfica para rejeitar um relacionamento de coluna inferido, selecione o objeto de dados, clique com o botão direito do mouse em um objeto de dados e selecione uma das seguintes opções:
   - **Rejeitar Todos os Relacionamentos com** > **Chaves Primárias Inferidas**. Escolha essa opção para rejeitar todos os relacionamentos entre as colunas com chaves primárias inferidas no objeto de dados e as colunas com chaves externas inferidas nos outros objetos de dados conectados.
   - **Rejeitar Todos os Relacionamentos com** > **Chaves Externas Inferidas**. Escolha essa opção para rejeitar todos os relacionamentos entre as colunas com chaves externas inferidas no objeto de dados e as colunas com chaves primárias inferidas nos outros objetos de dados conectados.
   - **Rejeitar Todos os Relacionamentos com** > **Chaves Primárias e Externas Inferidas**. Escolha essa opção para rejeitar todos os relacionamentos entre as colunas com chaves primárias inferidas no objeto de dados e as colunas com chaves externas inferidas nos outros objetos de dados conectados, bem como todos os relacionamentos entre as colunas com chaves externas inferidas no objeto de dados e as colunas com chaves primárias inferidas nos outros objetos de dados conectados.
4. Na exibição gráfica, clique com o botão direito do mouse em um objeto de dados e selecione **Exibir Relacionamentos de Coluna**.
5. Selecione um relacionamento de objeto de dados que você fazer curadoria.
6. Para aprovar o relacionamento de coluna, clique com o botão direito do mouse e clique em **Aprovar**.

   O status da linha é alterado para **Aprovado**.
7. Para restaurar o status inferido do relacionamento de coluna inferido, clique duas vezes e clique em **Redefinir**.
8. Para exibir os relacionamentos de coluna rejeitados, clique com o botão direito do mouse em uma das linhas e selecione **Mostrar Rejeitados**.
9. Para ocultar os tipos de dados rejeitados, clique com o botão direito do mouse em uma das linhas e selecione **Ocultar Rejeitados**.

Confirmando os Resultados no Repositório do Modelo

Depois de executar um perfil, você pode salvar os relacionamentos de coluna entre objetos de dados no repositório do Modelo. Você pode confirmar os relacionamentos no repositório do Modelo de perfil a partir da exibição tabular da exibição **Criação de Perfil de Chave Externa**.

1. Abra um perfil depois de executá-lo.
2. Verifique se você está na exibição **Criação de Perfil de Chave Externa**.
3. Clique com o botão direito do mouse em um objeto de dados no painel esquerdo e selecione **Exibir Relacionamentos de Colunas**.

   Você pode exibir os relacionamentos de colunas no objeto de dados selecionado com colunas em outros objetos de dados.
4. Clique com o botão direito do mouse em uma linha e selecione **Aprovar**.

Exibição Domínios de Dados

A exibição **Domínios de Dados** lista os domínios de dados e estatísticas de coluna correspondentes que a ferramenta Developer descobre como parte da descoberta empresarial. Você pode verificar as colunas, fazer uma busca detalhada em linhas e ver resultados do perfil de objeto de dados na exibição **Domínios de Dados**.

Exibindo Resultados da Descoberta do Domínio de Dados

Você pode exibir os resultados da descoberta do domínio de dados na guia **Domínios de Dados**. É possível pesquisar domínios de dados e exibir domínios de dados classificados por grupos de domínios de dados.

1. Execute um perfil para realizar a descoberta empresarial.
2. Abra o perfil.
3. Clique na guia **Domínios de Dados** para exibir os resultados da descoberta do domínio de dados.

   Os resultados do perfil do objeto de dados aparecem no painel direito.
4. Digite uma parte do nome do domínio de dados no campo de pesquisa para localizar um domínio de dados específico.

   Selecione **Mostrar grupo do domínio de dados na hierarquia** para classificar a lista de domínios de dados pelo grupo do domínio de dados.

Verificando Resultados da Descoberta do Domínio de Dados

Quando você executa um perfil, ele analisa uma amostra da fonte de dados para inferir os resultados do perfil. Você pode executar o perfil em todas as linhas dos dados de origem para verificar os resultados de inferência.

1. Abra um perfil depois de executá-lo.
2. Clique na guia **Domínios de Dados** para exibir os resultados.

   É possível exibir os resultados da descoberta do domínio de dados no painel direito.
3. Selecione uma coluna que você queira verificar no painel direito.
4. Clique com o botão direito do mouse na coluna e selecione **Verificar** para executar o perfil em todas as linhas da fonte de dados.

   Você poderá ver uma alteração no valor **% de Conformidade dos Dados** ou **Contagem de linhas em conformidade** depois de verificar os resultados.

Fazendo Buscas Detalhadas em Linhas

Você pode fazer uma busca detalhada nos resultados da descoberta do domínio de dados para fazer uma análise de dados mais aprofundada.

1. Abra um perfil depois de executá-lo.
2. Clique na guia **Domínios de Dados** para exibir os resultados.

   É possível exibir os resultados da descoberta do domínio de dados no painel direito.
3. Selecione uma linha no painel direito em que você deseja fazer uma busca detalhada.
4. Clique com o botão direito do mouse na coluna e selecione **Busca Detalhada** para fazer uma busca detalhada nas linhas de origem.

Exibindo Resultados do Perfil de Objeto de Dados na Exibição Domínios de Dados

Você pode exibir os resultados do perfil de um objeto de dados selecionado na exibição **Domínios de Dados**.

1. Verifique se você está na exibição **Domínios de Dados**.
2. Selecione um domínio de dados no painel **Domínios com Perfil Criado**.
3. No painel **Colunas** à direita, selecione uma coluna.
4. Clique com o botão direito do mouse na coluna e selecione **Abrir Perfil de Objeto de Dados**.

   Os resultados do perfil de objeto de dados são exibidos em uma guia.

Exibição Perfil de Coluna

A exibição **Perfil de Coluna** mostra um resumo dos resultados do perfil de coluna para perfis de objeto de dados único que a Developer tool executa como parte da descoberta empresarial. Você pode exibir as estatísticas de coluna como os valores exclusivos, valores nulos, tipos de dados e valores máximo e mínimo para cada coluna nos objetos de dados.

Exibindo Resultados do Perfil de Objeto de Dados

A descoberta empresarial inclui a execução de um perfil de objeto de dados perfil para descobrir estatísticas de dados da coluna, chaves primárias, chaves candidatas e domínios de dados. Você pode exibir os resultados do perfil de um objeto de dados selecionado na exibição **Perfil de Coluna**.

1. Verifique se você está na exibição **Perfil de Coluna**.
2. Selecione um objeto de dados no painel **Objetos de Dados com Perfil Criado**.
3. No painel **Colunas** à direita, selecione uma coluna.
4. Clique com o botão direito do mouse na coluna e selecione **Exibir Perfil de Objeto de Dados**.

   Os resultados do perfil de objeto de dados são exibidos em uma guia. Os resultados do perfil da coluna são exibidos por padrão.
5. Clique em **Inferência de Chave Primária** para exibir os resultados do perfil de chave primária.
6. Clique em **Inferência de Dependência Funcional** para exibir os resultados da descoberta de dependência funcional.
7. Clique em **Descoberta do Domínio de Dados** para exibir os resultados da descoberta do domínio de dados.

Exibindo Resultados do Perfil de Coluna Durante a Execução da Descoberta Empresarial

O tempo necessário para concluir a descoberta empresarial depende do número de tarefas de perfil, o tamanho da fonte de dados e o tipo de perfil. Quando a ferramenta Developer continua a executar as tarefas

de descoberta de dados, você pode ver os resultados dos perfis de coluna que ela conclui nos estágios iniciais da descoberta de dados.

1. Após executar o perfil, clique em **Perfis** na janela **Propriedades**.
2. Selecione o perfil de coluna cujos resultados você deseja ver. Verifique se o status da execução do perfil é **Êxito** na janela **Propriedades**.
3. Clique em **Abrir** para exibir os resultados em outra guia.
4. Na seção **Resultados**, selecione **Criação de Perfil de Coluna** para exibir os resultados no painel direito.

Exibindo Resultados da Descoberta do Domínio de Dados Durante a Execução da Descoberta Empresarial

Quando a ferramenta Developer continua a executar as tarefas de descoberta de dados incluídas na descoberta empresarial, você pode ver os resultados da descoberta do domínio de dados que ela conclui nos estágios iniciais da descoberta empresarial.

1. Após começar a executar o perfil, clique em **Perfis** na janela **Propriedades**.
2. Selecione o perfil para o qual você deseja exibir os resultados do domínio de dados. Verifique se o status da execução do perfil é **Êxito** no painel **Propriedades**.
3. Clique em **Abrir** para exibir os resultados em outra guia.
4. Na seção **Resultados**, selecione **Criação de Perfil de Descoberta do Domínio de Dados** para exibir os resultados no painel direito.

Exibindo o Status da Descoberta Empresarial em Tempo de Execução

A exibição **Progresso** da ferramenta Developer exibe o progresso das operações como uma execução de perfil. Você pode ver o status em tempo de execução das tarefas de descoberta empresarial na exibição **Progresso**.

1. Após executar um perfil para realizar a descoberta empresarial nas fontes de dados, clique no botão **Exibição do Progresso** no canto inferior direito da ferramenta Developer.

   O painel **Progresso** será exibido se ainda não estiver aberto.
2. Clique no link **Execução da Descoberta Empresarial : Exibir Status das Tarefas** para abrir a caixa de diálogo de subtarefas.

   A caixa de diálogo lista várias tarefas de perfil que fazem parte da descoberta empresarial. Você pode exibir o nome do perfil, o tipo e seu status.
3. Clique no cabeçalho da coluna para classificar as tarefas de perfil. Por exemplo, para classificar as tarefas de perfil por seu status, clique no cabeçalho da coluna **Status**.
4. Se precisar cancelar uma tarefa de perfil específica, selecione a tarefa e clique em **Cancelar**.

   O status da tarefa cancelada muda para **Encerrado**.

Arquivos de Exportação de Descoberta Empresarial

Depois de executar um perfil de descoberta empresarial, você pode exportar as informações, incluindo todos os relacionamentos de objeto de dados, os domínios de dados e os resultados de tarefa de chave externa individual. Você pode salvar a imagem gráfica dos relacionamentos de objetos de dados como um arquivo `.jpg`.

Quando você exporta os resultados de perfil, a ferramenta Developer salva todos os resultados da descoberta empresarial em vários arquivos do Microsoft Excel. Você pode exibir relacionamentos de objeto de dados, resultados do perfil de coluna, resultados da descoberta do domínio de dados, entidades e resultados de tarefa de chave externa individual em arquivos separados.

Exportando Resultados da Descoberta Empresarial

Você pode exportar a lista de entidades e a composição de cada entidade, todos os objetos de dados e relacionamentos de objeto de dados de uma entidade no nível da coluna, domínios de dados e resultados da criação de perfil de coluna.

1. Execute um perfil para realizar a descoberta empresarial.
2. Na exibição **Relacionamentos**, **Domínios de Dados** ou **Perfil de Coluna**, clique no ícone **Exportar** na área superior direita da janela.

   A caixa de diálogo **Exportar dados para um arquivo** é exibida.
3. Digite o nome do arquivo. Opcionalmente, use o nome de arquivo padrão.
4. Em **Salvar**, escolha **Salvar no Cliente** e clique em **Procurar** para selecionar uma localização e salvar o arquivo localmente no seu computador. Por padrão, o Informatica Developer grava o arquivo em uma localização definida nas propriedades do Data Integration Service do Informatica Administrator.
5. Clique em **OK**.

CAPÍTULO 27

Glossário Comercial Desktop no Informatica Developer

Este capítulo inclui os seguintes tópicos:

- Pesquisa no Business Glossary, 217
- Pesquisar um Termo Comercial, 217
- Personalizando Teclas de Acesso para Pesquisar um Termo Comercial, 218

Pesquisa no Business Glossary

Pesquise o significado de um nome de objeto da Developer tool como um termo comercial no Business Glossary Desktop para compreender seu requisito comercial e sua implementação atual.

Um glossário comercial é um conjunto de termos que usam linguagem comercial para definir os conceitos para usuários comerciais. Um termo comercial fornece a definição e o uso comercial de um conceito. O Business Glossary Desktop é um cliente que se conecta ao Serviço do Metadata Manager, que hospeda o glossário comercial. Use o Business Glossary Desktop para pesquisar termos comerciais em um glossário comercial.

Se o Business Glossary Desktop estiver instalado na sua máquina, você poderá selecionar um objeto na Developer tool e usar teclas de acesso ou o menu Pesquisar para pesquisar o nome desse objeto no glossário comercial. É possível pesquisar nomes de objetos nas exibições da Developer tool, como a exibição **Object Explorer**, ou nomes de colunas, perfis e portas de transformação no editor.

Por exemplo, um desenvolvedor deseja encontrar um termo comercial em um glossário comercial que corresponda ao objeto de dados Sales_Audit na Developer tool. O desenvolvedor deseja exibir os detalhes do termo comercial para compreender os requisitos comerciais e a implementação atual do objeto Sales_Audit na Developer tool. Isso pode ajudar o desenvolver a compreender o significado do objeto de dados e quais alterações talvez precisem ser implementadas nesse objeto.

Pesquisar um Termo Comercial

Pesquise um nome de objeto da ferramenta Developer no Business Glossary Desktop como um termo comercial para compreender seu requisito comercial e sua implementação atual.

O Glossário Comercial Desktop deve estar instalado no computador.

1. Selecione um objeto.
2. Escolha entre usar teclas de acesso ou o menu Pesquisar para abrir o Glossário Comercial Desktop.
   - Para usar teclas de acesso, use a seguinte combinação de teclas de acesso:

     `CTRL+Shift+F`
   - Para usar o menu Pesquisar, clique em **Pesquisar** > **Glossário Comercial**.

   O **Glossário Comercial Desktop** é exibido e mostra os termos comerciais que correspondem ao nome de objeto.

Personalizando Teclas de Acesso para Pesquisar um Termo Comercial

Personalize teclas de acesso para alterar a combinação de teclas que abrem o Business Glossary Desktop.

1. No menu da ferramenta Developer, clique em **Janela** > **Preferências** > **Geral** > **Teclas**.
2. Para localizar ou pesquisar **Pesquisar Glossário Comercial** na lista de comandos, selecione uma das seguintes opções:
   - Para procurar as teclas, insira Pesquisar Glossário Comercial na caixa de pesquisa.
   - Para percorrer até as teclas, role para localizar o comando **Pesquisar Glossário Comercial** na coluna **Comando**.
3. Clique no comando **Pesquisar Glossário Comercial**.
4. Clique em **Desassociar Comando**.
5. No campo **Associação**, insira uma combinação de teclas.
6. Clique em **Aplicar** e em **OK**.

ÍNDICE

A

análise de associação
  visão geral 201

C

Configuração Sqoop
  criação de perfil 39, 147
criação de perfil
  arquitetura 19
  gerenciamento de versões e bloqueios 27
  visão geral 17
Criação de Perfil de Mapplet e Mapeamento
  Visão Geral 174
criando um perfil de coluna
  perfis 41
criando uma regra de expressão
  regras 49
curadoria
  conceitos 32
  Informatica Analyst 74
  Informatica Developer 167
  processo 33
  tarefas 33

D

descoberta de chave externa
  visão geral 200
descoberta de chave primária
  visão geral 148
descoberta de dados
  processo 20
  visão geral 22
descoberta de dependência funcional
  visão geral 149
descoberta de sobreposição
  executando 204
  resultados 203
  visão geral 203
descoberta do domínio de dados
  processo 31
  visão geral 29
  visão geral do Informatica Analyst 104
  visão geral do Informatica Developer 177
descoberta empresarial
  editando 198
  editando na ferramenta Analyst 119
  executando no Informatica Analyst 118
  exibição criação de perfil de chave externa 208
  exibição domínios de dados 213
  exibição perfil de coluna 214
  exibição relacionamentos 207
descoberta empresarial *()*
  exibição tabular 211
  exibindo relacionamentos de objeto de dados 208
  processo 191
  status em tempo de execução 215
  visão geral 190
descoberta empresarial de perfil
  executando 199
descoberta empresarial na ferramenta Analyst
  conflito de tipo de dados 124
  exibição perfis 125
  exibição resumida 122
  processo 116
  visão geral 115
domínio de dados
  criando a partir de resultados do perfil no Informatica Analyst 106
  criando a partir de resultados do perfil no Informatica Developer 179
  criando no Informatica Analyst 105
  criando no Informatica Developer 179
  localizar no Informatica Developer 180
  visão geral 30
domínios de dados
  exportando 181
  importando 180
  localizar no Informatica Analyst 106

E

exceção
  detectando 63
exportar
  linhagem de scorecard para XML 172

F

filtros
  visão geral 50

G

glossário do domínio de dados
  Informatica Analyst 104
  Informatica Developer 178
  visão geral 31
gráficos de tendência
  custo 91
  exibindo 92
  exportar do Informatica Analyst 93
  pontuação 91
grupo do domínio de dados
  criando no Informatica Analyst 105
  criando no Informatica Developer 178
  visão geral 30

I

Informatica Analyst
gerenciamento de versões e bloqueios 41
regras 46
resultados de perfil de coluna 55
resultados do perfil de coluna 67
visão geral de perfis da coluna 36, 69
Informatica Developer
exibições de perfil 141
regras 169
visão geral do perfil 139

L

linhagem de scorecard
exibir no Informatica Analyst 103
exibir no Informatica Developer 172

O

objeto de dados de arquivo simples
sincronizando 44
objeto de dados de tabela
sincronizando 45
objeto de mapeamento
executar um perfil 174
objetos de dados com scorecards
Informatica Analyst 96
opções de configuração
descoberta empresarial na ferramenta Analyst 116
opções de descoberta do domínio de dados
Informatica Developer 182
opções de perfil
descoberta empresarial 192

P

painel de métricas cumulativas
Informatica Analyst 97
painel de scorecards
Informatica Analyst 94
painel de scorecards por projeto
Informatica Analyst 95
painel de tendência de execução do scorecard
Informatica Analyst 96
perfil
componentes 23
formatos Avro ou Parquets 158
formatos XML e JSON 157, 158
perfil da coluna
busca detalhada 73
perfil de chave externa
descobrindo 200
perfil de coluna
Informatica Developer 144
opções 27
perfil do sistema operacional 40, 151
processo 37
visão geral 26
perfil de descoberta empresarial
criando scripts DDL 205
scripts DDL 205
perfil de objeto de dados
comentários 154
perfis
criando um filtro 50
criando um perfil de coluna 41
editando um filtro 53
editar um perfil de coluna 42
executando 43, 68, 69, 112
Perfis de Mapplet e Mapeamento
Visão Geral 174
perfis de objeto de dados
criando vários perfis 153
criar um perfil único 152
descoberta empresarial 197
visão geral 143
pesquisa de descoberta
pré-requisitos 127
pesquisa de descoberta na ferramenta Analyst
processo 127
pesquisar
glossário comercial 217
projetos
termo comercial do Metadata Manager 136

R

regras
aplicando no PowerCenter Express 170
aplicar no Informatica Developer 170
aplicar uma regra predefinida 47
criando uma regra de expressão 49
criar no Informatica Developer 170
expressão 48
pré-requisitos 169
predefinido 47
regras de expressão
processo 48
regras predefinidas
processo 47
resultados da descoberta do domínio de dados
exportando do Informatica Analyst 114
exportando do Informatica Developer 189
exportando no Informatica Analyst 113
Resultados da descoberta do domínio de dados
Informatica Analyst 112
Informatica Developer 186
resultados da descoberta empresarial
exportando 216
salvando como uma imagem 210
visão geral 206
resultados da descoberta empresarial no Informatica Analyst
visão geral 121
resultados da pesquisa de descoberta
interface 130
resultados da pesquisa de descoberta na ferramenta Analyst
visão geral 129
resultados de perfil
adicionando comentários no Informatica Developer 154
termos comerciais 79
tipos de dados de coluna 166
resultados de perfil de coluna na ferramenta Analyst
detalhes da coluna 58, 72
interface 70
resumo 56
resultados de scorecard
exportando 99
exportar do Informatica Analyst 99
exportar para o Excel 99

resultados do perfil
aprovando domínios de dados 113
aprovando domínios de dados no Informatica Developer 188
aprovando tipos de dados 75
aprovando tipos de dados no Informatica Developer 167
comentários 79
Excel 76
exibição detalhada 59
exibição resumida 57
exportando 75
exportando do Informatica Analyst 76
exportar no Informatica Developer 168
fazendo curadoria dos relacionamentos de coluna no Informatica Developer 212
fazer busca detalhada 74
marcas 80
padrões de coluna 63
rejeitando domínios de dados 113
rejeitando domínios de dados no Informatica Developer 189
rejeitando tipos de dados 75
rejeitando tipos de dados na Developer tool 167
resumo 70, 73
tipos de dados de coluna 61
valores de coluna 64
resultados do perfil de coluna
Informatica Developer 163
perfil de coluna 163
resultados do perfil de coluna na ferramenta Analyst
interface 57, 71
resultados do perfil de descoberta do domínio de dados
Microsoft Excel 114

S

scorecard
configurar notificações 101
scorecard *()*
definir as configurações globais de notificação 102
scorecards
adicionando colunas a um scorecard 84
criar um grupo de métrica 88
custo de dados inválidos 87
custo fixo 87
custo variável 87
definir limites 88
editando 86
editar um grupo de métrica 89
excluir um grupo de métrica 90
executando 85
exibindo 86
fazer busca detalhada 90
gráfico de tendências 90
grupos de métrica 88
Informatica Analyst 81
Informatica Developer 171
métricas 87
mover pontuações 89
notificações 100
pesos de métrica 87
processo do Informatica Analyst 82
visão geral 28

T

termo comercial
pesquisar um termo comercial 137
termo comercial do Metadata Manager
gerenciar termo comercial 137
projetos 136
termos comerciais
personalizando teclas de acesso 218
pesquisando 217