Informatica®
10.1

Profil-Handbuch

Informatica Profil-Handbuch
10.1
Juni 2016

© Copyright Informatica LLC 2014, 2018

Diese Software und die zugehörige Dokumentation enthalten proprietäre Informationen der Informatica LLC, werden unter einem Lizenzvertrag mit Einschränkungen hinsichtlich Verwendung und Veröffentlichung zur Verfügung gestellt und sind urheberrechtlich geschützt. Das Zurückentwickeln (Reverse Engineering) der Software ist untersagt. Ohne ausdrückliche schriftliche Genehmigung der Informatica LLC darf kein Teil dieses Dokuments zu irgendeinem Zweck vervielfältigt oder übertragen werden, unabhängig davon, auf welche Art und Weise oder mit welchen Mitteln (elektronisch, mechanisch, durch Fotokopieren, Aufzeichnen usw.) dies geschieht. Diese Software ist möglicherweise durch US-amerikanische und/oder internationale Patente und weitere angemeldete Patente geschützt.

Die Verwendung, Vervielfältigung oder Veröffentlichung der Software durch die US-Regierung unterliegt den Bestimmungen des jeweiligen Softwarelizenzvertrags sowie ggf. den Bestimmungen in DFARS 227.7202-1(a) und 227.7702-3(a) (1995), DFARS 252.227-7013 © (1)(ii) (OCT. 1988), FAR 12.212(a) (1995), FAR 52.227-19 oder FAR 52.227-14 (ALT III).

Die in diesem Produkt und in dieser Dokumentation enthaltenen Informationen können jederzeit ohne vorherige Ankündigung geändert werden. Sollten Sie mit diesem Produkt oder dieser Dokumentation Probleme haben, teilen Sie uns dies bitte schriftlich mit.

Informatica, Informatica Platform, Informatica Data Services, PowerCenter, PowerCenterRT, PowerCenter Connect, PowerCenter Data Analyzer, PowerExchange, PowerMart, Metadata Manager, Informatica Data Quality, Informatica Data Explorer, Informatica B2B Data Transformation, Informatica B2B Data Exchange Informatica On Demand, Informatica Identity Resolution, Informatica Application Information Lifecycle Management, Informatica Complex Event Processing, Ultra Messaging, Informatica Master Data Management und Live Data Map sind Marken oder eingetragene Marken der Informatica LLC in den USA und anderen Ländern. Alle weiteren Produkt- und Firmennamen sind möglicherweise Markennamen oder Warenzeichen der jeweiligen Eigentümer.

Teile dieser Software und/oder Dokumentation sind durch die Urheberrechte Dritter geschützt, einschließlich und ohne Einschränkung: Copyright DataDirect Technologies. Alle Rechte vorbehalten. Copyright © Sun Microsystems. Alle Rechte vorbehalten. Copyright © RSA Security Inc. Alle Rechte vorbehalten. Copyright © Ordinal Technology Corp. Alle Rechte vorbehalten. Copyright © Aandacht c.v. Alle Rechte vorbehalten. Copyright Genivia, Inc. Alle Rechte vorbehalten. Copyright Isomorphic Software. Alle Rechte vorbehalten. Copyright © Meta Integration Technology, Inc. Alle Rechte vorbehalten. Copyright © Intalio. Alle Rechte vorbehalten. Copyright © Oracle. Alle Rechte vorbehalten. Copyright © Adobe Systems Incorporated. Alle Rechte vorbehalten. Copyright © DataArt, Inc. Alle Rechte vorbehalten. Copyright © ComponentSource. Alle Rechte vorbehalten. Copyright © Microsoft Corporation. Alle Rechte vorbehalten. Copyright © Rouge Wave Software, Inc. Alle Rechte vorbehalten. Copyright © Teradata Corporation. Alle Rechte vorbehalten. Copyright © Yahoo! Inc. Alle Rechte vorbehalten. Copyright © Glyph & Cog, LLC. Alle Rechte vorbehalten. Copyright © Thinkmap, Inc. Alle Rechte vorbehalten. Copyright © Clearpace Software Limited. Alle Rechte vorbehalten. Copyright © Information Builders, Inc. Alle Rechte vorbehalten. Copyright © OSS Nokalva, Inc. Alle Rechte vorbehalten. Copyright Edifecs, Inc. Alle Rechte vorbehalten. Copyright Cleo Communications, Inc. Alle Rechte vorbehalten. Copyright © International Organization for Standardization 1986. Alle Rechte vorbehalten. Copyright © ej-technologies GmbH . Alle Rechte vorbehalten. Copyright © Jaspersoft Corporation. Alle Rechte vorbehalten. Copyright © International Business Machines Corporation. Alle Rechte vorbehalten. Copyright © yWorks GmbH. Alle Rechte vorbehalten. Copyright © Lucent Technologies. Alle Rechte vorbehalten. Copyright © Universität von Toronto. Alle Rechte vorbehalten. Copyright © Daniel Veillard. Alle Rechte vorbehalten. Copyright © Unicode, Inc. Copyright IBM Corp. Alle Rechte vorbehalten. Copyright © MicroQuill Software Publishing, Inc. Alle Rechte vorbehalten. Copyright © PassMark Software Pty Ltd. Alle Rechte vorbehalten. Copyright © LogiXML, Inc. Alle Rechte vorbehalten. Copyright © 2003-2010 Lorenzi Davide. Alle Rechte vorbehalten. Copyright © Red Hat, Inc. Alle Rechte vorbehalten. Copyright © The Board of Trustees of the Leland Stanford Junior University. Alle Rechte vorbehalten. Copyright © EMC Corporation. Alle Rechte vorbehalten. Copyright © Flexera Software. Alle Rechte vorbehalten. Copyright © Jinfonet Software. Alle Rechte vorbehalten. Copyright © Apple Inc. Alle Rechte vorbehalten. Copyright © Telerik Inc. Alle Rechte vorbehalten. Copyright © BEA Systems. Alle Rechte vorbehalten. Copyright © PDFlib GmbH. Alle Rechte vorbehalten. Copyright © Orientation in Objects GmbH. Alle Rechte vorbehalten. Copyright © Tanuki Software, Ltd. Alle Rechte vorbehalten. Copyright © Ricebridge. Alle Rechte vorbehalten. Copyright © Sencha, Inc. Alle Rechte vorbehalten. Copyright © Scalable Systems, Inc. Alle Rechte vorbehalten. Copyright © jQWidgets. Alle Rechte vorbehalten. Copyright © Tableau Software, Inc. Alle Rechte vorbehalten. Copyright© MaxMind, Inc. Alle Rechte vorbehalten. Copyright © TMate Software s.r.o. Alle Rechte vorbehalten. Copyright © MapR Technologies Inc. Alle Rechte vorbehalten. Copyright © Amazon Corporate LLC. Alle Rechte vorbehalten. Copyright © Highsoft. Alle Rechte vorbehalten. Copyright © Python Software Foundation. Alle Rechte vorbehalten. Copyright © BeOpen.com. Alle Rechte vorbehalten. Copyright © CNRI. Alle Rechte vorbehalten.

Dieses Produkt enthält Software, die von der Apache Software Foundation (http://www.apache.org/) entwickelt wurde, und andere Software, die unter den Bedingungen des Apache-Lizenzvertrags lizenziert ist („Lizenz“). Eine Kopie dieser Lizenzen finden Sie unter http://www.apache.org/licenses/. Sofern nicht gesetzlich vorgeschrieben oder schriftlich vereinbart, erfolgt der Vertrieb der Software unter der Lizenz auf der BASIS „WIE BESEHEN” OHNE GARANTIEN ODER KONDITIONEN IRGENDEINER ART, weder ausdrücklich noch impliziert. Berechtigungen und Einschränkungen für bestimmte Sprachen finden Sie in der Lizenz.

Dieses Produkt enthält Software, die von Mozilla (http://www.mozilla.org/) entwickelt wurde, Software Copyright The JBoss Group, LLC. Alle Rechte vorbehalten; Software Copyright © 1999-2006 by Bruno Lowagie und Paulo Soares, und andere Software, die gemäß den verschiedenen Versionen des GNU Lesser General Public License Agreement unter http://www.gnu.org/licenses/lgpl.html lizenziert ist. Die Materialien werden „wie besehen“ kostenlos von Informatica bereitgestellt, ohne ausdrückliche oder stillschweigende Gewährleistung, einschließlich, jedoch nicht beschränkt auf die stillschweigenden Gewährleistungen der Handelsüblichkeit und der Eignung für einen bestimmten Zweck.

Das Produkt enthält ACE(TM) und TAO(TM) Software, Copyright Douglas C. Schmidt und seine Forschungsgruppe an der Washington University, University of California, Irvine und Vanderbilt University, Copyright (©) 1993-2006. Alle Rechte vorbehalten.

Dieses Produkt enthält Software, die von OpenSSL Project zur Verwendung im OpenSSL Toolkit entwickelt wurde (Copyright The OpenSSL Project. Alle Rechte vorbehalten). Die erneute Verteilung dieser Software unterliegt den unter „http://www.openssl.org“ und „http://www.openssl.org/source/license.html“ verfügbaren Bedingungen.

Dieses Produkt enthält urheberrechtlich geschützte Curl-Software (Copyright 1996-2013, Daniel Stenberg, <daniel@haxx.se>). Alle Rechte vorbehalten. Die mit dieser Software verbundenen Berechtigungen und Einschränkungen unterliegen den unter „http://curl.haxx.se/docs/copyright.html“ verfügbaren Bedingungen. Die Erlaubnis, diese Software für jeden beliebigen Zweck gegen Gebühr oder kostenlos zu verwenden, zu kopieren, zu ändern und zu verteilen, wird hiermit erteilt, sofern die oben genannten urheberrechtlichen Hinweise und diese Erlaubnis in allen Exemplaren angegeben werden.

Das Produkt enthält urheberrechtlich geschützte Software, Copyright 2001-2005 (©) MetaStuff, Ltd. Alle Rechte vorbehalten. Die mit dieser Software verbundenen Berechtigungen und Einschränkungen unterliegen den unter „http://www.dom4j.org/ license.html“ verfügbaren Bedingungen.

Das Produkt enthält urheberrechtlich geschützte Software, Copyright © 2004-2007, The Dojo Foundation. Alle Rechte vorbehalten. Die mit dieser Software verbundenen Berechtigungen und Einschränkungen unterliegen den unter „http://dojotoolkit.org/license“ verfügbaren Bedingungen.

Dieses Produkt enthält urheberrechtlich geschützte ICU-Software, Copyright International Business Machines Corporation und andere. Alle Rechte vorbehalten. Die mit dieser Software verbundenen Berechtigungen und Einschränkungen unterliegen den unter „http://source.icu-project.org/repos/icu/icu/trunk/license.html“ verfügbaren Bedingungen.

Dieses Produkt enthält urheberrechtlich geschützte Software, Copyright © 1996-2006 Per Bothner. Alle Rechte vorbehalten. Das Ihnen erteilte Recht, diese Materialien zu verwenden, unterliegt den unter „http:// www.gnu.org/software/ kawa/Software-License.html“ verfügbaren Bedingungen.

Dieses Produkt enthält urheberrechtlich geschützte OSSP UUID-Software (Copyright © 2002 Ralf S. Engelschall, Copyright © 2002 The OSSP Project Copyright © 2002 Cable & Wireless Deutschland). Die mit dieser Software verbundenen Berechtigungen und Einschränkungen unterliegen den unter „http://www.opensource.org/licenses/mit-license.php“ verfügbaren Bedingungen.

Dieses Produkt enthält Software, die von Boost (http://www.boost.org/) oder unter der Softwarelizenz von Boost entwickelt wurde. Die mit dieser Software verbundenen Berechtigungen und Einschränkungen unterliegen den unter „http://www.boost.org/LICENSE_1_0.txt“ verfügbaren Bedingungen.

Dieses Produkt enthält urheberrechtlich geschützte Software, Copyright © 1997-2007 University of Cambridge. Die mit dieser Software verbundenen Berechtigungen und Einschränkungen unterliegen den unter http://www.pcre.org/license.txt einsehbaren Bedingungen.

Dieses Produkt enthält urheberrechtlich geschützte Software, Copyright © 2007 The Eclipse Foundation. Alle Rechte vorbehalten. Die mit dieser Software verbundenen Berechtigungen und Einschränkungen unterliegen den unter „http://www.eclipse.org/org/documents/epl-v10.php“ und „http://www.eclipse.org/org/documents/edl-v10.php“ verfügbaren Bedingungen.

Dieses Produkt enthält Software gemäß den Lizenzbedingungen unter 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.

Dieses Produkt enthält Software, die unter der Academic Free License (http://www.opensource.org/licenses/afl-3.0.php), der Common Development Distribution License (http://www.opensource.org/licenses/cddl1.php), der Common Public License (http://www.opensource.org/licenses/cpl1.0.php), den Sun Binary Code License Agreement Supplemental License Terms, der BSD License (http:// www.opensource.org/licenses/bsd-license.php), der neuen BSD License (http://opensource.org/licenses/BSD-3-Clause), der MIT License (http://www.opensource.org/licenses/mit-license.php), der Artistic License (http://www.opensource.org/licenses/artistic-license-1.0) und der Initial Developer's Public License Version 1.0 (http://www.firebirdsql.org/en/initial-developer-s-public-license-version-1-0/) lizenziert ist.

Dieses Produkt enthält urheberrechtlich geschützte Software, Copyright © 2003-2006 Joe Walnes, 2006-2007 XStream Committers. Alle Rechte vorbehalten. Die mit dieser Software verbundenen Berechtigungen und Einschränkungen unterliegen den unter „http://xstream.codehaus.org/license.html“ verfügbaren Bedingungen. Dieses Produkt enthält Software, die von der Indiana University Extreme! Lab. entwickelt wurde. Weitere Informationen finden Sie unter http://www.extreme.indiana.edu/.

Dieses Produkt enthält Software, Copyright © 2013 Frank Balluffi und Markus Moeller. Alle Rechte vorbehalten. Die mit dieser Software verbundenen Berechtigungen und Einschränkungen unterliegen den Bedingungen der MIT-Lizenz.

Weitere Informationen über die Patente finden Sie unter https://www.informatica.com/legal/patents.html.

HAFTUNGSAUSSCHLUSS: Informatica LLC stellt diese Dokumentation „wie besehen“ bereit, ohne ausdrückliche oder stillschweigende Gewährleistung, einschließlich, jedoch nicht beschränkt auf die Gewährleistungen der Nichtverletzung der Rechte von Dritten, der Handelsüblichkeit oder Eignung für einen bestimmten Zweck. Informatica LLC garantiert nicht die Fehlerfreiheit dieser Software oder Dokumentation. Die in dieser Software oder Dokumentation bereitgestellten Informationen können technische Ungenauigkeiten oder Druckfehler enthalten. Die in dieser Software und in dieser Dokumentation enthaltenen Informationen können jederzeit ohne vorherige Ankündigung geändert werden.

HINWEISE

Dieses Informatica-Produkt (die „Software“) umfasst bestimmte Treiber (die „DataDirect-Treiber“) von DataDirect Technologies, einem Betreiber von Progress Software Corporation („DataDirect“), die folgenden Bedingungen und Bestimmungen unterliegen:

1. DIE DATADIRECT-TREIBER WERDEN „WIE GESEHEN“ OHNE JEGLICHE GEWÄHRLEISTUNG, WEDER AUSDRÜCKLICH NOCH STILLSCHWEIGEND, BEREITGESTELLT, EINSCHLIESSLICH, JEDOCH NICHT BESCHRÄNKT AUF DIE STILLSCHWEIGENDEN GEWÄHRLEISTUNGEN DER HANDELSÜBLICHKEIT, EIGNUNG FÜR EINEN BESTIMMTEN ZWECK UND DER NICHTVERLETZUNG VON RECHTEN DRITTER.
2. IN KEINEM FALL SIND DATADIRECT ODER DRITTANBIETER DEM ENDBENUTZER GEGENÜBER HAFTBAR FÜR UNMITTELBARE, MITTELBARE, KONKRETE, NEBEN-, FOLGE- ODER ANDERE SCHÄDEN, DIE SICH AUS DER VERWENDUNG DER ODBC-TREIBER ERGEBEN, UNABHÄNGIG DAVON, OB SIE IM VORAUS ÜBER DIE MÖGLICHKEIT SOLCHER SCHÄDEN INFORMIERT WORDEN SIND ODER NICHT. DIESE BESCHRÄNKUNGEN GELTEN FÜR ALLE KLAGEGEGENSTÄNDE, EINSCHLIESSLICH, JEDOCH NICHT BESCHRÄNKT AUF VERTRAGSBRUCH, GEWÄHRLEISTUNGSBRUCH, FAHRLÄSSIGKEIT, KAUSALHAFTUNG, TÄUSCHUNG UND ANDERE UNERLAUBTE HANDLUNGEN.

Publikationsdatum: 2018-06-08

Inhalt

Einleitung 9

Informatica-Ressourcen. 9

Informatica-Netzwerk. 9

Informatica-Wissensdatenbank. 10

Informatica-Dokumentation. 10

Informatica-Produktverfügbarkeitsmatrizen. 10

Informatica Velocity. 10

Informatica Marketplace. 10

Globaler Kundensupport von Informatica. 10

Teil I: Einführung in Profile. 12

Kapitel 1: Einführung in Profile. 13

Einführung in Profile - Übersicht. 13

Profiling-Prozess. 14

Profiling-Tools. 14

Profil-Komponenten. 14

Kapitel 2: Spaltenprofilkonzepte. 16

Spaltenprofilkonzepte - Übersicht Profile - Übersicht. 16

Spaltenprofiloptionen. 17

Repository-Profilsperren und Verwaltung versionierter Profile. 17

Scorecards. 18

Kapitel 3: Wiederherstellungskonzepte. 19

Wiederherstellungskonzepte - Übersicht. 19

Wiederherstellungsaufgaben. 20

Teil II: Profiling mit Informatica Analyst. 21

Kapitel 4: Spaltenprofile in Informatica Analyst. 22

Spaltenprofile in Informatica Analyst - Übersicht. 22

Spalten-Profilingprozess. 23

Profil-Optionen. 23

Stichprobenoptionen. 24

Drilldown-Optionen. 24

Laufzeitumgebungsoptionen. 24

Betriebssystemprofile in Informatica Analyst - Übersicht. 25

Auswählen eines Betriebssystemprofils. 25

Repository-Objektsperren und teambasierte Entwicklung - Übersicht. 25

Erstellen eines Spaltenprofils. 26
Bearbeiten eines Spaltenprofils. 28
Ausführen eines Profils. 28
Synchronisieren eines Einfachdatei-Datenobjekts. 29
Synchronisieren eines relationalen Datenobjekts. 30

Kapitel 5: Regeln in Informatica Analyst. 31

Regeln in Informatica Analyst - Übersicht. 31
Regeln in einem Spaltenprofil. 31
Vordefinierte Regeln. 32
Vordefinierte Regeln-Prozess. 32
Anwenden einer vordefinierten Regel. 32
Expressionsregeln. 33
Expressionsregeln-Prozess. 33
Erstellen einer Expressionsregel. 34

Kapitel 6: Spaltenprofilergebnisse in Informatica Analyst. 36

Spaltenprofilergebnisse in Informatica Analyst - Übersicht. 36
Zusammenfassungsansicht. 37
Eigenschaften der Zusammenfassungsansicht. 38
Standardfilter in der Zusammenfassungsansicht. 39
Detailansicht. 39
Bereiche der Detailansicht. 40
Statistiken. 41
Datenvorschau. 42
Datentypen. 42
Sonderfälle. 43
Muster. 44
Werte. 45
Typen von Profilausführungen. 48
Aktuelle Profilausführung. 48
Historische Profilausführung. 48
Zusammengefasste Profilausführung. 49
Auswählen einer Profilausführung. 49
Vergleichen mehrerer Profilergebnisse - Übersicht. 50
Vergleichen mehrerer Profilergebnisse. 50
Zusammenfassungsansicht für den Profilergebnisvergleich 51
Detailansicht des Profilergebnisvergleichs 54
Spaltenprofil-Drilldown. 55
Drilldown an Zeilendaten. 55
Filter auf Drilldown-Daten anwenden. 55
Wiederherstellung im Analyst-Tool. 56
Genehmigen von Datentypen und Datendomänen. 56

Ablehnen von Datentypen und Datendomänen. . . . 56
Exportdateien des Spaltenprofils in Informatica Analyst. . . . 57
Profilieren von Exportergebnissen in eine CSV-Datei. . . . 57
Profilieren von Exportergebnissen in Microsoft Excel. . . . 57
Exportieren von Profilergebnissen aus Informatica Analyst. . . . 58

Kapitel 7: Scorecards in Informatica Analyst. . . . 59

Scorecards in Informatica Analyst - Übersicht. . . . 59
Informatica Analyst Scorecard-Prozess. . . . 60
Erstellen einer Scorecard in Informatica Analyst. . . . 61
Hinzufügen von Spalten zu einer vorhandenen Scorecard. . . . 62
Ausführen einer Scorecard. . . . 63
Anzeigen einer Scorecard. . . . 64
Bearbeiten einer Scorecard. . . . 64
Maße. . . . 65
Metrikgewichtungen. . . . 65
Wert der Datenqualität. . . . 65
Definieren von Schwellenwerten. . . . 66
Metrikgruppen. . . . 67
Erstellen einer Metrikgruppe. . . . 67
Verschieben von Scores in eine Metrikgruppe. . . . 67
Bearbeiten einer Metrikgruppe. . . . 68
Löschen einer Metrikgruppe. . . . 68
Spalten-Drilldown. . . . 69
Trenddiagramme. . . . 69
Score-basiertes Trenddiagramm. . . . 70
Kostenbasiertes Trenddiagramm. . . . 70
Anzeigen von Trenddiagrammen. . . . 71
Exportieren von Trenddiagrammen. . . . 72
Scorecard-Exportdateien in Informatica Analyst. . . . 73
Scorecard-Exportergebnisse in Microsoft Excel. . . . 73
Exportieren von Scorecard-Ergebnissen aus Informatica Analyst. . . . 73
Scorecard-Benachrichtigungen. . . . 74
Vorlage für Benachrichtigungs-E-Mail. . . . 74
Einrichten der Scorecard-Benachrichtigungen. . . . 75
Konfigurieren von globalen Einstellungen für Scorecard-Benachrichtigungen. . . . 76
Scorecard-Herkunft. . . . 76
Anzeigen der Scorecard-Herkunft in Informatica Analyst. . . . 77

Teil III: Profiling mit Informatica Developer. . . . 78

Kapitel 8: Datenobjektprofile. . . . 79

Spaltenprofile in Informatica Developer. . . . 79

Filteroptionen. 79
Stichprobenoptionen. 80
Erstellen eines einzelnen Datenobjektprofils Erstellen eines Profils. 81
Erstellen mehrerer Datenobjektprofile. 82
Synchronisieren eines Einfachdatei-Datenobjekts. 82
Synchronisieren eines nicht-relationalen Datenobjekts. 83

Kapitel 9: Spaltenprofile in halbstrukturierten Datenquellen. 84

Spaltenprofile in halbstrukturierten Datenquellen - Übersicht. 84
Spaltenprofil in einer JSON- oder XML-Datenquelle. 85
Erstellen eines Datenobjekts in einer JSON- oder XML-Datenquelle. 85
Spaltenprofil in einer halbstrukturierten Datenquelle in HDFS. 86
Spaltenprofil in einer JSON- oder XML-Datenquelle in HDFS. 86
Spaltenprofil in einer Avro- oder Parquet-Datenquelle in HDFS. 86
Erstellen einer HDFS-Verbindung. 87
Erstellen eines komplexen Dateidatenobjekts in einer JSON- oder XML-Datei in HDFS. 87
Erstellen eines Datenobjekts in einer Avro- oder Parquet-Datenquelle. 88
Ausführen eines Spaltenprofils in halbstrukturierten Datenquellen. 89

Kapitel 10: Regeln in Informatica Developer. 91

Regeln in Informatica Developer - Übersicht Richtlinien für Regeln. 91
Erstellen einer Regel in Informatica Developer. 92
Anwenden einer Regel in Informatica Developer Eine Regel anwenden. 92

Kapitel 11: Spaltenprofilergebnisse in Informatica Developer. 93

Spaltenprofilergebnisse in Informatica Developer Spaltenprofilergebnisse. 93
Spaltenwert-Eigenschaften. 94
Spaltenmuster-Eigenschaften. 95
Spaltenstatistik-Eigenschaften. 95
Eigenschaften des Spaltendatentyps. 96
Wiederherstellung im Informatica Developer Wiederherstellung in Informatica Developer. 97
Genehmigen von DatentypenGenehmigen von Datentypen in Informatica Developer. 97
Ablehnen von DatentypenAblehnen von Datentypen in Informatica Developer. 97
Exportieren von Profilergebnissen aus Informatica Developer. 98

Kapitel 12: Scorecards in Informatica Developer. 99

Scorecards in Informatica Developer - Übersicht. 99
Erstellen einer Scorecard. 99
Exportieren einer Ressourcendatei zur Scorecard-Herkunftskontrolle. 100
Anzeigen der Scorecard-Herkunft aus Informatica Developer. 100

Kapitel 13: Mapplet- und Mapping-Profiling. 102

Mapplet- und Mapping-Profiling - Übersicht Mapplet- und Mapping-Profile. 102

Ausführen eines Profils für ein Mapplet- oder Mapping-Objekt. 102
Vergleichen von Profilen für Mapping- oder Mapplet-Objekte. 103
Generieren eines Mapping aus einem Profil. 103

Index. 105

Einleitung

Das *Profil-Handbuch* von Informatica richtet sich an Benutzer von Informatica Analyst und Informatica Developer. Es enthält Informationen über das Arbeiten mit Profilen zum Analysieren des Inhalts und der Struktur von Daten.

Verwenden Sie Spaltenprofile, um die Charakteristika von Spalten, z. B. Werthäufigkeit, Prozentsätze und Muster, in einer Datenquelle zu bestimmen.

Informatica-Ressourcen

Informatica-Netzwerk

Im Informatica-Netzwerk finden Sie den globalen Kundensupport von Informatica, die Informatica-Wissensdatenbank und andere Produktressourcen. Für den Zugriff auf das Informatica-Netzwerk besuchen Sie https://network.informatica.com.

Als Mitglied können Sie:

- zentral auf alle Ihre Informatica-Ressourcen zugreifen.
- Durchsuchen Sie die Wissensdatenbank nach Produktressourcen, einschließlich Dokumentation, häufig gestellter Fragen und bewährter Methoden.
- Zeigen Sie Informationen zur Produktverfügbarkeit an.
- Ihre Support-Fälle prüfen.
- Ihr lokales Informatica-Netzwerk für Benutzergruppen suchen und mit anderen Benutzern zusammenarbeiten.

Als Mitglied können Sie:

- zentral auf alle Ihre Informatica-Ressourcen zugreifen.
- Durchsuchen Sie die Wissensdatenbank nach Produktressourcen, einschließlich Dokumentation, häufig gestellter Fragen und bewährter Methoden.
- Zeigen Sie Informationen zur Produktverfügbarkeit an.
- Ihr lokales Informatica-Netzwerk für Benutzergruppen suchen und mit anderen Benutzern zusammenarbeiten.

Informatica-Wissensdatenbank

Verwenden Sie die Informatica-Wissensdatenbank, um das Informatica-Netzwerk nach Produktressourcen, wie z. B. Dokumentation, Ratgeberartikeln, bewährten Methoden und PAMs, zu durchsuchen.

Für den Zugriff auf die Wissensdatenbank besuchen Sie https://kb.informatica.com. Wenn Sie Fragen, Kommentare oder Ideen zur Wissensdatenbank haben, wenden Sie sich per E-Mail an das Team der Informatica-Wissensdatenbank unter KB_Feedback@informatica.com.

Informatica-Dokumentation

Navigieren Sie zur Informatica-Wissensdatenbank unter https://kb.informatica.com/_layouts/ProductDocumentation/Page/ProductDocumentSearch.aspx, um die aktuelle Dokumentation für Ihr Produkt abzurufen.

Wenn Sie Fragen, Kommentare oder Ideen zu dieser Dokumentation haben, wenden Sie sich per E-Mail an das Informatica-Dokumentationsteam unter infa_documentation@informatica.com.

Informatica-Produktverfügbarkeitsmatrizen

Produktverfügbarkeitsmatrizen (PAMs) geben die Versionen der Betriebssysteme, Datenbanken und anderen Typen von Datenquellen und Zielen an, die in einer Produktversion unterstützt werden. Als Mitglied des Informatica-Netzwerks können Sie unter https://network.informatica.com/community/informatica-network/product-availability-matrices auf PAMs zugreifen.

Informatica Velocity

Bei Informatica Velocity handelt es sich um eine Sammlung von Tipps und bewährten Methoden, die von den professionellen Informatica-Diensten entwickelt wurden. Informatica Velocity basiert auf der Praxiserfahrung aus Hunderten von Datenmanagementprojekten und umfasst das kollektive Wissen unserer Berater, die mit Unternehmen aus der ganzen Welt an der Planung, Entwicklung, Bereitstellung und Wartung erfolgreicher Datenmanagementlösungen gearbeitet haben.

Als Mitglied des Informatica-Netzwerks können Sie unter http://velocity.informatica.com auf Informatica Velocity-Ressourcen zugreifen.

Wenn Sie Fragen, Anregungen oder Ideen zu Informatica Velocity haben, wenden Sie sich an die professionellen Informatica-Dienste unter ips@informatica.com.

Informatica Marketplace

Informatica Marketplace ist ein Forum, das Lösungen zur Erweiterung und Verbesserung Ihrer Informatica-Implementierungen bereitstellt. Indem Sie die zahlreichen Lösungen von Informatica-Entwicklern und -Partnern nutzen, können Sie Ihre Produktivität steigern und die Implementierungsdauer Ihrer Projekte verkürzen. Zugriff auf den Informatica Marketplace erhalten Sie unter https://marketplace.informatica.com.

Globaler Kundensupport von Informatica

Sie können sich telefonisch oder über den Online-Support mit einem globalen Support-Center im Informatica-Netzwerk in Verbindung setzen.

Die Telefonnummer des globalen Kundensupports von Informatica vor Ort finden Sie auf der Informatica-Website unter folgender Verknüpfung:
http://www.informatica.com/us/services-and-training/support-services/global-support-centers.

Als Mitglied des Informatica-Netzwerks können Sie den Online-Support unter http://network.informatica.com verwenden.

Teil I: Einführung in Profile

Dieser Teil enthält die folgenden Kapitel:

- Einführung in Profile, 13
- Spaltenprofilkonzepte, 16
- Wiederherstellungskonzepte, 19

KAPITEL 1

Einführung in Profile

Dieses Kapitel umfasst die folgenden Themen:

- Einführung in Profile - Übersicht, 13
- Profiling-Prozess, 14
- Profiling-Tools, 14
- Profil-Komponenten, 14

Einführung in Profile - Übersicht

Erstellen und Ausführen eines Profils, um den Inhalt, die Qualität und die Struktur der Datenquellen einer Anwendung, eines Schemas oder eines Unternehmens zu finden. Der Inhalt der Datenquelle schließt Werthäufigkeiten und Datentypen ein. Die Struktur der Datenquelle schließt Schlüssel und funktionale Abhängigkeiten ein.

Als Teil des Erkennungsprozesses können Sie Profile erstellen und ausführen. Ein Profil ist ein Repository-Objekt, das alle Datenunregelmäßigkeiten über Datenquellen im Unternehmen und versteckte Datenprobleme, die Datenprojekte gefährden können, findet und analysiert. Die Ausführung eines Profils für eine beliebige Datenquelle im Unternehmen vermittelt Ihnen ein gutes Verständnis für die Stärken und Schwächen ihrer Daten und Metadaten.

Mithilfe von Informatica Analyst und Informatica Developer können Sie die Quell- und Metadaten analysieren. Analytiker und Entwickler können diese Tools nutzen, um zusammenzuarbeiten, um Datenqualitätsprobleme zu identifizieren und um Datenbeziehungen zu analysieren. Je nach Ihrer Arbeitsplatzrolle können Sie die Funktionen des Analyst Tools oder des Developer Tools nutzen. Der Profilierungsgrad, den Sie durchführen können, unterscheidet sich je nach verwendetem Tool.

Die folgenden Aufgaben können sowohl im Developer-Tool als auch im Analyst-Tool ausgeführt werden:

- Durchführen von Spalten-Profiling. Der Prozess schließt das Erkennen der Anzahl an eindeutigen Werten, Null-Werten und Datenmustern in einer Spalte ein.
- Hinzufügen von Regeln zu Spaltenprofilen.
- Stellen Sie die abgeleiteten Datentypen in den Profilergebnissen wiederher.
- Verwenden von Scorecards, um die Datenqualität zu überwachen.
- Generieren eines Mappings aus einem Profil.

Profiling-Prozess

Wenn Sie mit einem Datenintegrationsprojekt beginnen, ist Profiling oft der erste Schritt. Sie können Profile erstellen, um den Inhalt, die Qualität und die Struktur von Datenquellen zu analysieren. Als Teil des Profiling-Prozesses erkunden Sie die Metadaten von Datenquellen.

Sie können unterschiedliche Profile für verschiedene Arten von Datenanalyse wie z. B. ein Spaltenprofil verwenden. Sie können Datenqualitätsprobleme aufdecken und dokumentieren. Schließen Sie die folgenden Aufgaben ab, um das Profiling auszuführen:

1. Finden und Analysieren des Inhalts der Daten in den Datenquellen. Umfasst Datentypen, Werthäufigkeit, Musterhäufigkeit und Datenstatistiken, wie Mindestwert und Maximalwert.
2. Prüfen von Profiling-Ergebnissen.
3. Erstellen von Referenzdaten.
4. Drilldown der Profilergebnisse.
5. Dokumentieren von Datenproblemen.
6. Erstellen und Ausführen von Regeln.
7. Erstellen von Scorecards, um die Datenqualität zu überwachen.

Profiling-Tools

Verwenden Sie mehrere Informatica-Tools zum Verwalten des Profiling-Prozesses.

Mit den folgenden Tools können Sie den Profiling-Prozess verwalten:

Informatica Administrator

Verwalten von Benutzern, Gruppen, Berechtigungen und Rollen. Sie können den Analyst Service administrieren und Berechtigungen für Projekte und Objekte in Informatica Analyst verwalten. Mit diesem Tool können Sie die Zugangsberechtigungen in Informatica Developer kontrollieren.

Informatica Developer

In diesem Tool erstellen Sie Profile und führen sie aus, um die Metadaten von einer oder mehreren Datenquellen zu finden und zu analysieren. Profile werden mit einem Assistenten erstellt.

Informatica Analyst

Sie können ein Spaltenprofil im Analyst-Tool ausführen. Nachdem Sie ein Profil ausgeführt haben, können Sie ein Drilldown für Datenzeilen in einer Datenquelle durchführen.

Profil-Komponenten

Ein Profil hat mehrere Komponenten, die Sie verwenden können, um den Inhalt und die Struktur der Datenquellen effektiv zu analysieren.

Ein Profil hat folgende Komponenten:

Filter

Erstellt eine Teilmenge der ursprünglichen Daten, die spezifische Kriterien erfüllt. Anschließend können Sie ein Profil über diese Beispieldaten ausführen.

Regel

Geschäftslogik, die Bedingungen definiert, welche auf Daten angewendet werden, wenn ein Profil ausgeführt wird. Fügen Sie eine Regel zum Profil hinzu, um Daten zu validieren.

Tag

Metadaten, die ein Objekt im Model Repository anhand der Unternehmensanwendung definieren. Erstellen Sie Tags, um Objekte entsprechend ihrer Unternehmensanwendung zu gruppieren.

Kommentar

Beschreibung des Profils. Mit Hilfe von Kommentaren können Sie mit anderen Benutzern des Analyst-Tools bzw. des Developer-Tools Informationen über Profile gemeinsam nutzen.

Scorecard

Eine Scorecard ist die grafische Darstellung von gültigen Werten für eine Spalte oder die Ausgabe einer Regel in Profilergebnissen. Mit Scorecards können Sie den Fortschritt der Datenqualität messen.

KAPITEL 2

Spaltenprofilkonzepte

Dieses Kapitel umfasst die folgenden Themen:

- Spaltenprofilkonzepte - Übersicht Profile - Übersicht, 16
- Spaltenprofiloptionen, 17
- Repository-Profilsperren und Verwaltung versionierter Profile, 17
- Scorecards, 18

Spaltenprofilkonzepte - Übersicht Profile - Übersicht

Ein Spaltenprofil bestimmt die Charakteristika der Spalte in einer Datenquelle, wie Werthäufigkeit, Prozentsätze und Muster.

Anhand des Spalten-Profiling ist es möglich, folgende Erkenntnisse über die Daten zu gewinnen:

- Die Anzahl der eindeutigen, nicht eindeutigen und Nullwerte in jeder Spalte, ausgedrückt als Zahl und Prozentsatz.
- Das Muster der Daten in jeder Spalte und die Häufigkeit, mit der diese Werte vorkommen.
- Statistiken zu den Spaltenwerten, wie Maximal- und Minimallänge der Werte und erster und letzter Wert in jeder Spalte.
- Dokumentierte Datentypen, abgeleitete Datentypen und mögliche Konflikte zwischen den dokumentierten und abgeleiteten Datentypen.
- Sonderfälle bei Mustern und Werthäufigkeiten.

Sie können die folgenden Optionen beim Erstellen oder Bearbeiten eines Spaltenprofils konfigurieren:

- Spaltenprofiloptionen: Sie können die Spalten auswählen, für die ein Profil ausgeführt werden soll, sowie eine Stichproben- und Drilldown-Option festlegen.
- Bearbeiten, löschen oder fügen Sie Filter und Regeln hinzu.

In den Profilergebnissen können Sie einem Profil und den Spalten in einem Profil Kommentare und Tags hinzufügen. Sie können Spalten Geschäftsbegriffe zuweisen.

Das Modellrepository sperrt Profile mit den Repository-Profilsperren, damit Benutzer keine Arbeiten überschreiben können. Das Versionsverwaltungssystem speichert mehrere Versionen eines Profils und weist jeder Version eine Versionsnummer zu. Sie können ein Profil auschecken und nach der Durchführung von Änderungen wieder einchecken. Sie können das Auschecken eines Profils rückgängig machen, bevor Sie es wieder einchecken.

Erstellen von Scorecards, um die Datenqualität zu prüfen. Bevor und nachdem Sie Regeln auf Profile anwenden, erstellen Sie Scorecards, damit Sie eine grafische Darstellung der gültigen Werte für Spalten anzeigen können.

Verwenden Sie den Scheduler-Dienst zum Planen von Profil- und Scorecard-Ausführungen, die zu einem bestimmten Zeitpunkt oder in Intervallen ausgeführt werden sollen. Der Scheduler-Dienst verwaltet Pläne für Profile, Scorecards, bereitgestellte Zuordnungen und bereitgestellte Arbeitsabläufe. Sie können Pläne in Informatica Administrator erstellen, verwalten und ausführen.

Spaltenprofiloptionen

Wenn Sie ein Profil erstellen, können Sie den Profilassistenten zum Definieren von Filtern, Regeln, Drilldown-Optionen, Stichprobenoptionen und Verbindungen verwenden. Mit diesen Optionen wird festgelegt, wie das Profil Zeilen aus der Datenquelle liest.

Sie können die folgenden Optionen in einem Spaltenprofil, einem Datendomänen- oder Enterprise-Erkennungsprofil definieren:

- Filter. Sie können Filter erstellen und auf ein Profil anwenden.
- Regeln. Beim Erstellen eines Profils können Sie Regeln hinzufügen. Sie können die Regeln wiederverwenden, die Sie im Analyst Tool oder Developer Tool erstellen.
- Drilldown-Optionen. Sie können aktuelle Daten in der Datenquelle oder Profildaten lesen, die im Profiling-Warehouse zwischengespeichert werden.
- Stichprobenoptionen. Sie können eine Stichprobenoption auswählen, um die Anzahl der Zeilen festzulegen, für die ein Profil ausgeführt werden soll.
- Verbindung. Sie können „Nativ“, „Hive“ oder „Hadoop“ als Laufzeitumgebung auswählen.

Repository-Profilsperren und Verwaltung versionierter Profile

Das Modellrepository sperrt Profile, damit Benutzer keine Arbeiten überschreiben können. Wenn Sie mit der Bearbeitung eines Profils beginnen, wird das Profil gesperrt, damit andere Benutzer keine Änderungen darin speichern können. Die Sperre wird aufgehoben, wenn Sie das Profil speichern. Durch die Verwaltung versionierter Profile werden Versionen eines Profils erstellt. Sie können sich die Versionshistorie ansehen.

Das Modellrepository sperrt ein Profil, wenn Sie es im Developer Tool oder Analyst Tool bearbeiten. Falls das Tool unerwartet beendet wird, wird die Sperre beibehalten. Wenn Sie wieder eine Verbindung mit dem Modellrepository herstellen, können Sie sich so die Profile ansehen, die Sie gesperrt haben. Sie können mit der Bearbeitung der Profile fortfahren oder die Profile entsperren.

Wenn das Modellrepository in ein Versionsverwaltungssystem integriert ist, können Sie Versionen eines Profils verwalten. Sie können Profile beispielsweise aus- und einchecken, Auscheckvorgänge rückgängig machen, spezifische historische Versionen eines Profils sowie die von Ihnen ausgecheckten Profile anzeigen. Weitere Informationen zu Repository-Objektsperren und zur Verwaltung von versionierten Objekten im Analyst Tool finden Sie im *Analyst Tool-Handbuch*. Weitere Informationen zu Repository-Sperren von Objekten und zur Verwaltung von versionierten Objekten im Developer Tool finden Sie im *Developer Tool-Handbuch*.

Scorecards

Eine Scorecard ist die grafische Darstellung der gültigen Werte für eine Spalte oder die Ausgabe einer Regel in Profilergebnissen. Mit Scorecards können Sie den Fortschritt der Datenqualität messen. Sie können eine Scorecard aus einem Profil erstellen und den Fortschritt der Datenqualität im Laufe der Zeit überwachen.

Eine Scorecard beinhaltet mehrere Komponenten, z. B. Metriken, Metrikgruppen und Schwellenwerte. Nachdem Sie ein Profil ausgeführt haben, können Sie die Quellspalten als Metriken zu einer Scorecard hinzufügen und die gültigen Werte für die Metriken konfigurieren. Mit Scorecards können Unternehmen den Wert der Datenqualität ermitteln, indem sie die Kosten für fehlerhafte Daten auf der Metrik- und Scorecard-Ebene verfolgen. Um die Kosten für fehlerhafte Daten für eine Metrik zu ermitteln, weisen Sie der Metrik eine Kosteneinheit zu und richten Sie feste oder variable Kosten ein. Wenn Sie die Scorecard ausführen, enthalten die Scorecard-Ergebnisse die Kosten für fehlerhafte Daten für jede Metrik sowie den Gesamtkostenwert für alle Metriken.

Verwenden Sie eine Metrikgruppe, um verwandte Scores in einer Scorecard zu einem Satz zu kategorisieren. Ein Schwellenwert gibt den Bereich fehlerhafter Daten als Prozentsatz an, der für Spalten in einem Datensatz akzeptabel ist. Sie können Schwellenwerte für gute, akzeptable oder inakzeptable Datenbereiche einstellen.

Geben Sie bei der Ausführung einer Scorecard an, ob Sie für die Score-Metriken einen Drilldown basierend auf Live- oder zwischengespeicherten Daten durchführen möchten. Nachdem Sie eine Scorecard ausgeführt und die Scores angezeigt haben, können Sie einen Drilldown für jede Metrik durchführen, um gültige und ungültige Datensätze zu erkennen. Außerdem können Sie für jede Metrik oder Metrikgruppe in einer Scorecard die Scorecard-Herkunft anzeigen lassen. Zur effektiven Verfolgung der Datenqualität können Sie score- und kostenbasierte Trenddiagramme verwenden. In diesen Diagrammen wird die Änderung der Scores sowie der Kosten für fehlerhafte Daten über einen bestimmten Zeitraum überwacht.

Das Profiling Warehouse speichert die Scorecard-Statistiken und Konfigurationsdaten. Sie können auch eine Drittanbieter-Anwendung so konfigurieren, dass sie die Scorecard-Ergebnisse liest und Berichte erzeugt. Sie können die Scorecard-Ergebnisse auch in einer Webanwendung, einem Portal oder Bericht anzeigen, wie z. B. einem Business Intelligence-Bericht.

KAPITEL 3

Wiederherstellungskonzepte

Dieses Kapitel umfasst die folgenden Themen:

- Wiederherstellungskonzepte - Übersicht, 19
- Wiederherstellungsaufgaben, 20

Wiederherstellungskonzepte - Übersicht

Wiederherstellung ist der Prozess des Validierens und Verwaltens erkannter Metadaten einer Datenquelle, um die Metadaten für ihre Verwendung und das Reporting anzupassen.

Sie können die folgenden abgeleiteten Profilergebnisse wiederherstellen:

- Datentypen
- Datendomänen
- Primärschlüssel
- Fremdschlüssel

Sie stellen abgeleitete Profilergebnisse wieder her, um die Metadaten von Spalten, Datendomänen und Datenobjektbeziehungen in den Datenbanken und Schemas genau darzustellen. Sie können dann die wichtigsten Metadaten finden, indem Sie mithilfe der Enterprise-Erkennungssuche in mehreren Repositorys gleichzeitig nach Informationen suchen. Sie können die wichtigsten Metadaten ebenfalls finden, wenn Sie das Fremdschlüssel-Beziehungsdiagramm in den Enterprise-Erkennungsergebnissen anzeigen.

Sie können die spezifischen Metadateninferenzen, die ein Profil als Teil der Profilausführung generiert, wiederherstellen. Sie können die abgeleiteten Datentypen in den Spaltenprofil- und Datendomänenerkennungsergebnissen beispielsweise genehmigen oder ablehnen. Sie können die abgeleiteten Primärschlüssel und Fremdschlüssel in den Enterprise-Erkennungsergebnissen ebenfalls genehmigen oder ablehnen.

Wiederherstellungsaufgaben

Sie können Profilergebnisse nach der Profilausführung wiederherstellen. Sie können eine Wiederherstellungsentscheidung rückgängig machen, die Sie vor einer Profilausführung getroffen haben.

Sie können die folgenden Wiederherstellungsaufgaben im Analyst-Tool ausführen:

- Genehmigen Sie die abgeleiteten Datentypen für mehrere Spalten oder Datendomänen oder lehnen Sie sie ab.
- Stellen Sie den Ableitungsstatus genehmigter oder abgelehnter Datentypen wieder her.
- Stellen Sie genehmigte oder abgelehnte Datendomänen im abgeleiteten Status wieder her.
- Abgelehnte Ergebniszeilen anzeigen oder verbergen
- Schließen Sie Spalten basierend auf bestimmten Metadateneinstellungen, wie z. B. genehmigten Datentypen und Datendomänen, aus Profilausführungen aus.

Sie können die folgenden Wiederherstellungsaufgaben im Developer-Tool ausführen:

- Genehmigen Sie die abgeleiteten Datentypen für mehrere Spalten oder lehnen Sie sie ab.
- Stellen Sie den Ableitungsstatus genehmigter oder abgelehnter Datentypen wieder her.
- Stellen Sie genehmigte oder abgelehnte Datendomänen im abgeleiteten Status wieder her.
- Abgelehnte Ergebniszeilen anzeigen oder verbergen
- Datenobjekte in den Primärschlüssel-Erkennungsergebnissen genehmigen oder ablehnen
- Enterprise-Erkennungsergebnisse, einschließlich Fremdschlüssel-Erkennungsergebnisse genehmigen oder ablehnen
- Schließen Sie Spalten basierend auf bestimmten Metadateneinstellungen, wie z. B. genehmigten Datentypen und Datendomänen, aus Profilausführungen aus.

Teil II: Profiling mit Informatica Analyst

Dieser Teil enthält die folgenden Kapitel:

- Spaltenprofile in Informatica Analyst, 22
- Regeln in Informatica Analyst, 31
- Spaltenprofilergebnisse in Informatica Analyst, 36
- Scorecards in Informatica Analyst, 59

KAPITEL 4

Spaltenprofile in Informatica Analyst

Dieses Kapitel umfasst die folgenden Themen:

- Spaltenprofile in Informatica Analyst - Übersicht, 22
- Spalten-Profilingprozess, 23
- Profil-Optionen, 23
- Betriebssystemprofile in Informatica Analyst - Übersicht, 25
- Repository-Objektsperren und teambasierte Entwicklung - Übersicht, 25
- Erstellen eines Spaltenprofils, 26
- Bearbeiten eines Spaltenprofils, 28
- Ausführen eines Profils, 28
- Synchronisieren eines Einfachdatei-Datenobjekts, 29
- Synchronisieren eines relationalen Datenobjekts, 30

Spaltenprofile in Informatica Analyst - Übersicht

Wenn Sie ein Profil erstellen, wählen Sie die Spalten in dem Datenobjekt aus, in dem ein Profil ausgeführt werden soll. Sie können die Stichproben- und Drilldown-Optionen für schnelleres Profiling konfigurieren. Sie können eine Laufzeitumgebung auswählen. Wenn Sie ein Profil erstellen, können Sie dem Profil Regeln und Filter hinzufügen. Nach dem Ausführen des Profils können Sie die Profiling-Statistik untersuchen, um die Daten zu verstehen.

Sie haben die Möglichkeit, breite Tabellen und Einfachdateien mit maximal 1.000 Spalten zu profilieren. Wenn Sie ein Profil erstellen oder ausführen, können Sie alle Spalten oder jede Spalte für ein Profil auswählen. Sie können alle Spalten auswählen, für die ein Drilldown durchgeführt werden soll, und Werthäufigkeiten für diese Spalten anzeigen.

Sie können Spaltenprofile mithilfe der folgenden Methoden in Informatica Analyst erstellen:

- Im Arbeitsbereich **Bibliothek** mit der rechten Maustaste auf das Datenobjekt klicken, um ein Profil zu erstellen.
- Die Standardoptionen verwenden, um ein standardmäßiges Spaltenprofil zu erstellen.
- Die Einstellungen für das Profil anpassen, um ein benutzerdefiniertes Profil zu erstellen.

Hinweis: Sie können das Profil in Avro-, JSON-, Parquet- und XML-Datenquellen anzeigen und ausführen. Sie können ein Spaltenprofil in Avro-, JSON-, Parquet- und XML-Datenquellen in Informatica Developer erstellen und bearbeiten.

Spalten-Profilingprozess

Als Teil des Spalten-Profilingprozesses können Sie entweder alle Quellspalten zum Profiling aufnehmen oder bestimmte Spalten auswählen. Sie können außerdem die Standardprofiloptionen übernehmen oder die Stichprobenoptionen, Drilldown-Optionen und die Laufzeitumgebung konfigurieren.

Die folgenden Schritte beschreiben den Spalten-Profilingprozess:

1. Wählen Sie einen Namen, eine Beschreibung und einen Speicherort für das Spaltenprofil aus.
2. Wählen Sie ein importiertes Datenobjekt oder eine externe Quelle aus, für die das Profil ausgeführt werden soll.
3. Optional können Sie die Quelldaten in der Vorschau anzeigen.
4. Wählen Sie die Spalten aus, für die das Profil ausgeführt werden soll.
5. Geben Sie an, ob das Profil mit den Standardoptionen erstellt werden soll, oder ändern Sie die Standardoptionen. Zu den konfigurierbaren Optionen gehören Stichprobenoptionen, Drilldown-Optionen und die Laufzeitumgebung.
6. Fügen Sie Regeln und Filter optional hinzu, wenn Sie das Profil erstellen.
7. Führen Sie das Profil aus.

Hinweis: Berücksichtigen Sie die folgenden Regeln und Richtlinien für Spaltennamen und das Profiling von mehrsprachigen und Unicode-Daten:

- Sie können mehrsprachige Daten aus verschiedenen Quellen profilieren und die Profilergebnisse basierend auf den Gebietsschemaeinstellungen des Browsers anzeigen. Das Analyst Tool ändert die Datentypen „Datum/Uhrzeit“, „Numerisch“ und „Dezimal“ basierend auf dem Gebietsschema des Browsers.
- Sortieren an mehrsprachigen Daten. Sie können mehrsprachige Daten sortieren. Die Sortierreihenfolge wird im Analysetool basierend auf dem Browser-Gebietsschema angezeigt.
- Um Unicode-Daten in einer DB2-Datenbank zu profilieren, legen Sie in der Datenbank die Datenbankumgebungsvariable DB2CODEPAGE fest und starten den Datenintegrationsdienst neu.

Profil-Optionen

Zu den Profiloptionen gehören Optionen für Datenstichproben und Optionen für den Daten-Drilldown. Sie können diese Optionen konfigurieren, wenn Sie ein Spaltenprofil für ein Datenobjekt erstellen oder bearbeiten.

Verwenden Sie den Arbeitsbereich **Erkennung** zum Konfigurieren der Profiloptionen. Sie können ein Profil mit den Standardoptionen für Spalten und Stichproben sowie mit Drilldown-Optionen erstellen. Verwenden Sie die Drilldown-Optionen, um Live-Daten oder zwischengespeicherte Daten auszuwählen.

Stichprobenoptionen

Stichprobenoptionen bestimmen die Anzahl der Zeilen, für die das Analyst Tool ein Profil ausführt. Sie können Stichprobenoptionen konfigurieren, wenn Sie ein Profil definieren oder wenn Sie ein Profil ausführen.

Die folgende Tabelle beschreibt die Stichprobenoptionen für ein Profil:

Option	Beschreibung
Alle Zeilen	Wählt alle Zeilen im Datenobjekt aus.
Stichprobe erste <number> Zeilen	Die Anzahl der Zeilen, für die Sie das Profil ausführen möchten. Das Analyst Tool wählt die Zeilen aus den ersten Zeilen in der Quelle aus.
Zufällige Stichprobe <number> Zeilen	Die Anzahl der Zeilen für eine zufällige Stichprobe, für die das Profil ausgeführt werden soll. Zufällige Stichproben zwingen das Analyst Tool zum Durchführen eines Drilldowns für zwischengespeicherte Daten. Beachten Sie, dass sich diese Option auf die Drilldown-Leistung auswirken kann.
Zufallsstichprobe	Größe der zufälligen Stichprobe basierend auf der Anzahl der Zeilen im Datenobjekt. Zufällige Stichproben zwingen das Analyst Tool zum Durchführen eines Drilldowns für zwischengespeicherte Daten. Beachten Sie, dass sich diese Option auf die Drilldown-Leistung auswirken kann.
Genehmigte Datentypen und Datendomänen in den nachfolgenden Profilausführungen aus der Datentyp- und Datendomänenableitung ausschließen	Schließt den genehmigten Datentyp oder die genehmigte Datendomäne bei der nächsten Profilausführung aus der Datentyp- und Datendomänenableitung aus.

Drilldown-Optionen

Sie können Drilldown-Optionen konfigurieren, wenn Sie ein Profil definieren oder bearbeiten.

Die folgende Tabelle beschreibt die Drilldown-Optionen für ein Profil:

Optionen	Beschreibung
Live	Führt einen Drilldown für Live-Daten zum Lesen aktueller Daten in der Datenquelle durch.
Zwischengespeichert	Führt einen Drilldown für zwischengespeicherte Daten zum Lesen von Profildaten durch, die im Profiling-Warehouse zwischengespeichert wurden.
Spalten auswählen	Kennzeichnet Spalten für den Drilldown, die nicht für das Profiling ausgewählt wurden.

Laufzeitumgebungsoptionen

Sie können die Option „Nativ“, „Hive“ oder „Hadoop“ als Laufzeitumgebung für ein Spaltenprofil auswählen. Informatica Analyst legt die Laufzeitumgebung in der Profildefinition fest, nachdem Sie eine

Laufzeitumgebung ausgewählt haben. Die Laufzeitumgebung hat keine Auswirkungen auf die Profilergebnisse.

In der folgenden Tabelle werden die Laufzeitumgebungsoptionen für ein Spaltenprofil beschrieben:

Option	Beschreibung
Nativ	Das Analyst Tool sendet die Profiljobs an das Profilerstellungsdienst-Modul. Das Profilerstellungsdienst-Modul teilt die Profiljobs dann in mehrere Zuordnungen auf. Der Datenintegrationsdienst führt diese Zuordnungen aus und schreibt die Profilergebnisse in das Profiling-Warehouse.
Hive	Der Datenintegrationsdienst übergibt die Profillogik an die Hive-Engine auf dem Hadoop-Cluster, um die Profile auszuführen.
Hadoop	Der Datenintegrationsdienst übergibt die Profillogik an die Blaze-Engine auf dem Hadoop-Cluster, um Profile auszuführen.

Betriebssystemprofile in Informatica Analyst - Übersicht

Sie können ein Betriebssystemprofil im Analyst Tool auswählen. Nach Auswahl eines Betriebssystemprofils erstellt der Datenintegrationsdienst die Spaltenprofile, Enterprise-Erkennungsprofile und Scorecards basierend auf der Berechtigung des Betriebssystemprofilbenutzers und führt diese aus.

Das Analyst Tool verwendet das Standardprofil zum Ausführen von Profilen und Scorecards. Wenn nur ein Betriebssystemprofil vorhanden ist, wird dieses standardmäßig ausgewählt. Wenn mehrere Betriebssystemprofile vorhanden sind, können Sie eines der Betriebssystemprofile auswählen.

Auswählen eines Betriebssystemprofils

Sie können ein Betriebssystemprofil in Informatica Analyst auswählen. Der Datenintegrationsdienst verwendet die Berechtigungen des Betriebssystemprofilbenutzers zum Ausführen der Profiling-Jobs.

1. Klicken Sie im Kopfzeilenbereich von Informatica Analyst auf **<Benutzername>** > **Einstellungen**.

   Das Dialogfeld **Einstellungen** wird angezeigt.
2. Wählen Sie ein Betriebssystemprofil aus. Klicken Sie auf **Speichern**.

Repository-Objektsperren und teambasierte Entwicklung - Übersicht

Das Modellrepository sperrt Profile, damit Benutzer nicht die Arbeit anderer Benutzer überschreiben können. Wenn das Modellrepository in ein Versionsverwaltungssystem integriert ist, speichert es mehrere Versionen von Objekten und weist einer Version eine Versionsnummer zu. Sie können Profile aus- und einchecken und

Auscheckvorgänge rückgängig machen. Sie können eine spezifische Version eines von Ihnen ausgecheckten Profils anzeigen.

Wenn Sie im Analyst Tool mit der Bearbeitung eines Profils beginnen, wird das Profil vom Modellrepository gesperrt, damit andere Benutzer das Profil nicht bearbeiten können. Beim Speichern des Profils wird die Sperre beibehalten. Wenn Sie das Profil schließen, wird die Sperre durch das Modellrepository aufgehoben.

Das Modellrepository verhindert, dass Profile von anderen Teammitgliedern bei der Verwaltung von versionierten Objekten überschrieben werden. Wenn Sie versuchen, ein von einem anderen Benutzer ausgechecktes Profil zu bearbeiten, erhalten Sie eine Benachrichtigung, in der der Benutzer angegeben ist, der das Profil ausgecheckt hat. Sie können ein ausgechecktes Profil im schreibgeschützten Modus öffnen oder es unter einem anderen Namen speichern.

Sie können im Dialogfeld „Profileigenschaften“ eine Version des Profils auswählen, um die Profildefinition für diese Version anzuzeigen. Auf die Option „Profileigenschaften“ können Sie im Menü „Aktionen“ zugreifen. Weitere Informationen zu Repository-Objektsperren und zur Verwaltung von versionierten Objekten finden Sie im *Analyst Tool-Handbuch*.

Erstellen eines Spaltenprofils

Sie können ein benutzerdefiniertes Profil oder ein Standardprofil erstellen. Wenn Sie ein benutzerdefiniertes Profil erstellen, können Sie die Spalten, Stichprobenzeilen und Drilldown-Optionen konfigurieren. Wenn Sie ein Standardprofil erstellen, wird das Spaltenprofil und die Datendomänenerkennung für den gesamten Datensatz mit allen Datendomänen ausgeführt.

1. Klicken Sie im Arbeitsbereich **Erkennung** auf **Profil** oder wählen Sie im Kopfzeilenbereich **Neu** > **Profil** aus.

   Hinweis: Sie können im Arbeitsbereich **Bibliothek** mit der rechten Maustaste auf das Datenobjekt klicken und ein Profil erstellen. In diesem Profil werden der Profilname, der Name des Speicherorts und das Datenobjekt aus den Datenobjekteigenschaften extrahiert. Sie können ein Standardprofil erstellen oder die Einstellungen anpassen, um ein benutzerdefiniertes Profil zu erstellen.

   Der Assistent **Neues Profil** wird eingeblendet.
2. Die Option **Einzelquelle** ist standardmäßig ausgewählt. Klicken Sie auf **Weiter**.
3. Geben Sie im Bildschirm **Allgemeine Eigenschaften angeben** einen Namen und eine optionale Beschreibung für das Profil ein. Wählen Sie im Feld „Speicherort“ das Projekt oder den Ordner aus, in dem das Profil erstellt werden soll. Klicken Sie auf **Weiter**.
4. Klicken Sie im Bildschirm **Quelle auswählen** auf **Auswählen**, um ein Datenobjekt auszuwählen, oder auf **Neu**, um ein Datenobjekt zu importieren. Klicken Sie auf **Weiter**.
   - Wählen Sie im Dialogfeld **Datenobjekt auswählen** ein Datenobjekt aus. Klicken Sie auf **OK**.
     Im Bereich „Eigenschaften“ werden die Eigenschaften des ausgewählten Datenobjekts angezeigt. Im Bereich „Datenvorschau“ werden die Spalten im Datenobjekt angezeigt.
   - Im Dialogfeld **Neues Datenobjekt** können Sie eine Verbindung, ein Schema, eine Tabelle oder eine Ansicht auswählen, um ein Profil dafür zu erstellen. Zudem haben Sie die Möglichkeit, einen Speicherort auszuwählen und einen Ordner zu erstellen, um das Datenobjekt zu importieren. Klicken Sie auf **OK**.
5. Wählen Sie im Bildschirm **Quelle auswählen** die Spalten aus, für die ein Profil ausgeführt werden soll. Wählen Sie optional **Name** aus, um alle Spalten auszuwählen. Klicken Sie auf **Weiter**.

Alle Spalten sind standardmäßig ausgewählt. Das Analyst Tool listet Spalteneigenschaften auf, wie z. B. Name, Datentyp, Genauigkeit, Skalierung, Größenordnung und auf Null setzbar, und trägt zum Primärschlüssel für jede Spalte bei.

6. Legen Sie im Bildschirm **Einstellungen angeben** die Ausführung eines Spaltenprofils, einer Datendomänenerkennung oder eines Spaltenprofils mit Datendomänenerkennung fest. Standardmäßig ist die Option für das Spaltenprofil ausgewählt.
   - Wählen Sie **Spaltenprofil ausführen** zum Ausführen eines Spaltenprofils aus.
   - Wählen Sie **Datendomänenerkennung ausführen** zur Durchführung der Datendomänenerkennung aus. Wählen Sie im Bereich **Datendomäne** die zu erkennenden Datendomänen und ein Konformitätskriterium aus und im Dialogfeld **Spaltenauswahl für Datendomänenerkennung bearbeiten** die Spalten für die Datendomänenerkennung.
   - Wählen Sie **Spaltenprofil ausführen** und **Datendomänenerkennung ausführen** aus, um das Spaltenprofil mit Datendomänenerkennung auszuführen. Wählen Sie die Optionen für Datendomänen im Bereich **Datendomäne** aus.

     Hinweis: Die Spalten, die Sie auswählen, werden standardmäßig für das Spaltenprofil und die Datendomänenerkennung verwendet. Klicken Sie auf **Bearbeiten**, um Spalten für die Datendomänenerkennung auszuwählen bzw. ihre Auswahl aufzuheben.
   - Wählen Sie Daten, Spalten oder Daten und Spalten für die Datendomänenerkennung aus.
   - Wählen Sie eine Stichprobenoption aus. Sie können im Bereich **Spaltenprofil ausführen für** als Stichprobenoption **Alle Zeilen (vollständige Analyse)**, **Erste Stichprobe**, **Zufallsstichprobe** oder **Zufallsstichprobe (auto)** auswählen. Diese Option gilt für das Spaltenprofil und die Datendomänenerkennung.
   - Wählen Sie eine Drilldown-Option aus. Sie können im Bereich **Drilldown** die Drilldown-Option **Live** oder **Zwischengespeichert** auswählen bzw. **Aus**, um den Drilldown zu deaktivieren. Optional können Sie auf **Spalten auswählen** klicken, um Spalten für den Drilldown auszuwählen. Sie können die Datentyp- und Datendomäneninferenz für Spalten mit einem genehmigten Datentyp oder einer genehmigten Datendomäne weglassen.
   - Wählen Sie **Nativ**, **Hive** oder **Hadoop** als Laufzeitumgebung aus. Bei Verwendung von „Hive“ klicken Sie auf **Auswählen**, um eine Hive-Verbindung im Dialogfeld **Hive-Verbindung auswählen** auszuwählen. Bei Verwendung von „Hadoop“ klicken Sie auf **Auswählen**, um eine Hadoop-Verbindung im Dialogfeld **Hadoop-Verbindung auswählen** auszuwählen.
7. Klicken Sie auf **Weiter**.

   Der Bildschirm **Regeln und Filter angeben** wird geöffnet.
8. Im Bildschirm **Regeln und Filter angeben** können Sie die folgenden Aufgaben durchführen:
   - Erstellen, bearbeiten oder löschen einer Regel. Sie können vorhandene Regeln auf das Profil anwenden.
   - Erstellen, bearbeiten oder löschen eines Filters.

     Hinweis: Wenn Sie eine Scorecard für dieses Profil erstellen, können Sie die Filter, die Sie für das Profil anlegen, wiederverwenden.
9. Klicken Sie auf **Speichern und beenden**, um das Profil zu erstellen, oder auf **Speichern und ausführen**, um das Profil zu erstellen und auszuführen.

Bearbeiten eines Spaltenprofils

Sie können ein Spaltenprofil ändern, nachdem Sie es ausführen.

1. Wählen Sie im Arbeitsbereich **Bibliothek** das Projekt mit dem Profil aus oder wählen Sie das Profil im Bereich **Objekte** aus.
2. Klicken Sie auf den Profilnamen.

   Die Zusammenfassungsansicht wird im Arbeitsbereich **Erkennung** angezeigt.
3. Falls das Versionsverwaltungssystem aktiviert ist, klicken Sie auf **Aktionen** > **Auschecken**, um das Profil auszuchecken.
4. Klicken Sie auf **Aktionen** > **Profil bearbeiten**.

   Der **Profilassistent** wird angezeigt.
5. Wählen Sie je nach den von Ihnen gewünschten Änderungen eine der folgenden Seitenoptionen aus:
   - **Allgemeine Eigenschaften angeben**. Ändern Sie die allgemeinen Eigenschaften, wie z. B. Name, Beschreibung und Speicherort.
   - **Quelle auswählen**. Wählen Sie eine andere passende Datenquelle und Spalten aus, für die Sie das Profil ausführen.
   - **Einstellungen angeben**. Führen Sie das Spaltenprofil oder das Spaltenprofil mit Datendomänenerkennung aus. Wählen Sie die zu erkennenden Datendomänen aus und bearbeiten Sie die Datendomänenerkennungs-, Stichproben- und Drilldown-Optionen.
   - **Regeln und Filter angeben**. Erstellen, bearbeiten oder löschen Sie Regeln und Filter.
6. Klicken Sie auf **Speichern und beenden**, um die Bearbeitung des Profils abzuschließen, oder auf **Speichern und ausführen**, um das Profil zu bearbeiten und auszuführen.
7. Falls das Versionsverwaltungssystem aktiviert ist, müssen Sie die folgenden Aufgaben durchführen:
   - Klicken Sie auf **Speichern und beenden**, um die Bearbeitung des Profils abzuschließen.
   - Klicken Sie in der Zusammenfassungsansicht auf **Einchecken**, um das Profil einzuchecken.
   - Klicken Sie auf **Aktionen** > **Profil ausführen**, um das Profil auszuführen.

Ausführen eines Profils

Führen Sie ein Profil aus, um eine Datenquelle auf Inhalt und Struktur hin zu analysieren, und wählen Sie Spalten und Regeln für den Drilldown aus. Drilldown ist bei Live- oder zwischengespeicherten Daten für Spalten und Regeln möglich. Wenn Sie ein Profil nur für eine einzelne Spalte oder Regel ausführen möchten, müssen Sie erstmalig ein Profil oder einen Lauf für alle Quellspalten ausgeführt haben.

1. Wählen Sie im Arbeitsbereich **Bibliothek** im Bereich „Projekte“ das Projekt bzw. den Ordner aus, in dem das Profil enthalten ist. Alternativ können Sie das Profil im Bereich **Objekte** auswählen.
2. Klicken Sie auf **Aktionen** > **Öffnen**.

   Die Zusammenfassungsansicht wird im Arbeitsbereich **Erkennung** angezeigt.
3. Klicken Sie auf **Aktionen** > **Profil ausführen.**

   Das Analyst Tool führt eine Profilausführung durch und zeigt die Profilergebnisse in der Zusammenfassungsansicht an.

4. Klicken Sie in der Zusammenfassungsansicht auf eine Spalte, um die Spaltenergebnisse anzuzeigen.

   Die Detailansicht wird angezeigt.

Synchronisieren eines Einfachdatei-Datenobjekts

Sie können die Änderungen an einer externen Einfachdatei-Datenquelle mit deren Datenobjekt im Analyst-Tool synchronisieren. Zum Synchronisieren der Datenobjekte verwenden Sie den Assistenten zum **Synchronisieren von Einfachdateien**.

1. Öffnen Sie den **Bibliotheks**-Workspace.
2. Wählen Sie im Abschnitt **Projekte** ein Einfachdatei-Datenobjekt aus einem Projekt aus.

   Das Analyst-Tool zeigt die Datenvorschau für die Einfachdatei auf der Registerkarte **Datenvorschau** an.
3. Klicken Sie auf die Registerkarte **Eigenschaften**.
4. Klicken Sie im Aktionsmenü auf **Synchronisieren**.

   Der Assistent **Einfachdatei synchronisieren** wird angezeigt.
5. Sie können einen Speicherort durchsuchen oder einen Netzwerkpfad zum Importieren der Einfachdatei eingeben.
   - Um einen Speicherort zu durchsuchen, klicken Sie auf **Datei auswählen**, um die Einfachdatei aus einem Verzeichnis auszuwählen, auf die Ihr Computer zugreifen kann.
   - Um einen Netzwerkpfad einzugeben, wählen Sie **Netzwerkpfad eingeben** aus und konfigurieren Sie den Dateipfad und Dateinamen.
6. Klicken Sie auf **Weiter**.
7. Wählen Sie, ob eine delimitierte Einfachdatei oder eine Einfachdatei mit fester Breite importiert werden soll.
   - Akzeptieren Sie zum Importieren einer delimitierten Einfachdatei die Option **Delimitiert**.
   - Wählen Sie zum Importieren einer Einfachdatei mit fester Breite die Option **Mit fester Breite**.
8. Klicken Sie auf **Weiter**.
9. Konfigurieren Sie die Einfachdateioptionen für die delimitierte Einfachdatei oder Einfachdatei mit fester Breite.
10. Klicken Sie auf **Weiter**.
11. Ändern Sie optional die Spaltenattribute.
12. Klicken Sie auf **Weiter**.
13. Akzeptieren Sie den Standardnamen oder geben Sie einen anderen Namen für die Einfachdatei ein.
14. Geben Sie optional eine Beschreibung ein.
15. Klicken Sie auf **Fertig stellen**.

    Eine Synchronisierungsnachricht fordert Sie auf, die Aktion zu bestätigen.
16. Klicken Sie auf **Ja**, um die Einfachdatei zu synchronisieren.

    Eine Nachricht, die besagt, dass die Synchronisierung abgeschlossen ist, wird angezeigt. Um Details der Metadaten-Änderungen anzuzeigen, klicken Sie auf **Details anzeigen**.
17. Klicken Sie auf **OK**.

Synchronisieren eines relationalen Datenobjekts

Sie können die Änderungen an einer externen relationalen Datenquelle mit deren Tabellendatenobjekt synchronisieren. Externe Datenquelländerungen sind Hinzufügen, Ändern und Entfernen von Quellspalten und Regelspalten.

1. Öffnen Sie den **Bibliotheks**-Workspace.
2. Wählen Sie im Abschnitt **Projekte** ein Tabellendatenobjekt aus einem Projekt aus.

   Das Analyst-Tool zeigt die Datenvorschau für die Tabelle auf der Registerkarte **Datenvorschau** an.
3. Klicken Sie auf die Registerkarte **Eigenschaften**.
4. Klicken Sie im Aktionsmenü auf **Synchronisieren**.

   Sie werden aufgefordert, die Aktion zu bestätigen.
5. Um den Synchronisationsvorgang abzuschließen, klicken Sie auf **OK**.

   Eine Synchronisations-Statusnachricht wird angezeigt.
6. Eine Nachricht, die besagt, dass die Synchronisierung abgeschlossen ist, wird angezeigt.

   Um Details der Metadaten-Änderungen anzuzeigen, klicken Sie auf **Details anzeigen**.
7. Klicken Sie auf **OK**.

KAPITEL 5

Regeln in Informatica Analyst

Dieses Kapitel umfasst die folgenden Themen:

- Regeln in Informatica Analyst - Übersicht, 31
- Vordefinierte Regeln, 32
- Expressionsregeln, 33

Regeln in Informatica Analyst - Übersicht

Eine Regel ist eine Geschäftslogik, die Bedingungen definiert, die bei Ausführung eines Profils auf Quelldaten angewendet werden. Sie können eine Regel zum Profil hinzufügen, um Daten zu validieren.

Sie möchten eine Regel unter Umständen in unterschiedlichen Situationen verwenden. Sie können eine Regel zur Bereinigung von einer oder mehreren Datenspalten hinzufügen. Sie können eine Lookup-Regel hinzufügen, die Informationen enthält, die die Quellendaten nicht zur Verfügung stellen. Sie können eine Regel hinzufügen, um eine Bereinigungsregel für ein Datenqualitäts- oder Datenintegrationsprojekt zu validieren.

Wenn Sie einem Profil eine Regel hinzufügen, können Sie eine Regel erstellen oder eine Regel anwenden.

Für ein Profil können Sie folgende Regeltypen anwenden bzw. erstellen:

- **Expressionsregeln** Verwenden Sie Expressionsfunktionen und Spalten zur Definition der Regellogik. Expressionsregeln erstellen Sie im Analysetool. Der Analytiker kann eine Expressionsregel erstellen und sie zu einer wiederverwendbaren Regel machen, die von anderen Analytikern in mehreren Profilen verwendet werden können.
- **Vordefinierte Regeln** Umfasst wiederverwendbare Regeln, die ein Entwickler im Developer Tool erstellt. Regeln, die ein Entwickler im Developer Tool als Mapplets erstellt, können im Analysetool als wiederverwendbare Regeln erscheinen.

Regeln in einem Spaltenprofil

Sie können Regeln beim Anlegen eines Spaltenprofils erstellen.

Nachdem Sie eine Regel zu einem Profil hinzugefügt haben, können Sie das Profil erneut für die Regelspalte ausführen. Das Analyst Tool zeigt die Profilergebnisse für die Regelspalte in der Zusammenfassungsansicht an. Sie können die Spaltenergebnisse für eine Regel in der Detailansicht anzeigen. Die Ausgabe einer Regel kann in einer oder mehreren virtuellen Spalten erfolgen. Die virtuellen Spalten befinden sich in den Profilergebnissen. Das Analyst Tool führt ein Profil für die virtuellen Spalten aus. Sie verwenden beispielsweise eine vordefinierte Regel, die eine Spalte mit Vor- und Nachnamen in die virtuellen Spalten FIRST_NAME und LAST_NAME aufteilt. Das Analyst Tool führt das Profil für die Spalten FIRST_NAME und LAST_NAME aus.

Hinweis: Wenn Sie ein Regelobjekt löschen, das von anderen Objekttypen referenziert wird, erscheint im Analysetool eine Meldung mit einer Liste der referenzierten Objekttypen. Ermitteln Sie vor dem Löschen einer Regel die Auswirkungen des Löschvorgangs.

Vordefinierte Regeln

Vordefinierte Regeln sind Regeln, die im Developer Tool erstellt oder mit dem Developer Tool und dem Analyst Tool zur Verfügung gestellt wurden. Vordefinierte Regeln müssen Sie auf die Analyst-Tool-Profile anwenden, um Quelldaten zu ändern oder zu validieren.

Vordefinierte Regeln verwenden zum Definieren der Regellogik Umwandlungen. Vordefinierte Regeln können mit Mehrfachprofilen eingesetzt werden. In einem Model Repository ist eine vordefinierte Regel ein Mapplet mit einer Eingabegruppe, einer Ausgabegruppe und Umwandlungen für die Definition der Regellogik.

Vordefinierte Regeln-Prozess

Zum Anwenden einer vordefinierten Regel auf ein Profil verwenden Sie den **Neuen Regelassistenten**.

Sie können die folgenden Schritte durchführen, um eine vordefinierte Regel anzuwenden:

1. Öffnen Sie ein Profil.
2. Wählen Sie eine vordefinierte Regel.
3. Prüfen Sie die Parameter der Regel.
4. Wählen Sie die Eingabespalte. Sie können mehrere Spalten auswählen, wenn Sie die Regel auf mehr als eine Spalte anwenden wollen.
5. Konfigurieren Sie die Profiling-Optionen.

Anwenden einer vordefinierten Regel

Beim Anwenden einer vordefinierten Regel wählen Sie die Regel und konfigurieren die Eingabe- und Ausgabespalten für die Regel. Eine vordefinierte Regel wenden Sie an, wenn Sie eine als wiederverwendbar deklarierte Regel oder eine von einem Entwickler erstellte Regel benutzen möchten.

1. Wählen Sie im Arbeitsbereich **Bibliothek** das Projekt mit dem Profil aus oder wählen Sie das Profil im Bereich **Objekte** aus.
2. Klicken Sie auf **Aktionen** > **Öffnen**, um das Profil zu öffnen.

   Die Zusammenfassungsansicht wird im Arbeitsbereich **Erkennung** angezeigt.
3. Klicken Sie auf **Aktionen** > **Profil bearbeiten**.

   Der **Profilassistent** wird angezeigt.
4. Klicken Sie auf **Regeln und Filter angeben**.
5. Klicken Sie im Bildschirm **Regeln und Filter angeben** auf **Aktionen** > **Vorhandene Regel anwenden** im Bereich **Regeln**.

   Das Dialogfeld **Regelassistenten anwenden** wird geöffnet.
6. Wählen Sie eine Regel aus und klicken Sie auf **Weiter**.
7. Klicken Sie auf **Hinzufügen**.

   Das Dialogfeld **Spalten für Eingabeport auswählen** wird angezeigt.

8. Wählen Sie ein Feld und eine Eingabespalte aus. Klicken Sie auf **OK**.

   Die Eingabe- und Ausgabespalten werden im Dialogfeld **Regelassistenten anwenden** angezeigt.

9. Klicken Sie im Dialogfeld **Regelassistenten anwenden** auf **OK**.

   Die Regel wird im Bildschirm **Regeln und Filter angeben** angezeigt.

Expressionsregeln

Expressionsregeln verwenden Expressionsfunktionen und Spalten zum Definieren der Regellogik. Erstellen Sie Expressionsregeln, und fügen Sie sie zu einem Profil im Analyst Tool hinzu.

Expressionsregeln verwenden Sie zum Ändern oder Validieren von Werten für Spalten in einem Profil. Sie können eine oder mehrere Expressionsregeln für die Verwendung in einem Profil erstellen. Expressionsfunktionen sind SQL-ähnliche Funktionen zum Umwandeln von Quelldaten. Sie können die Expressionsregellogik mit folgenden Funktionstypen erstellen:

- Zeichen
- Umwandlung
- Datenbereinigung
- Datum
- Codierung
- Finanz
- Numerisch
- Wissenschaftlich
- Spezial
- Test

Expressionsregeln-Prozess

Verwenden Sie den **Profilassistenten** zum Erstellen einer Ausdrucksregel sowie zum Hinzufügen der Regel zu einem Profil.

Den Expressions-Editor können Sie verwenden, um Expressionsfunktionen hinzuzufügen, Spalten als Eingabe in die Funktionen zu konfigurieren, die Expression zu validieren und den Rückgabetyp, die Präzision und die Skala zu konfigurieren.

Nach dem Erstellen und Validieren einer Ausdrucksregel können Sie den Genauigkeitswert der Ausgaberegelspalte bearbeiten. Der Genauigkeitswert der Ausgaberegelspalte ist standardmäßig auf 10 festgelegt. Er wird abgeschnitten, wenn die Ausgaberegelspalte den festgelegten Genauigkeitswert überschreitet.

Die Ausgabe einer Expressionsregel ist eine virtuelle Spalte, die den Namen der Regel als Spaltenname verwendet. Das Analyst Tool führt ein Spaltenprofil für die virtuelle Spalte aus. Beispielsweise verwenden Sie eine Expressionsregel zum Validieren einer Postleitzahl. Die Regel gibt 1 zurück, wenn die Postleitzahl gültig ist, und 0, wenn die Postleitzahl nicht gültig ist. Informatica Analyst führt ein Spaltenprofil für die Ausgabewerte 0 und 1 der Regel aus.

Sie können die folgenden Schritte durchführen, um eine Expressionsregel zu erstellen:

1. Öffnen Sie ein Profil.

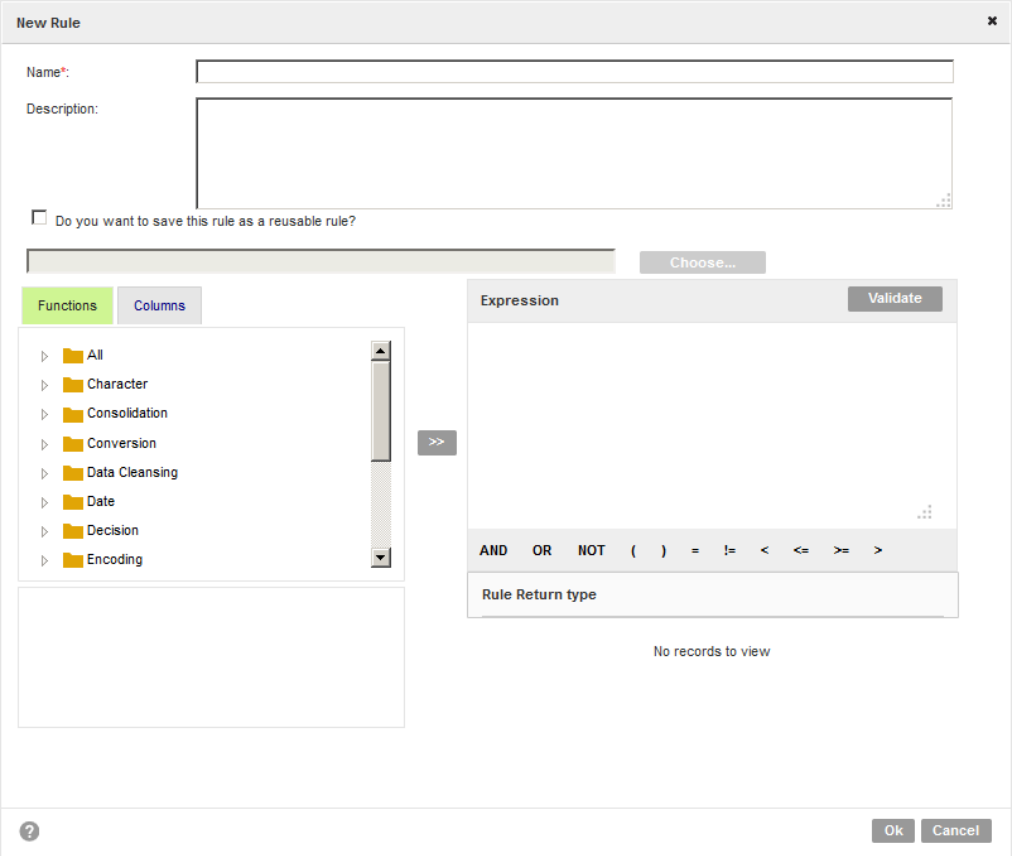
2. Konfigurieren Sie die Regellogik mithilfe von Expressionsfunktionen und Spalten als Parameter.
3. Konfigurieren Sie die Profiling-Optionen.

Erstellen einer Expressionsregel

Verwenden Sie den **Profilassistenten** zum Erstellen einer Ausdrucksregel sowie zum Hinzufügen der Regel zu einem Profil. Erstellen Sie eine Ausdrucksregel zum Validieren von Werten für Spalten in einem Profil.

1. Öffnen Sie ein Profil.
2. Klicken Sie in der Zusammenfassungsansicht auf **Aktionen** > **Profil bearbeiten**, um den **Profilassistenten** zu öffnen.
3. Klicken Sie auf **Regeln und Filter angeben**.
4. Klicken Sie im Bereich „Regeln“ auf **Aktionen** > **Regel hinzufügen**.

   Das Dialogfeld **Neue Regel** wird geöffnet.



5. Geben Sie im Dialogfeld **Neue Regel** einen Namen und eine optionale Beschreibung für die Regel ein. Sie können eine Regel im Bereich „Funktionen“ oder im Bereich „Spalten“ erstellen.
   - Wählen Sie im Bereich „Funktionen“ eine Funktionskategorie aus und klicken Sie auf die Schaltfläche mit den nach rechts weisenden Pfeilen (>>). Geben Sie im Dialogfeld Parameter an und klicken Sie auf **OK**.
     Die Funktion wird zusammen mit den Spalten und Werten im Bereich „Ausdruck“ angezeigt.
   - Wählen Sie im Bereich „Spalten“ eine Spalte aus und klicken Sie auf die Schaltfläche mit den nach rechts weisenden Pfeilen (>>). Die Spalte wird im Bereich „Ausdruck“ angezeigt. Fügen Sie Funktionen, Ausdrücke und Werte hinzu, um eine Regel zu erstellen.

6. Klicken Sie zum Überprüfen der Regel auf **Validieren**.
7. Sie können die Regel optional als wiederverwendbare Regel deklarieren und den Projekt- und Ordnerspeicherort konfigurieren. Wenn Sie eine Regel als wiederverwendbare Regel definieren, können Sie oder andere Benutzer die Regel als vordefinierte Regel in einem anderen Profil verwenden.
8. Klicken Sie auf **OK**.

   Der Bildschirm **Regeln und Filter angeben** wird mit der Regel im Bereich „Regeln“ angezeigt.

KAPITEL 6

Spaltenprofilergebnisse in Informatica Analyst

Dieses Kapitel umfasst die folgenden Themen:

- Spaltenprofilergebnisse in Informatica Analyst - Übersicht, 36
- Zusammenfassungsansicht, 37
- Detailansicht, 39
- Statistiken, 41
- Typen von Profilausführungen, 48
- Vergleichen mehrerer Profilergebnisse - Übersicht, 50
- Spaltenprofil-Drilldown, 55
- Wiederherstellung im Analyst-Tool, 56
- Exportdateien des Spaltenprofils in Informatica Analyst, 57

Spaltenprofilergebnisse in Informatica Analyst - Übersicht

Zeigen Sie Profilergebnisse an, um den Inhalt, die Struktur und die Datenqualität zu verstehen und zu analysieren. Sie können alle Spalten und Regeln in einem Profil in der Zusammenfassungsansicht anzeigen. Sie können die Eigenschaften einer Spalte oder Regel ausführlich in der Detailansicht anzeigen.

Sie können die Profilergebnisse im Arbeitsbereich **Erkennung** anzeigen. In der Kopfzeile der Ansicht werden der Profiltyp, die Anzahl der Spalten im Profil, die Anzahl der Regeln im Profil, Stichprobendaten sowie das Datum und die Uhrzeit der Erstellung angezeigt.

In der Zusammenfassungsansicht können Sie die Eigenschaften aller Spalten als Wert, horizontales Balkendiagramm oder Prozentsatz anzeigen. Sie können Spalteneigenschaften anzeigen, wie z. B. eindeutige, nicht eindeutige und Nullwerte, Muster, Datentypen und Datendomänen. Sie können die Profilergebnisse in der Zusammenfassungsansicht basierend auf den Standardfiltern anzeigen.

In der Detailansicht können Sie eindeutige, nicht eindeutige und Nullwerte, abgeleitete Datentypen, abgeleitete Datendomänen, abgeleitete Muster, Werte sowie Geschäftsbegriffe anzeigen und die Daten in Bereichen in der Vorschau anzeigen.

Sie können Profilergebnisse für die aktuelle, historische und zusammengefasste Ausführung anzeigen. Sie können Profilergebnisse für zwei Profilausführungen vergleichen und die Ergebnisse in der

Zusammenfassungs- und Detailansicht anzeigen. Sie können Profilstatistiken anzeigen und die Daten verwalten. Zu den Profilstatistiken gehören Werte, Muster, Datentypen, Sonderfälle sowie Statistiken für Spalten und Regeln. Sie können Datenerkennung und Drilldown für Daten durchführen.

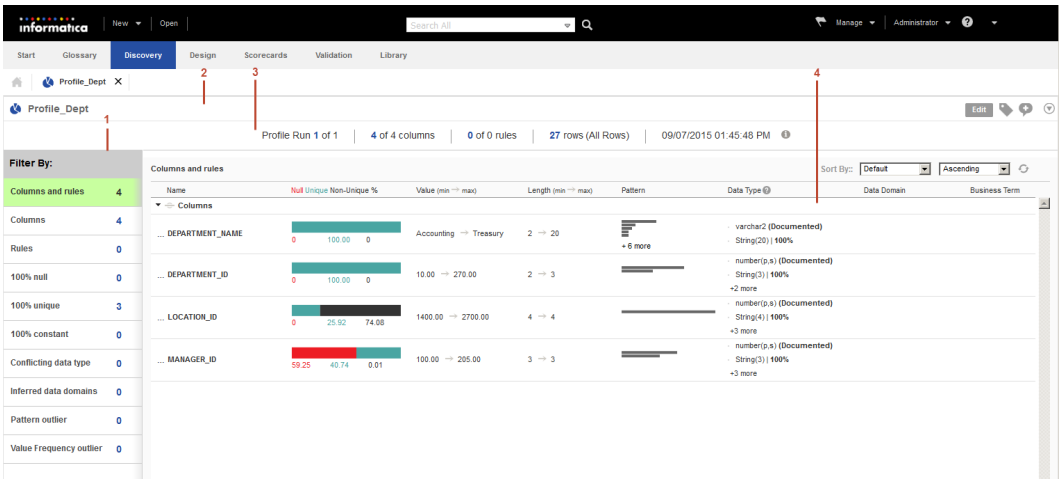
Hinweis: Sie können ein Profil in Avro-, JSON-, Parquet- und XML-Datenquellen anzeigen und ausführen. Sie können Profilergebnisse für die aktuelle, historische und zusammengefasste Ausführung anzeigen und Profilergebnisse für zwei Profilausführungen vergleichen.

Sie können Werthäufigkeiten, Musterhäufigkeiten, Drilldown-Daten, Kommentare, Tags und Geschäftsbegriffe in eine CSV-Datei exportieren. Sie können die Zusammenfassungsinformationen des Profils in eine Microsoft Excel-Datei exportieren, um zur weiteren Analyse alle Daten in einer Datei anzuzeigen. Sie können die Regelinformationen in den Profilergebnissen anzeigen. Welche Profilergebnisse angezeigt werden, ist von der Profilkonfiguration und den Stichprobenoptionen abhängig.

Zusammenfassungsansicht

Die Zusammenfassung der Profilergebnisse wird in der Zusammenfassungsansicht in einem Gitterformat angezeigt. Sie können die Standardfilter in der Zusammenfassungsansicht verwenden, um bestimmte Statistiken anzuzeigen. Wenn Sie beispielsweise „Regeln“ auswählen, werden in der Zusammenfassungsansicht alle Regeln im Profil angezeigt.

Die folgende Abbildung zeigt ein Beispiel für eine grafische Ansicht der Zusammenfassungsansicht:



1. Standardfilter. Sie können die Profilergebnisse in der Zusammenfassungsansicht basierend auf den Standardfiltern anzeigen.
2. Kopfzeile des Profils. Sie können den Profilnamen in der Kopfzeile anzeigen. Über die Schaltfläche „Bearbeiten“ können Sie das Profil bearbeiten. Mit den Tag- und Kommentarsymbolen können Sie Tags bzw. Kommentare hinzufügen oder bearbeiten. Zudem haben Sie die Möglichkeit, die Optionen im Menü „Aktionen“ auswählen.
3. Kopfzeile der Zusammenfassungsansicht. In der Kopfzeile der Zusammenfassungsansicht können Sie profilspezifische Informationen anzeigen. Sie können die Profilausführungsnummer, die Gesamtanzahl der Profilausführungen, die Anzahl der Spalten und Regeln sowie die Anzahl der Zeilen im Profil anzeigen.
4. Zusammenfassungsansicht. Sie können die Eigenschaften für alle Spalten und Regeln im Profil anzeigen.

In der Zusammenfassungsansicht können Sie das Profil ausführen oder bearbeiten, Sonderfälle des Musters oder der Werthäufigkeit ermitteln, einer Scorecard Spalten hinzufügen, eine Profilausführung auswählen, zwei Profilausführungen vergleichen, Profilergebnisse bzw. Ergebnisse der Datendomänenerkennung in ein Microsoft Excel-Arbeitsblatt exportieren, die Inferenzergebnisse mehrerer Spalten überprüfen, Kommentare und Tags hinzufügen bzw. löschen oder Profileigenschaften anzeigen.

Eigenschaften der Zusammenfassungsansicht

In der Zusammenfassungsansicht werden die Eigenschaften für alle Spalten und Regeln im Profil angezeigt. Die Zusammenfassungsansicht bietet eine visuelle Darstellung der Eigenschaften. Sie können auf jede Zusammenfassungseigenschaft klicken, um die Werte der Eigenschaft zu sortieren.

In der folgenden Tabelle werden die Zusammenfassungseigenschaften der Profilergebnisse beschrieben:

Eigenschaft	Beschreibung
Name	Zeigt den Namen der Spalte oder Regel im Profil an.
Null Eindeutig Nicht eindeutig %	Zeigt die eindeutigen, nicht eindeutigen und Nullwerte als Prozentsätze für eine Spalten- oder Regelausgabe an. Sie können die Werte in einem horizontalen Balkendiagramm anzeigen.
Muster	Zeigt die Mehrfachmuster in der Spalte als horizontale Balkendiagramme an. Sie können die Musterzeichen und die Anzahl ähnlicher Muster in einer Spalte als Prozentsatz anzeigen, wenn Sie den Mauszeiger über das Balkendiagramm bewegen.
Wert	Zeigt die Mindest- und Maximalwerte in der Spalten- oder Regelausgabe an.
Länge	Zeigt die minimale und maximale Länge der Werte in der Spalten- oder Regelausgabe an.
Datentyp	Zeigt den dokumentierten Datentyp der Spalte oder Regel an. Zeigt die abgeleiteten Datentypen an, wenn Sie den Mauszeiger über das Feld bewegen. Das Analyst Tool kann die folgenden Datentypen ableiten: - Zeichenfolge - Varchar - Dezimal - Ganzzahl - Datum Sie können auch den Konformitätsprozentsatz basierend auf den abgeleiteten Datentypen anzeigen. **Hinweis:** Das Analyst Tool kann den Datentyp nicht aus den Werten einer numerischen Spalte ableiten, deren Genauigkeit größer als 38 ist. Das Analyst Tool kann den Datentyp nicht aus den Werten einer Zeichenfolgenspalte ableiten, deren Genauigkeit größer als 255 ist. Wenn Sie für eine Datumsspalte ein Spaltenprofil mit einem Jahreswert vor 1800 erstellen, wird als abgeleiteter Datentyp unter Umständen eine Zeichenfolge mit fester Länge angezeigt. Ändern Sie den Standardwert für den Parameter Jahreszahlenminimum bei Bedarf in InferDateTimeConfig.xml.
Datendomäne	Zeigt die Namen der mit der Spalte verknüpften Datendomänen mit dem Konformitätsprozentsatz und der Anzahl an konformen Zeilen an.
Geschäftsbegriff	Zeigt den der Spalte zugewiesenen Geschäftsbegriff an.

Standardfilter in der Zusammenfassungsansicht

Sie können die Profilergebnisse in der Zusammenfassungsansicht basierend auf den Standardfiltern anzeigen.

In der Zusammenfassungsansicht werden standardmäßig die Profilergebnisse für alle Quellspalten, virtuellen Spalten und Regelspalten angezeigt. Im Bereich „Filtern nach“ wird die Anzahl der Spalten angezeigt, auf die die Standardfilter angewendet werden können.

In der Zusammenfassungsansicht können Sie die Profilergebnisse anzeigen, indem Sie die folgenden Standardfilteroptionen verwenden:

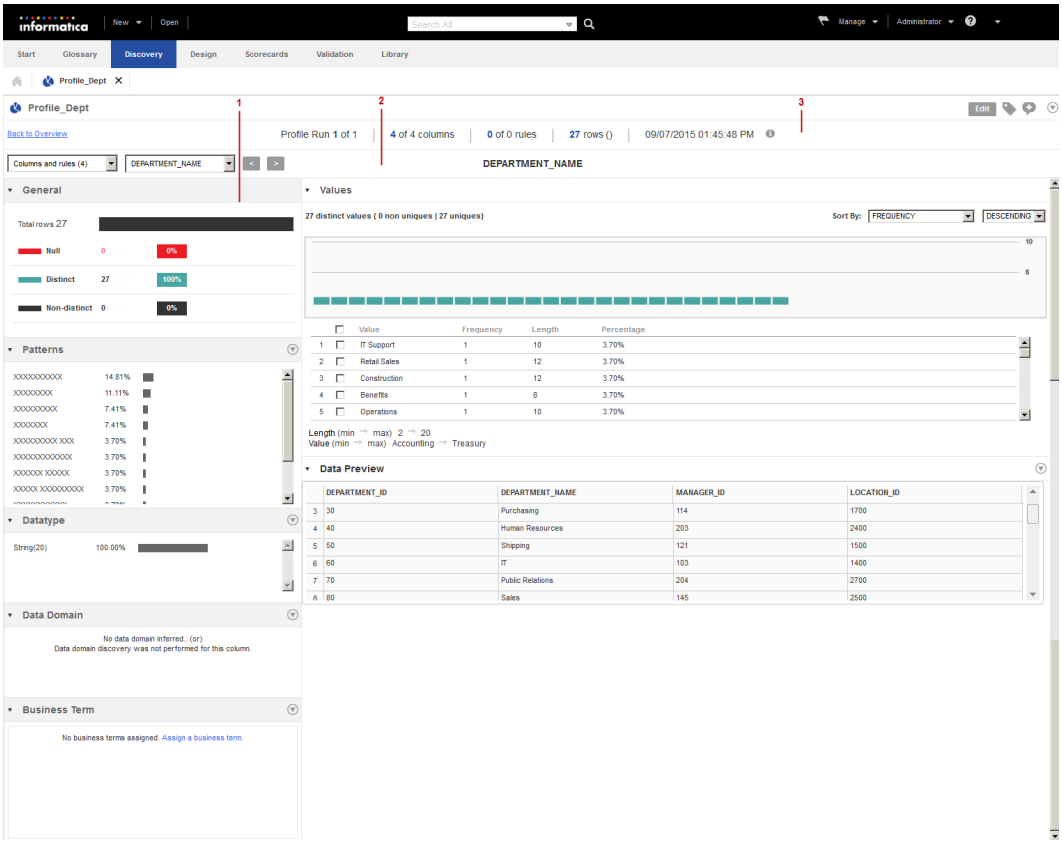
Standardfilteroption	Beschreibung
Spalten und Regeln	Zeigt die Profilergebnisse für die Quell- und Regelspalten an. Sie können die Quell- und Regelspalten zur Anzeige der Ergebnisse erweitern und reduzieren.
Spalten	Zeigt die Profilergebnisse für die Quellspalten an.
Regeln	Zeigt die Profilergebnisse für die Regelspalten an.
100% null	Zeigt die Profilergebnisse für die Spalten an, die 100% Nullwerte aufweisen.
100 % unterschiedlich	Zeigt die Profilergebnisse für die Spalten an, die 100 % unterschiedliche Werte aufweisen.
100% konstant	Zeigt die Profilergebnisse für die Spalten an, die denselben Wert für alle Datensätze aufweisen. Beispiel: Der Filter „100% konstant“ berücksichtigt die Profilergebnisse der Spalte „Land“, wenn diese nur den Wert „USA“ aufweist.
Kollidierende Datentypen	Zeigt die Profilergebnisse für Spalten an, bei denen der dokumentierte und der abgeleitete Datentyp nicht übereinstimmen. Der Filter zeigt beispielsweise die Spalte „CustomerTier“ an, weil „Ganzzahl (2)“ als dokumentierter Datentyp für die Spalte und „Zeichenfolge“ als abgeleiteter Datentyp verwendet werden.
Abgeleitete Datendomänen	Zeigt die Profilergebnisse für die Spalten an, bei denen die abgeleitete Datendomäne mit der konfigurierten Datendomäne übereinstimmt.
Sonderfall bei Mustern	Zeigt die Profilergebnisse für die Spalten an, die Sonderfälle bei Mustern aufweisen.
Sonderfall bei Werten und Häufigkeiten	Zeigt die Profilergebnisse für die Spalten an, die Sonderfälle bei Häufigkeiten aufweisen.

Detailansicht

Die Spaltenergebnisse werden in der Detailansicht angezeigt. Sie können sich die Spalteneigenschaften im Detail ansehen.

Die Detailansicht für eine Spalte wird angezeigt, nachdem Sie in der Zusammenfassungsansicht auf die Spalte geklickt haben.

Die folgende Abbildung zeigt ein Beispiel für eine grafische Ansicht von Spalteneigenschaften in der Detailansicht:



1. Bereiche. In Bereichen können Sie allgemeine Eigenschaften, Spaltenwerte, eine Datenvorschau, abgeleitete Muster, Datentypen und Datendomänen sowie Geschäftsbegriffe anzeigen.
2. Kopfzeile für Spaltendetails. Sie können die Spaltenergebnisse anzeigen, indem Sie die Spalte in der Dropdown-Liste auswählen oder die Navigationsschaltflächen verwenden.
3. Kopfzeile der Zusammenfassungsansicht. Sie können profilspezifische Informationen in der Kopfzeile der Zusammenfassungsansicht anzeigen. Sie können die Profilausführung, die Anzahl der Spalten, Regeln und Zeilen in der Profilausführung sowie das Datum und die Uhrzeit der Profilausführung anzeigen.

In der Detailansicht können Sie das Profil ausführen oder bearbeiten, die Spalte einer Scorecard hinzufügen, eine Profilausführung auswählen, zwei Profilausführungen vergleichen, die Profilergebnisse in ein Microsoft Excel-Arbeitsblatt exportieren, Werthäufigkeiten, Musterhäufigkeiten, Datentypen, Drilldown-Daten für ausgewählte Werte oder Muster in eine CSV-Datei exportieren, der Spalte Kommentare und Tags hinzufügen bzw. diese daraus löschen und Profileigenschaften anzeigen.

Über das Menü „Aktionen“ in jedem Bereich können Sie weitere Aktionen zu den Spalteneigenschaften ausführen. Sie können die Bereiche aus- oder einblenden.

Bereiche der Detailansicht

In der Detailansicht werden die folgenden Spalteneigenschaften in Bereichen angezeigt: Anzahl und Prozentsatz der eindeutigen, nicht eindeutigen und Nullwerte, Muster, abgeleitete Datentypen, abgeleitete Datendomänen, Werte, die Datenvorschau sowie verknüpfte Geschäftsbegriffe.

Wenn Sie auf die Spalte oder Regel klicken können, wird die Detailansicht für die Spalte oder Regel geöffnet.

In der folgenden Tabelle werden die Bereiche in der Detailansicht beschrieben:

Bereiche	Beschreibung
Allgemein	Zeigt die Anzahl der Zeilen mit eindeutigen, nicht eindeutigen und Nullwerten in verschiedenen Farben an. Sie können die Werte als Prozentsätze anzeigen. Sie können die Zunahme und Abnahme der allgemeinen Werte in aufeinanderfolgenden Profilausführungen als Sparkline anzeigen. Die Abweichungen in der Anzahl der eindeutigen, nicht eindeutigen und Nullwerte für die letzten fünf aufeinanderfolgenden Profilausführungen werden in einer Wortgrafik im Liniendiagramm angezeigt. Sie können die Anzahl und den Prozentsatz der Werte anzeigen, wenn Sie den Mauszeiger über die Wortgrafik für jedes Profil bewegen. Sie können der Spalte Tags und Kommentare hinzufügen.
Muster	Zeigt die Muster für die Spaltenwerte an. Die Häufigkeit, mit der die Muster in einer Spalte vorkommen, wird als horizontales Balkendiagramm und in Prozentsätzen angezeigt. Sie können einen Drilldown für ein Muster durchführen, ein Muster zu einer Referenztabelle hinzufügen oder eine Datendomäne mit dem ausgewählten Muster erstellen.
Datentyp	Zeigt die abgeleiteten Datentypen für die Spalte an. Die Häufigkeit der Datentypen in einer Spalte wird als horizontales Balkendiagramm und in Prozentsätzen angezeigt. Sie können einen Drilldown für einen Datentyp durchführen und den ausgewählten abgeleiteten Datentyp genehmigen, ablehnen oder zurücksetzen. Mit der Option **Abgelehnte einblenden** werden die abgelehnten abgeleiteten Datentypen angezeigt.
Datendomäne	Zeigt die abgeleiteten Datendomänen für die Spalte an. Sie können einen Drilldown für eine Datendomäne für konforme Zeilen, nicht konforme Zeilen oder Zeilen mit Nullwerten durchführen. Sie können den Datendomänenwert genehmigen, ablehnen oder zurücksetzen. Mit der Option **Abgelehnte einblenden** werden die abgelehnten Datendomänen angezeigt. Sie können den Datendomänenwert überprüfen.
Geschäftsbegriff	Zeigt den zugewiesenen Geschäftsbegriff für die Spalte an. Sie können einer Spalte einen Geschäftsbegriff zuweisen oder dessen Zuweisung aufheben.
Werte	Zeigt alle Werte in der Spalte zusammen mit der Häufigkeit, der Länge und dem Prozentsatz in einer grafischen Darstellung an. Sie können einen Drilldown für jeden Wert durchführen. Sie können den Wert zu einer Referenztabelle hinzufügen, eine Werthäufigkeitsregel und eine Datendomäne erstellen.
Datenvorschau	Zeigt die Drilldown-Daten für das ausgewählte Muster, den ausgewählten Datentyp, die ausgewählte Datendomäne oder den ausgewählten Wert an.

Statistiken

Sie können Statistiken, wie z. B. Werte, Muster, Datentypen, Datendomänen und Sonderfälle, für die Spalten und Regeln in einem Profil anzeigen.

Sie können Profilstatistiken in der Zusammenfassungsansicht und Spaltenstatistiken in der Zusammenfassungs- und Detailansicht anzeigen. Sie können Statistiken für die aktuelle, historische und zusammengefasste Profilausführung anzeigen. Sie können Profilergebnisse für zwei Profilausführungen vergleichen und die Statistiken für das Profil und die Spalten in der Zusammenfassungs- und Detailansicht anzeigen.

Datenvorschau

Sie können die Drilldown-Daten für das ausgewählte Muster, den ausgewählten Datentyp, die ausgewählte Datendomäne oder den ausgewählten Wert im Bereich „Datenvorschau“ anzeigen.

Sie können den Bereich „Datenvorschau“ in der Detailansicht anzeigen. Wenn Sie in der Zusammenfassungsansicht auf eine Spalte klicken, wird die Detailansicht angezeigt, wobei der Bereich „Datenvorschau“ standardmäßig ausgeblendet wird. Zum Anzeigen der Spaltendaten können Sie auf **Aktionen** > **Vorschau anzeigen** klicken.

In der folgenden Tabelle werden die Optionen des Menüs **Aktionen** im Bereich „Datenvorschau“ beschrieben:

Option	Beschreibung
Zu Filter hinzufügen	Erstellen Sie einen Drilldown-Filter zum Filtern der Drilldown-Daten, um Datenunregelmäßigkeiten in den Teilmengen der Profilergebnisse zu analysieren.
Filter speichern	Speichert den Drilldown-Filter.
Vorschau anzeigen	Zeigt die Quellzeilen an.
Daten exportieren	Exportiert die Drilldown-Ergebnisse in eine CSV- oder Microsoft Excel-Datei.

Datentypen

Die Datentypen enthalten alle abgeleiteten Datentypen für jede Spalte in den Profilergebnissen.

Sie können die Datentypen in der Zusammenfassungs- und Detailansicht anzeigen. In der Zusammenfassungsansicht können Sie den dokumentierten Datentyp und die abgeleiteten Datentypen anzeigen. Mit dem Filter **Kollidierender Datentyp** werden die Spalten angezeigt, bei denen ein Konflikt zwischen dem dokumentierten und dem abgeleiteten Datentyp besteht. In der Detailansicht können Sie die abgeleiteten Datentypen für die Spalte anzeigen. Die Häufigkeit der Datentypen in einer Spalte wird als horizontales Balkendiagramm und in Prozentsätzen angezeigt. Sie können für den ausgewählten abgeleiteten Datentyp einen Drilldown durchführen und den Datentyp genehmigen, ablehnen oder wiederherstellen. Mit der Option „Abgelehnte einblenden“ werden die abgelehnten abgeleiteten Datentypen angezeigt.

In der folgenden Tabelle werden die Eigenschaften für die Datentypen beschrieben:

Eigenschaft	Beschreibung
Datentyp	Zeigt die Liste der dokumentierten und abgeleiteten Datentypen für die Spalte im Profil an.
Häufigkeit	Zeigt an, wie oft ein Datentyp für eine Spalte angezeigt wird, ausgedrückt als Zahl.
Prozent	Zeigt den anteiligen Prozentsatz eines Datentyps pro Spalte an.

Eigenschaft	Beschreibung
Drilldown	Führt einen Drilldown für bestimmte Quellzeilen auf Basis eines Spaltendatentyps aus. **Hinweis:** Sie können keine Drilldown-Aktion durchführen, wenn Sie mehrere abgeleitete Datentypen auswählen.
Status	Gibt den Status des Datentyps an. Die Statusangaben lauten: „Abgeleitet“, „Genehmigt“ oder „Abgelehnt“. **Abgeleitet** Gibt den Datentyp der Spalte an, der vom Analyst Tool abgeleitet wurde. **Genehmigt** Gibt einen genehmigten Datentyp für die Spalte an. Wenn Sie einen Datentyp genehmigen, binden Sie den Datentyp an das Modellrepository. **Abgelehnt** Gibt einen abgelehnten Datentyp für die Spalte an.

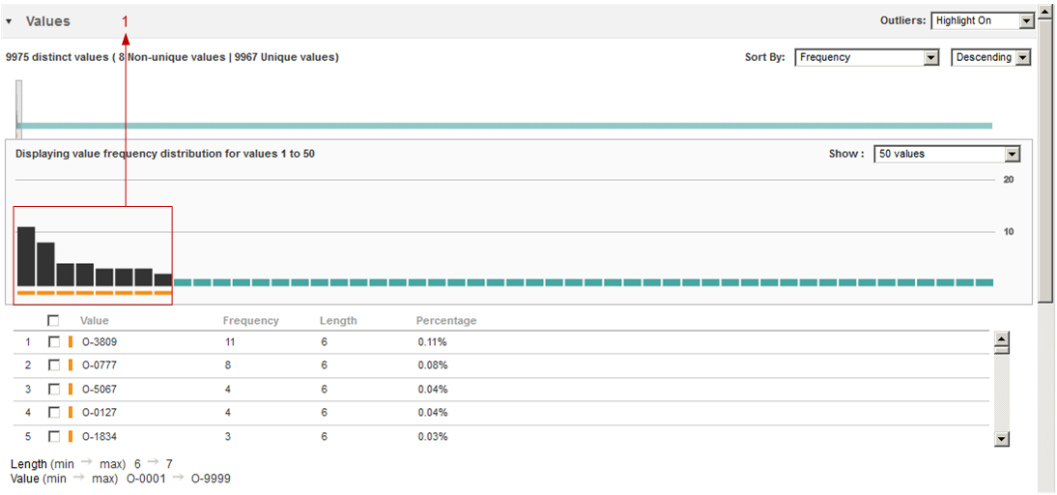
Sonderfälle

Ein Muster, ein Wert oder eine Häufigkeit für eine Spalte in den Profilergebnissen wird als Sonderfall angesehen, wenn das Muster, der Wert oder die Häufigkeit außerhalb des erwarteten Wertebereichs liegt.

Das Profiling-Plug-In im Datenintegrationsdienst führt einen Algorithmus zur Erkennung von Werten aus, die außerhalb des Bereichs der meisten Werte in der Spalte liegen. Alle Muster, Werte oder Häufigkeiten, die außerhalb des erwarteten Bereichs der meisten Werte in der Spalte liegen, werden als Sonderfall betrachtet.

Das Analyst Tool bestimmt standardmäßig keine Sonderfälle in den Profilergebnissen. In der Zusammenfassungsansicht können Sie auf den Sonderfall klicken, um die Sonderfallergebnisse anzuzeigen. Mit dem Filter „Sonderfall des Musters“ werden die Sonderfälle basierend auf den Mustern in der Spalte angezeigt. Der Filter „Sonderfall bei Werten und Häufigkeiten“ zeigt die Sonderfälle basierend auf den Werten oder Häufigkeiten in der Spalte an. Die Sonderfallerkennung findet im Hintergrund statt, so dass andere Aktionen in der Zusammenfassungsansicht ausgeführt werden können.

In der Detailansicht können Sie die Sonderfallwerte im Bereich „Werte“ anzeigen, wenn Sie die Option **Markierung Ein** in der Liste auswählen. Der Sonderfallwert wird als vertikaler Balken angezeigt, der orange unterstrichen ist. Wenn Sie nur den Sonderfallwert anzeigen möchten, wählen Sie die Option **Filter** in der Liste aus.



1. Sonderfallwerte. Ein Sonderfallwert wird als vertikaler Balken angezeigt, der orange unterstrichen ist.

Ausführen eines Sonderfalls

Führen Sie einen Sonderfall zur Erkennung von Mustern, Werten oder Häufigkeiten in einer Spalte aus, die außerhalb des erwarteten Wertebereichs liegen.

1. Klicken Sie in der Zusammenfassungsansicht auf **Aktionen** > **Sonderfall erkennen**.

   Der Wert von „Sonderfall des Musters“ und „Sonderfall der Werthäufigkeit“ im Bereich **Filtern nach** ändert sich von „N/V“ in die Anzahl der erkannten Sonderfälle.
2. Klicken Sie im Bereich **Filtern nach** auf **Sonderfall für Muster**.

   Die Spalten mit Sonderfällen bei Mustern werden in der Zusammenfassungsansicht angezeigt.
3. Klicken Sie in der Spalte **Filtern nach** auf **Sonderfall bei Werten und Häufigkeiten**.

   Die Spalten mit Sonderfällen bei Werten oder Häufigkeiten werden in der Zusammenfassungsansicht angezeigt.
4. Wählen Sie in der Detailansicht in der Dropdown-Liste für Sonderfälle die Option **Markierung Ein** aus.

   Die Sonderfälle werden im Bereich „Werte“ als vertikale Balken angezeigt, die orange unterstrichen sind.
5. Klicken Sie in der Dropdown-Liste „Sonderfälle“ auf **Filter**, um nur Sonderfallwerte anzuzeigen.

Muster

Sie können die Muster für die Spaltenwerte und die Häufigkeit, mit der die Muster vorkommen, in der Zusammenfassungs- und Detailansicht anzeigen.

In der Zusammenfassungsansicht können Sie die Mehrfachmuster in der Spalte als horizontale Balkendiagramme anzeigen. Sie können die Musterzeichen und die Anzahl ähnlicher Muster in einer Spalte als Prozentsatz anzeigen, wenn Sie den Mauszeiger über das Balkendiagramm bewegen. In der Detailansicht können Sie die Häufigkeit, mit der die Muster in einer Spalte vorkommen, als horizontales Balkendiagramm und in Prozentsätzen anzeigen. Sie können einen Drilldown durchführen, das Muster zu einer Referenztabelle hinzufügen oder eine Datendomäne mit dem ausgewählten Muster erstellen.

Im Profiling-Warehouse werden standardmäßig maximal 16.000 eindeutige Häufigkeitshöchstwerte, einschließlich NULL-Werten für Profilergebnisse, gespeichert. Wenn die Profilergebnisse mindestens einen NULL-Wert enthalten, kann das Analysetool NULL-Werte als Muster anzeigen.

Hinweis: Das Analysetool kann kein Muster für eine numerische Spalte mit einer Präzision größer 38 ableiten. Das Analysetool kann kein Muster für eine Zeichenfolgenspalte mit einer Präzision größer 255 ableiten.

Die Eigenschaften für die Spaltenmuster sind in folgender Tabelle beschrieben:

Eigenschaft	Beschreibung
Muster	Zeigt das Muster für die Spalte im Profil an.
Häufigkeit	Gibt an, wie oft ein Muster für einen Spalte angezeigt wird, ausgedrückt als Zahl.
Prozentsatz	Zeigt den anteiligen Prozentsatz eines Musters pro Spalte an.

Die folgende Tabelle beschreibt die Musterzeichen und was diese darstellen:

Zeichen	Beschreibung
9	Stellt ein beliebiges numerisches Zeichen dar. Informatica Analyst zeigt separat bis zu drei Zeichen im Format „9“ an. Das Tool zeigt mehr als drei Zeichen als Wert in Klammern an. Das Format „9(8)“ stellt beispielsweise einen numerischen Wert mit 8 Ziffern dar.
X	Stellt ein beliebiges alphabetisches Zeichen dar. Informatica Analyst zeigt separat bis zu drei Zeichen im Format „X“ an. Das Tool zeigt mehr als drei Zeichen als Wert in Klammern an. Das Format „X(6)“ kann beispielsweise den Wert·„Boston“ darstellen. **Hinweis:** Das Musterzeichen X unterscheidet nicht zwischen Groß-/Kleinschreibung und kann für Groß- oder Kleinbuchstaben aus der Datenquelle stehen.
p	Stellt „(“ dar, die öffnende Klammer.
q	Stellt „)“ dar, die schließende Klammer.

Hinweis: Spaltenmuster können auch Sonderzeichen enthalten. Beispiel: ~, [,], =, -, ?, =, {, *, -, >, < und $.

Werte

Sie können Werte für Spalten und die Häufigkeit anzeigen, mit der die Werte in der Spalte angezeigt werden.

Zeigen Sie Mindest- und Maximalwerte in einer Spalte in der Zusammenfassungsansicht an. In der Detailansicht können Sie die Werteigenschaften für eine Spalte anzeigen.

Werte in der Zusammenfassungsansicht

Sie können die Mindest- und Maximalwerte für alle Spalten und Regeln für die aktuelle, historische und zusammengefasste Profilausführung in der Zusammenfassungsansicht anzeigen.

Beispiel

Die Datenbank eines Einzelhandelsgeschäfts weist eine Spalte mit der Bezeichnung „Mitarbeiter-ID“ in der Mitarbeitertabelle auf, die mit Mitarbeiter-IDs von 100 bis 250 und Namen, wie z. B. Bob und Robert, befüllt ist. Wenn Sie ein Spaltenprofil für die Mitarbeitertabelle ausführen, zeigt die Spalte „Wert“ für die Mitarbeiter-ID in der Zusammenfassungsansicht 100 --> Robert an

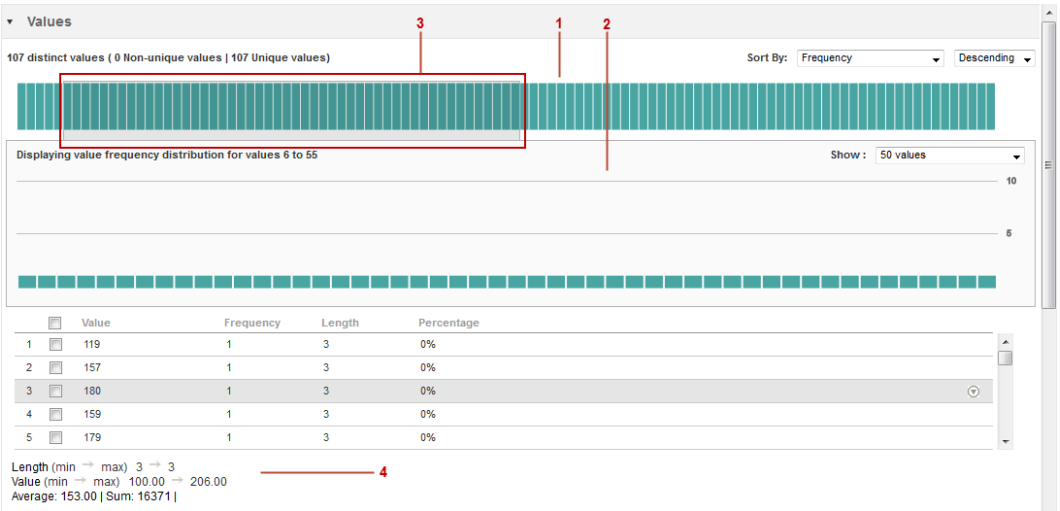
Werte in der Detailansicht

Zu den Spaltenwerten in der Detailansicht gehören Werte für eine Spalte und die Häufigkeit, mit der die Werte in der Spalte angezeigt werden.

Im Bereich **Werte** werden die Spaltenwerte als grafische Darstellung angezeigt. Sie können die Häufigkeit, Länge und den Prozentsatz jedes Werts anzeigen. Sie können die Werte basierend auf Wert oder Häufigkeit sortieren. Sie können einen Drilldown für die Daten durchführen, die Werte zu einer Referenztabelle hinzufügen und eine Werthäufigkeitsregel oder Datendomäne erstellen. Sie können die Nullwerte als roten vertikalen Balken, die Häufigkeit der Werte als schwarzen vertikalen Balken und die Sonderfallwerte als vertikale Balken mit orangefarbener Markierung anzeigen. Sie können die Sonderfälle markieren und deaktivieren oder die Ergebnisse filtern, um nur Sonderfallwerte in der Spalte anzuzeigen.
Der Bereich „Werte“ enthält das grafische Layout und Wertauswahlen.

Das grafische Layout ist in zwei Bereiche unterteilt.

Die folgende Abbildung zeigt den Bereich „Werte“ in der Detailansicht:



1. Oberer Bereich. Sie können die Werte als vertikales Balkendiagramm anzeigen. Sie können die Werte nach Häufigkeit und Wert sortieren. Sie können die Werte in auf- oder absteigender Reihenfolge sortieren. Sie können die Sonderfallwerte als vertikale Balken mit orangefarbenen Markierungen anzeigen.
2. Unterer Bereich. Sie können die Werte im Schieberegler im unteren Bereich anzeigen, in dem jeder Wert als vertikaler Balken dargestellt wird. Sie können einen Drilldown für den Wert durchführen, den Wert zu einer Referenztabelle hinzufügen und eine Werthäufigkeitsregel oder eine Datendomäne für den Wert erstellen. Sie können gleichzeitig 50, 75 oder 100 Werte anzeigen.
3. Schieberegler. Sie können den Schieberegler entlang der Werte im oberen Bereich ziehen. Im unteren Bereich werden die Werte im Schieberegler angezeigt.
4. Werteigenschaften. Im Abschnitt „Werteigenschaften“ werden die Werte und Eigenschaften angezeigt.

In der folgenden Tabelle werden die Bereiche im grafischen Layout beschrieben:

Bereich	Beschreibung
Oberer Bereich	Zeigt alle Werte als vertikales Balkendiagramm an. Sie können im oberen Bereich höchstens 16.000 Werte anzeigen. Über den Schieberegler können Sie eine Gruppe von Werten anzeigen.
Unterer Bereich	Zeigt die Werte für den Stapel an, den Sie im oberen Bereich ausgewählt haben. Das Analyst Tool zeigt standardmäßig 50 Werte an. Sie können 75 oder 100 Werte gleichzeitig anzeigen.

In der folgenden Tabelle werden die Eigenschaften für die Spaltenwerte im Abschnitt „Wert“ beschrieben:

Eigenschaft	Beschreibung
Wert	Zeigt eine Werteliste für den Stapel an, den Sie im oberen Bereich ausgewählt haben. **Hinweis:** Das Analyst Tool schließt die Datentypen „CLOB“, „BLOB“, „Roh“ und „Binär“ in Spaltenwerten aus.
Häufigkeit	Zeigt an, wie oft ein Wert in der Spalte angezeigt wird, ausgedrückt als Zahl.
Länge	Zeigt die Länge des Spaltenwerts an.
Prozentsatz	Zeigt den anteiligen Prozentsatz für einen Wert pro Spalte an.

In der folgenden Tabelle werden die Statistiken für die ausgewählte Spalte angezeigt:

Statistiken	Beschreibung
Länge (min - max)	Zeigt die Länge des kürzesten und des längsten Werts für die Spalte an.
Wert (min - max)	Zeigt die Minimal- und Maximalwerte in der Spalte an.
Durchschnitt	Zeigt den Durchschnitt der Werte für die Spalte an.
Summe	Zeigt die Summe aller Werte in der Spalte an.

Werte in der Detailansicht für den Profilergebnisvergleich

Im Bereich „Werte“ in der Detailansicht für den Profilergebnisvergleich werden Werteigenschaften angezeigt, wie z. B. Anzahl der unterschiedlichen Werte, Mindestwert, Maximalwert, maximale und minimale Länge, Durchschnitt, Standardabweichung und Summe der Werte.

In der Detailansicht einer Spalte für den Profilergebnisvergleich werden Werteigenschaften sowie der Wert und die Häufigkeit des Werts als horizontales Balkendiagramm angezeigt.

In der folgenden Tabelle werden die Eigenschaften für die Spaltenwerte in der Detailansicht beim Vergleich der Ergebnisse zweier Profilausführungen beschrieben.:

Eigenschaft	Beschreibung
Anzahl der unterschiedlichen Werte	Zeigt die Anzahl der unterschiedlichen Werte in der Spalte an.
Mindestwert	Zeigt den Mindestwert in der Spalte an.
Maximalwert	Zeigt den Maximalwert in der Spalte an.
Länge (Min - Max)	Zeigt die Länge des kürzesten und des längsten Werts für die Spalte an.
Durchschnitt	Zeigt den Durchschnitt der Werte für die Spalte an.
Standardabweichung	Zeigt die Standardabweichung bzw. Variabilität zwischen den Spaltenwerten für alle Spaltenwerte an.
Summe	Zeigt die Summe aller Werte in der Spalte an.

Typen von Profilausführungen

Sie können die Profilergebnisse für die aktuelle, historische und zusammengefasste Profilausführung anzeigen. Sie können die Ergebnisse der Profilausführung in der Zusammenfassungsansicht anzeigen.

Aktuelle Profilausführung

Zeigen Sie Profilergebnisse für die aktuelle Profilausführung im Profil in der Zusammenfassungsansicht an.

Sie können die Profilergebnisse für die letzte Profilausführung in der Zusammenfassungsansicht anzeigen, wenn Sie:

- ein Profil erstellen, speichern und ausführen.
- ein Profil öffnen, das Sie zuvor im Arbeitsbereich **Bibliothek** ausgeführt haben.
- in der Zusammenfassungs- bzw. Detailansicht für die konsolidierte Profilausführung auf den Link **Zurück zur letzten Profilausführung** klicken.
- in der Zusammenfassungs- bzw. Detailansicht für eine historische Profilausführung auf den Link **Zurück zur letzten Profilausführung** klicken.
- die letzte Profilausführung im Dialogfeld **Profilausführung auswählen** auswählen und auf **OK** klicken.

Historische Profilausführung

Zeigen Sie die Profilergebnisse für eine vorherige Profilausführung in der Zusammenfassungsansicht an.

Im Profiling-Warehouse werden die Profilergebnisse aller Profilausführungen eines Profils gespeichert. Sie können die Ergebnisse aus einer früheren Version der Profilausführung anzeigen, indem Sie die Profilausführung im Dialogfeld „Profilausführung auswählen“ auswählen.

Zusammengefasste Profilausführung

Zeigen Sie die aktuellen Profilergebnisse für alle Spalten im Profil in der Zusammenfassungsansicht an.

In der zusammengefassten Profilausführung können Sie die aktuellen Ergebnisse für alle Spalten im Profil anzeigen. Wenn Sie die zusammengefasste Profilausführung im Dialogfeld **Profilausführung auswählen** auswählen, ruft das Profiling-Warehouse die aktuellen Spaltenergebnisse aus allen Profilausführungen des Profils ab. Sie können die Ergebnisse in der Zusammenfassungsansicht anzeigen. In der Kopfzeile der Zusammenfassungsansicht wird eine inkrementelle Profilausführung angezeigt.

Beispiel

Als Datenanalyst können Sie die aktuellen Ergebnisse für alle Spalten im Profil anzeigen. Sie können beispielsweise die Spalten 1, 2 und 3 zur Durchführung von Profilausführung A und die Spalten 3, 4 und 5 für Profilausführung B festlegen. Zur Anzeige der aktuellen Ergebnisse für alle Spalten können Sie die zusammengefasste Profilausführung im Dialogfeld „Profilausführung auswählen“ auswählen. In der Zusammenfassungsansicht werden Ergebnisse für die Spalten 1 und 2 aus Ausführung A sowie Ergebnisse für die Spalten 3, 4 und 5 aus Ausführung B angezeigt.

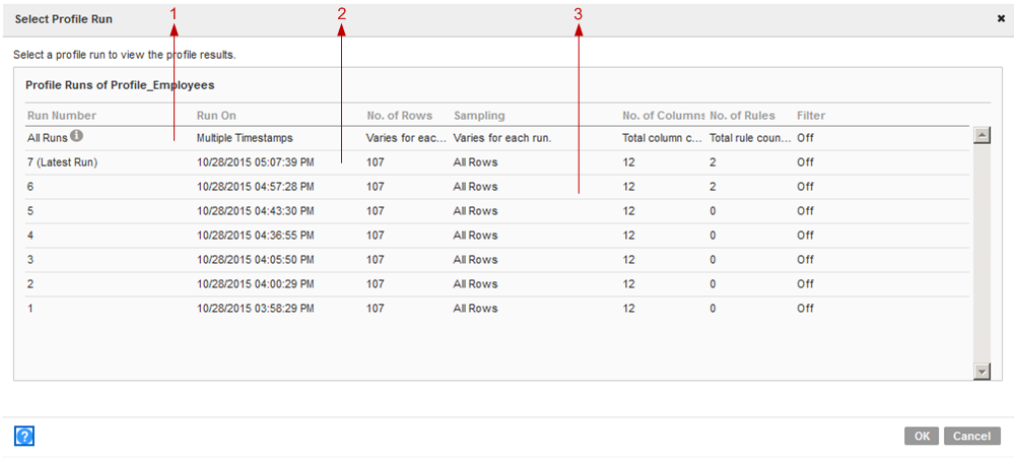
Auswählen einer Profilausführung

Sie können eine historische, die aktuelle oder eine zusammengefasste Profilausführung zum Anzeigen der Profilergebnisse auswählen. Sie können die Profilergebnisse in der Zusammenfassungsansicht und die Spaltenergebnisse in der Detailansicht anzeigen.

1. Wählen Sie im Arbeitsbereich **Bibliothek** das Projekt bzw. den Ordner aus, in dem das Profil enthalten ist, oder wählen Sie im Bereich **Objekte** das Profil aus.
2. Klicken Sie auf **Aktionen** > **Öffnen**, um das Profil zu öffnen.

   Die Zusammenfassungsansicht wird im Arbeitsbereich **Erkennung** angezeigt.
3. Klicken Sie in der Zusammenfassungsansicht auf **Aktionen** > **Profilausführung auswählen**.

   Das Dialogfeld **Profilausführung auswählen** wird angezeigt.
   Die folgende Abbildung zeigt das Dialogfeld **Profilausführung auswählen**.



1. Zusammengefasste Profilausführung. Wenn Sie diese Profilausführung auswählen, können Sie die aktuellen Profilergebnisse für jede Spalte in der Zusammenfassungsansicht anzeigen.
2. Aktuelle Profilausführung. Wenn Sie diese Profilausführung auswählen, können Sie die aktuellen Profilergebnisse für das Profil in der Zusammenfassungsansicht anzeigen.

3. Historische Profilausführung. Wenn Sie diese Profilausführung auswählen, können Sie die historischen Profilergebnisse für eine vorherige Profilausführung in der Zusammenfassungsansicht anzeigen.

4. Wählen Sie im Dialogfeld **Profilausführung auswählen** eine der Profilausführungen aus, um deren Profilergebnisse anzuzeigen:
   - Wenn Sie die Profilergebnisse für die letzte Profilausführung anzeigen möchten, wählen Sie die letzte Profilausführung aus und klicken Sie auf **OK**.
   - Wenn Sie die Profilergebnisse für eine historische Profilausführung anzeigen möchten, wählen Sie nicht die letzte, sondern eine andere Profilausführung aus und klicken Sie auf **OK**.
   - Wenn Sie die Profilergebnisse für ein konsolidiertes Profil anzeigen möchten, wählen Sie **Alle Ausführungen** aus und klicken Sie auf **OK**. Die aktuellen Profilergebnisse für alle Spalten werden in der Zusammenfassungsansicht angezeigt.

   Das Analyst Tool führt eine Profilausführung durch und zeigt die Profilergebnisse in der Zusammenfassungsansicht an.

5. Klicken Sie in der Zusammenfassungsansicht auf eine Spalte, um die Spaltenergebnisse anzuzeigen.

   Die Detailansicht wird angezeigt.

Vergleichen mehrerer Profilergebnisse - Übersicht

Sie können Profilergebnisse für zwei Profilausführungen vergleichen. Sie können die Ergebnisse des Vergleichs in der Zusammenfassungsansicht und Spaltenergebnisse in der Detailansicht anzeigen.

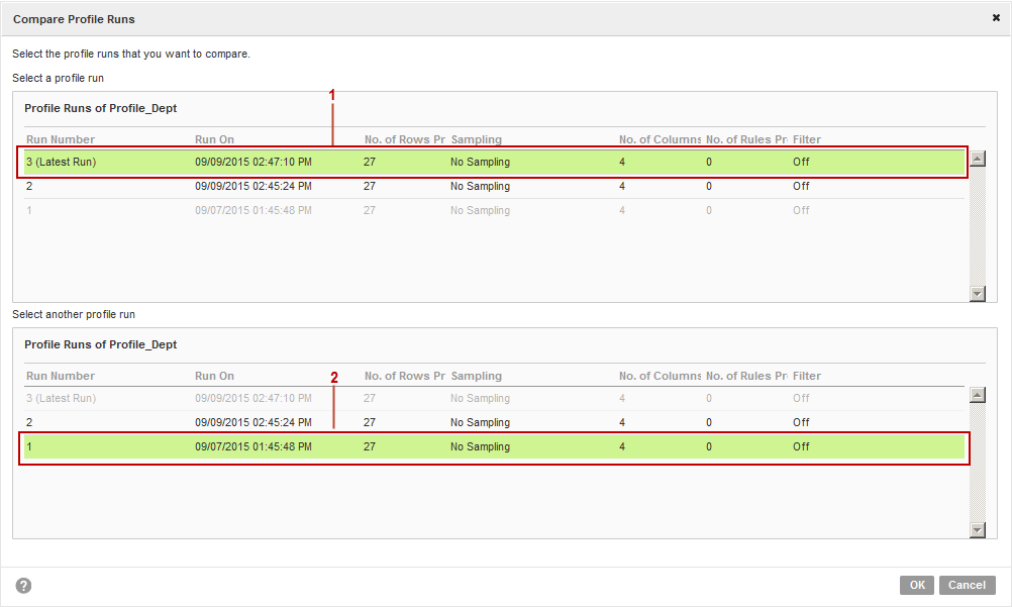
In der Zusammenfassungsansicht können Sie die Ergebnisse des Vergleichs für alle Spalten in beiden Profilausführungen anzeigen.

Vergleichen mehrerer Profilergebnisse

Wenn Sie zwei Profilausführungen vergleichen, können Sie den Profilergebnisvergleich in der Zusammenfassungsansicht anzeigen.

1. Klicken Sie in der Zusammenfassungsansicht auf **Aktionen** > **Profilausführung vergleichen**.

   Die folgende Abbildung zeigt das Dialogfeld **Profilausführungen vergleichen**.



1. Ausführung A. Wählen Sie eine Profilausführung als Ausführung A aus.
2. Ausführung B. Wählen Sie eine Profilausführung als Ausführung B aus.

Das Dialogfeld **Profilausführungen vergleichen** wird angezeigt.

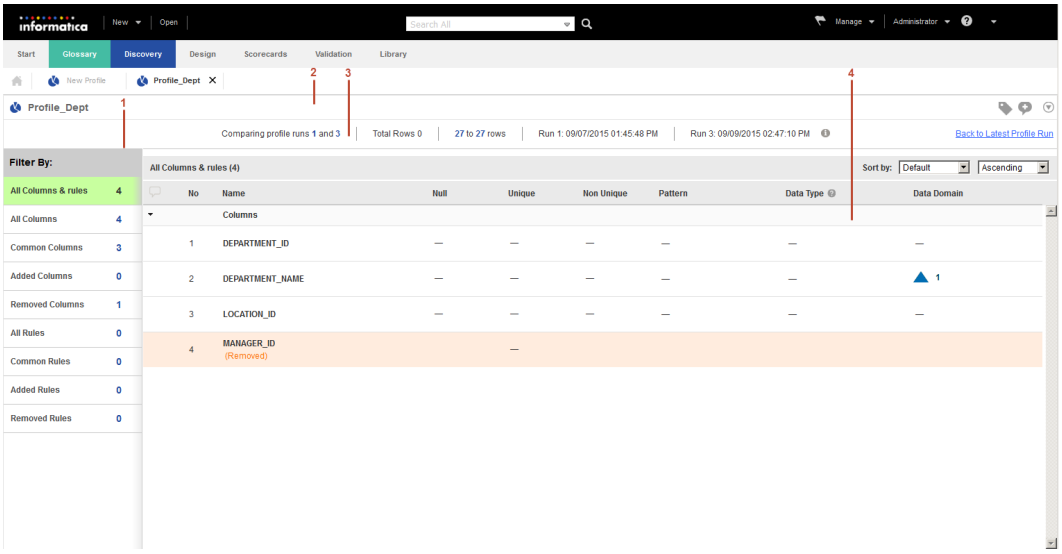
2. Wählen Sie ein Profil aus dem Bereich **Ausführung A** und ein anderes Profil aus dem Bereich **Ausführung B** aus.
3. Klicken Sie auf **OK**.

   In der Zusammenfassungsansicht wird eine zusammengefasste Anzeige der Profilergebnisse angezeigt.

Zusammenfassungsansicht für den Profilergebnisvergleich

Wenn Sie zwei Profilausführungen vergleichen, können Sie die Ergebnisse in einem Gitterformat in der Zusammenfassungsansicht anzeigen. Sie können die Standardfilter in der Zusammenfassungsansicht verwenden, um bestimmte Statistiken anzuzeigen.

Die folgende Abbildung zeigt die Ergebnisse eines Profilvergleichs für zwei Profilausführungen in der Zusammenfassungsansicht:



1. Standardfilter. Sie können die Ergebnisse des Profilvergleichs in der Zusammenfassungsansicht basierend auf den Standardfiltern anzeigen.
2. Kopfzeile des Profils. Sie können den Profilnamen in der Kopfzeile anzeigen.
3. Kopfzeile der Zusammenfassungsansicht. Sie können profilspezifische Informationen in der Kopfzeile der Zusammenfassungsansicht anzeigen. Sie können die Profilausführungen, die verglichen werden, eine höhere oder geringere Zeilenanzahl zwischen den Profilausführungen, die Anzahl der Zeilen im Profil sowie das Datum und die Uhrzeit der Profilausführungen anzeigen.
4. Zusammenfassungsansicht. Sie können den Vergleich zwischen den Spalten in beiden Profilausführungen anzeigen.

Eigenschaften der Zusammenfassungsansicht für den Profilergebnisvergleich

Zu den Eigenschaften der Zusammenfassungsansicht für den Profilergebnisvergleich gehören die Anzahl und der Prozentsatz der eindeutigen, nicht eindeutigen und Nullwerte, Muster, abgeleitete Datentypen und Datendomänen sowie verknüpfte Geschäftsbegriffe. Die Zusammenfassungsansicht bietet eine visuelle Darstellung der Eigenschaften. Sie können auf jede Zusammenfassungseigenschaft klicken, um die Werte der Eigenschaft zu sortieren.

In der Zusammenfassungsansicht weist der Datenintegrationsdienst allen Spalten und Regeln eine Zahl in absteigender Reihenfolge zu

Hinweis: Ein nach oben weisender Pfeil mit einer Zahl zeigt von Profilausführung zu Profilausführung die Vergrößerung der Werte einer Eigenschaft an. Ein nach unten weisender Pfeil mit einer Zahl zeigt von Profilausführung zu Profilausführung die Verkleinerung der Werte einer Eigenschaft an.

In der folgenden Tabelle werden die Zusammenfassungseigenschaften für den Profilergebnisvergleich beschrieben:

Eigenschaft	Beschreibung
Nummer	Zeigt die Nummer der Spalte oder Regel an.
Name	Zeigt den Namen der Spalte oder Regel im Profil an.
Null	Zeigt die Zunahme oder Abnahme der Nullwerte an.

Eigenschaft	Beschreibung
Eindeutig	Zeigt die Zunahme bzw. Abnahme der eindeutigen Werte an.
Nicht eindeutig	Zeigt die Zunahme bzw. Abnahme der nicht eindeutigen Werte an.
Muster	Zeigt die Änderung der Muster zwischen den Profilausführungen an.
Datentyp	Zeigt die Änderung zwischen den abgeleiteten Datentypen für die Spalte oder Regel in den beiden Profilausführungen an.
Datendomäne	Zeigt die Änderung zwischen den abgeleiteten Datendomänen an, die mit der Spalte oder Regel in den beiden Profilausführungen verknüpft sind.

Standardfilter für den Profilergebnisvergleich in der Zusammenfassungsansicht

Sie können die Profilergebnisse basierend auf den Standardfiltern in der Zusammenfassungsansicht anzeigen.

In der Zusammenfassungsansicht können Sie die Profilergebnisse anzeigen, indem Sie die folgenden Standardfilteroptionen verwenden:

Standardfilteroption	Beschreibung
Alle Spalten und Regeln	Zeigt die Profilergebnisse für die Quellspalten, virtuellen Spalten und Regelspalten an. Sie können die Quell- und Regelspalten zur Anzeige der Ergebnisse erweitern und reduzieren.
Alle Spalten	Zeigt die Profilergebnisse für die Quell- und virtuellen Spalten an.
Allgemeine Spalten	Zeigt die Spalten an, die in beiden Profilausführungsergebnissen zur Verfügung stehen.
Hinzugefügte Spalten	Zeigt die Spalten an, die in der aktuellen Profilausführung zur Verfügung stehen. Wenn Sie beispielsweise Ausführung 5 mit Ausführung 3 vergleichen, werden unter „Hinzugefügte Spalten“ die Spalten angezeigt, die in Ausführung 5, nicht aber in Ausführung 3 zur Verfügung stehen.
Entfernte Spalten	Zeigt die Spalten an, die in der historischen Profilausführung zur Verfügung stehen. Wenn Sie beispielsweise Ausführung 5 mit Ausführung 3 vergleichen, werden unter „Entfernte Spalten“ die Spalten angezeigt, die in Ausführung 3, nicht aber in Ausführung 5 zur Verfügung stehen.
Alle Regeln	Zeigt die Profilergebnisse für alle Regelspalten an.
Hinzugefügte Regeln	Zeigt die Regeln an, die in der aktuellen Profilausführung zur Verfügung stehen. Wenn Sie beispielsweise Ausführung 5 mit Ausführung 3 vergleichen, werden unter „Hinzugefügte Regeln“ die Regeln angezeigt, die in Ausführung 5, nicht aber in Ausführung 3 zur Verfügung stehen.
Entfernte Regeln	Zeigt die Regeln an, die in der historischen Profilausführung zur Verfügung stehen. Wenn Sie beispielsweise Ausführung 5 mit Ausführung 3 vergleichen, werden unter „Entfernte Regeln“ die Regeln angezeigt, die in Ausführung 3, nicht aber in Ausführung 5 zur Verfügung stehen.

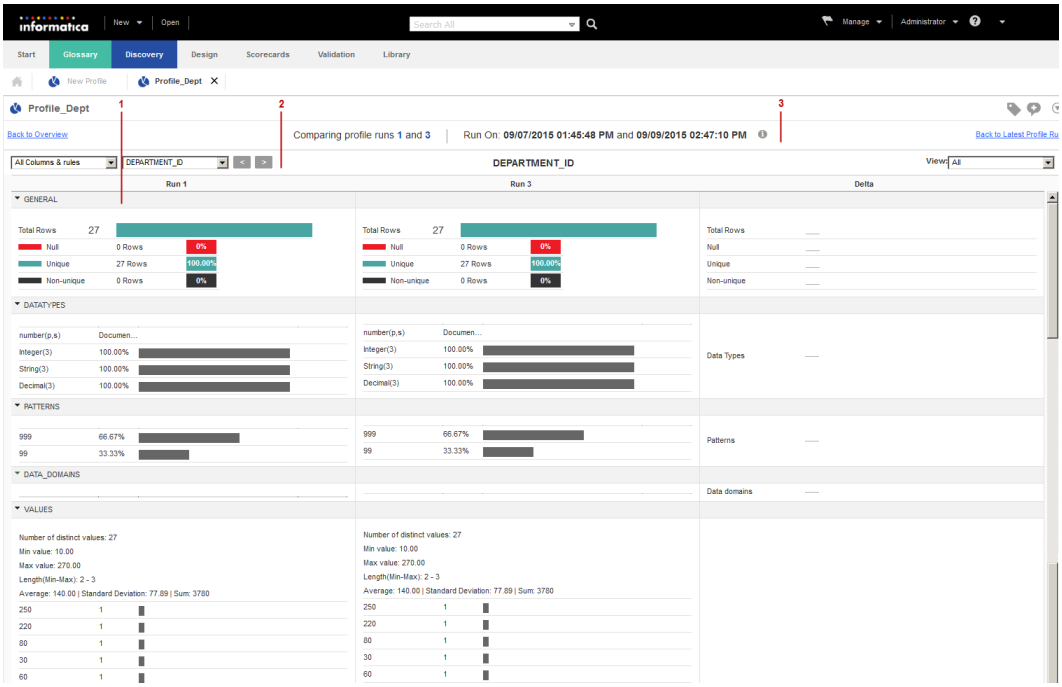
In der Zusammenfassungsansicht werden standardmäßig die Profilergebnisse für alle Quellspalten, virtuellen Spalten und Regelspalten angezeigt.

Detailansicht des Profilergebnisvergleichs

Spaltenergebnisse werden im Gitterformat in der Detailansicht angezeigt. Zu den Spaltendetails gehören allgemeine Informationen, wie z. B. eindeutige, nicht eindeutige und Nullwerte, Muster, Datentypen, Datendomänen, Geschäftsbegriffe, Werte und die Datenvorschau.

Die Detailansicht für eine Spalte wird angezeigt, wenn Sie auf den Spaltennamen klicken. Sie können die Spaltenergebnisse in Ausführung A und Ausführung B in getrennten Spalten anzeigen. Der Datenvergleich befindet sich in der Deltaspalte.

Die folgende Abbildung zeigt den Profilergebnisvergleich für eine Spalte in der Detailansicht:



1. Bereiche. Sie können die Profilergebnisse und Statistiken für die Spalte in den beiden Profilausführungen sowie die Deltainformationen für die Spalte in den beiden Profilausführungen in Bereichen anzeigen.
2. Kopfzeile des Profils. Sie können Spaltenergebnisse anzeigen, indem Sie die Spalte in der Dropdown-Liste auswählen oder die Navigationsschaltflächen verwenden. Mithilfe der Optionen in der Dropdown-Liste „Anzeigen“ können Sie den Spaltennamen und bestimmte Ergebnisse anzeigen.
3. Kopfzeile der Zusammenfassungsansicht. Sie können profilspezifische Informationen in der Kopfzeile der Zusammenfassungsansicht anzeigen. Sie können die Profilausführungen, die verglichen werden, sowie das Datum und die Uhrzeit der Profilausführungen anzeigen.

Bereiche der Detailansicht für den Profilergebnisvergleich

Die Detailansicht zeigt die Profilergebnisse und die Ergebnisse des Vergleichs für eine Spalte in den beiden Profilausführungen detailliert an.

In der Detailansicht werden die Spaltenergebnisse für Ausführung A und Ausführung B angezeigt. Der Datenvergleich befindet sich in der Deltaspalte. Zur Anzeige anderer Spaltenergebnisse können Sie einen Filter aus der Dropdown-Liste „Filter“ oder die Spalte aus der Dropdown-Liste „Spalte“ auswählen.

Spaltenprofil-Drilldown

Verwenden Sie die Drilldown-Optionen in einem Spaltenprofil, um basierend auf einem Spaltenwert einen Drilldown für bestimmte Zeilen in der Datenquelle durchzuführen. Sie können die aktuellen Daten in einer Datenquelle für den Drilldown lesen. Sie können aber auch Profildaten lesen, die im Profiling-Warehouse zwischengespeichert wurden. Wenn Sie mit zwischengespeicherten Profildaten einen Drilldown für eine bestimmte Zeile durchführen, erstellt das Analyst Tool einen Drilldown-Filter für den passenden Spaltenwert. Nach dem Drilldown können Sie den Drilldown-Filter bearbeiten, erneut aufrufen, zurücksetzen und speichern.

Sie können Spalten für den Drilldown auswählen, auch wenn Sie diese Spalten nicht für das Profiling ausgewählt haben. Sie können die aktuellen Daten in einer Datenquelle für den Drilldown lesen. Sie können aber auch Profildaten lesen, die im Profiling-Warehouse zwischengespeichert wurden. Nachdem Sie einen Drilldown für einen Spaltenwert durchgeführt haben, können Sie Drilldown-Daten für die ausgewählten Werte oder Muster in eine CSV-Datei an einem ausgewählten Speicherort exportieren. Obwohl Informatica Analyst die ersten 200 Werte für die Drilldown-Daten anzeigt, exportiert das Tool alle Werte in die CSV-Datei.

Drilldown an Zeilendaten

Nachdem Sie ein Profil ausgeführt haben, können Sie einen Drilldown für bestimmte Zeilen durchführen, die dem Spaltenwert, Datentyp oder Muster entsprechen.

1. Führen Sie ein Profil aus.

   Die Profilergebnisse werden in der Zusammenfassungsansicht angezeigt.
2. Klicken Sie in der Zusammenfassungsansicht auf einen Spaltennamen.

   Die Spaltenergebnisse werden in der Detailansicht angezeigt.
3. Klicken Sie in der Detailansicht im Bereich **Werte** mit der rechten Maustaste auf einen Wert und wählen Sie **Drilldown** aus.

   Die Drilldown-Daten werden im Bereich **Datenvorschau** angezeigt.

Filter auf Drilldown-Daten anwenden

Sie können die Drilldown-Daten iterativ filtern, um Unregelmäßigkeiten in den Daten in den Teilmengen der Profilergebnisse analysieren zu können.

1. Wählen Sie einen Spaltenwert auf der Registerkarte **Werte** aus.
2. Klicken Sie mit der rechten Maustaste und wählen Sie **Drilldown** aus.

   Die Drilldown-Ergebnisse werden im Bereich **Datenvorschau** angezeigt.
3. Wenn Sie eine Filterbedingung hinzufügen möchten, klicken Sie im Bereich **Datenvorschau** mit der rechten Maustaste auf einen Spaltenwert und wählen **Zu Filter hinzufügen** aus.

   Das Dialogfeld **Drilldown-Filter** wird mit der Filterbedingung geöffnet.
4. Fügen Sie die erforderlichen Filterbedingungen hinzu und klicken Sie auf **OK**.

   Drilldown-Filter können auf abgeleitete Datentypen nicht angewendet werden.
5. Klicken Sie auf **Aktionen** > **Filter speichern**, um den Filter zu speichern.
6. Klicken Sie auf **Aktionen** > **Aktualisieren**, um die Drilldown-Filter zu löschen.
7. Klicken Sie auf **Aktionen** > **Daten exportieren**, um die Drilldown-Daten in ein Microsoft Excel-Arbeitsblatt zu exportieren.

Wiederherstellung im Analyst-Tool

Wiederherstellung ist der Prozess des Validierens und Verwaltens erkannter Metadaten einer Datenquelle, um die Metadaten für ihre Verwendung und das Reporting anzupassen. Wenn Sie Metadaten im Analyst Tool wiederherstellen, können Sie die abgeleiteten Datentypen oder Datendomänen in Profilergebnissen genehmigen, ablehnen und zurücksetzen.

Sie können einen Datentyp und eine Datendomäne für eine Spalte genehmigen. Sie können die abgelehnten Datentypen und Datendomänen für eine Spalte ausblenden. Nachdem Sie einen abgeleiteten Datentyp oder eine abgeleitete Datendomäne genehmigt oder abgelehnt haben, können Sie den Datentyp oder die Datendomäne zurücksetzen, um den Ableitungsstatus wiederherzustellen.

Genehmigen von Datentypen und Datendomänen

Die Profilergebnisse umfassen die abgeleiteten Datentypen und Datendomänen für alle Spalten in der Datenquelle. Sie können im Analyst Tool einen einzelnen Datentyp und eine einzelne Datendomäne für jede Spalte auswählen und genehmigen.

1. Führen Sie ein Profil aus.

   Die Profilergebnisse werden in der Zusammenfassungsansicht angezeigt.
2. Klicken Sie in der Zusammenfassungsansicht auf einen Spaltennamen.

   Die Spaltenergebnisse werden in der Detailansicht angezeigt.
3. Wählen Sie in der Detailansicht einen Datentyp im Bereich **Datentyp** oder eine Datendomäne im Bereich **Datendomäne** aus.
4. Klicken Sie auf **Aktionen** > **Genehmigen**.
5. Wählen Sie zum Wiederherstellen des abgeleiteten Status des Datentyps oder der Datendomäne den Datentyp oder die Datendomäne aus und klicken Sie auf **Aktionen** > **Zurücksetzen**.

Ablehnen von Datentypen und Datendomänen

In der Detailansicht können Sie einen Datentyp oder eine Datendomäne ablehnen. Sie können die abgelehnten Datentypen und Datendomänen anzeigen oder ausblenden.

1. Führen Sie ein Profil aus.

   Die Profilergebnisse werden in der Zusammenfassungsansicht angezeigt.
2. Klicken Sie in der Zusammenfassungsansicht auf einen Spaltennamen.

   Die Spaltenergebnisse werden in der Detailansicht angezeigt.
3. Wählen Sie in der Detailansicht einen Datentyp im Bereich **Datentyp** oder eine Datendomäne im Bereich **Datendomäne** aus.
4. Klicken Sie auf **Aktionen** > **Ablehnen**.

   Das Analyst Tool entfernt den abgelehnten Datentyp aus der Liste der Datentypen.
5. Klicken Sie zum Anzeigen der abgelehnten Datensätze auf **Aktionen** > **Abgelehnte einblenden**.

Exportdateien des Spaltenprofils in Informatica Analyst

Sie können Spaltenprofilergebnisse in eine CSV-Datei oder einer Microsoft Excel-Datei exportieren. Der Export basiert darauf, ob Sie einen Teil der Profilergebnisse oder die Zusammenfassung aller Ergebnisse auswählen.

Sie können Werthäufigkeiten, Musterhäufigkeiten, Datentypen oder Drilldown-Daten für ausgewählte Werte und Muster in eine CSV-Datei exportieren. Sie können die Zusammenfassung der Profiling-Ergebnisse für alle Spalten in eine Microsoft Excel-Datei exportieren. Verwenden Sie die Berechtigung **Ergebnis-Drilldown und Export** des Datenintegrationsdiensts, um nach Benutzer oder Gruppe festzulegen, wer Profilergebnisse exportieren darf.

Profilieren von Exportergebnissen in eine CSV-Datei

Sie können Werthäufigkeiten, Musterhäufigkeiten, Datentypen oder Drilldown-Daten exportieren, um die Daten in einer Datei anzuzeigen. Das Analysetool speichert die Informationen in einer CSV-Datei.

Beim Exportieren abgeleiteter Spaltenmuster exportiert das Analysetool ein anderes Format des Spaltenmusters. Wenn Sie zum Beispiel das abgeleitete Spaltenmuster X(5) exportieren, wird das Spaltenmuster im folgenden Format vom Analysetool in der CSV-Datei angezeigt: XXXXX

Profilieren von Exportergebnissen in Microsoft Excel

Beim Exportieren der Zusammenfassung aller Profilergebnisse speichert das Analysetool die Informationen in mehreren Arbeitsblättern in einer Microsoft Excel-Datei. Das Analyst Tool speichert die Datei im XLSX-Format.

Die folgende Tabelle beschreibt die Informationen, die in jedem Arbeitsblatt in der Exportdatei angezeigt werden.

Tab	Beschreibung
Spaltenprofil	Zusammenfassungsinformationen, die nach Ausführung des Profils aus der Zusammenfassungsansicht exportiert wurden. Zu den Beispielen gehören Spaltennamen, Regelnamen, Anzahl der eindeutigen Werte, Anzahl der Nullwerte, abgeleitete Datentypen sowie Datum und Uhrzeit der letzten Profilausführung.
Werte	Werte für die Spalten und Regeln und die Häufigkeit, mit der die Werte für jede Spalte angezeigt werden.
Muster	Wertmuster für die Spalten und Regeln, für die Sie das Profil ausführen, und die Häufigkeit, mit der die Muster angezeigt werden.
Datentypen	Alle Datentypen für die Spalte, Häufigkeit der einzelnen Datentypen, Prozentsatz und Status des Datentyps, wie z. B. „Abgeleitet“, „Genehmigt“ oder „Abgelehnt“.
Statistiken	Statistiken über jede Spalte und Regel. Beispiele: Durchschnitt, Länge, obere und untere Werte und Standardabweichung.
Eigenschaften	Informationen der Eigenschaftenansicht. Dazu zählen Profilname, Typ, Stichprobenrichtlinie und Zeilenanzahl.

Exportieren von Profilergebnissen aus Informatica Analyst

Sie können die Ergebnisse eines Profils in eine CSV- oder XLSX-Datei exportieren, um die Daten in einer Datei anzuzeigen.

1. Wählen Sie im Arbeitsbereich **Bibliothek** das Projekt bzw. den Ordner aus, in dem das Profil enthalten ist.
2. Klicken Sie auf das Profil, um es zu öffnen.

   Die Profilergebnisse werden in der Zusammenfassungsansicht angezeigt.
3. Klicken Sie in der Zusammenfassungsansicht auf **Aktionen** > **Daten exportieren**.

   Das Dialogfeld **Daten in Datei exportieren** wird angezeigt.
4. Geben Sie im Dialogfeld **Daten in eine Datei exportieren** einen Dateinamen ein. Optional können Sie den Standarddateinamen verwenden.
5. Wählen Sie **Alle (Zusammenfassung, Werte, Muster, Datentypen, Statistiken, Eigenschaften)** oder **Ergebnisse der Datendomänenerkennung** und anschließend eine **Codepage** aus. Klicken Sie auf **OK**.

   Die Daten werden in ein Microsoft Excel-Arbeitsblatt exportiert.
6. Klicken Sie in der Zusammenfassungsansicht auf eine Spalte.

   Die Spaltenergebnisse werden in der Detailansicht angezeigt.
7. Klicken Sie in der Detailansicht auf **Aktionen** > **Daten exportieren**.

   Das Dialogfeld **Daten in Datei exportieren** wird angezeigt.
8. Geben Sie im Dialogfeld **Daten in eine Datei exportieren** einen Dateinamen ein. Optional können Sie den Standarddateinamen verwenden.
9. Wählen Sie eine der folgenden Optionen aus:
   - Alle (Zusammenfassung, Werte, Muster, Statistiken, Eigenschaften)
   - Werthäufigkeiten für die ausgewählte Spalte.
   - Musterhäufigkeiten für die ausgewählte Spalte.
   - Datentypen für die ausgewählte Spalte.
   - Drilldown-Daten für die ausgewählten Werte.
   - Drilldown-Daten für die ausgewählten Muster.
   - Drilldown-Daten für die ausgewählten Datentypen.
10. Geben Sie ein Dateiformat ein. Das Format lautet **Excel** für die Option **Alle** und **CSV** für die übrigen Optionen. Sie können den Feldnamen als erste Zeile in der Datei exportieren.
11. Wählen Sie die Codeseite der Datei.
12. Klicken Sie auf **OK**.

    Die Daten werden in die Datei exportiert.

KAPITEL 7

Scorecards in Informatica Analyst

Dieses Kapitel umfasst die folgenden Themen:

- Scorecards in Informatica Analyst - Übersicht, 59
- Informatica Analyst Scorecard-Prozess, 60
- Erstellen einer Scorecard in Informatica Analyst, 61
- Hinzufügen von Spalten zu einer vorhandenen Scorecard, 62
- Ausführen einer Scorecard, 63
- Anzeigen einer Scorecard, 64
- Bearbeiten einer Scorecard, 64
- Maße, 65
- Metrikgruppen, 67
- Spalten-Drilldown, 69
- Trenddiagramme, 69
- Scorecard-Exportdateien in Informatica Analyst, 73
- Scorecard-Benachrichtigungen, 74
- Scorecard-Herkunft, 76

Scorecards in Informatica Analyst - Übersicht

Eine Scorecard ist eine grafische Darstellung der gültigen Werte einer Spalte in einem Profil. Sie können Scorecards erstellen, um ein Drilldown an Live-Daten oder zwischengespeicherten Daten durchzuführen.

Mit Scorecards können Sie den Fortschritt der Datenqualität messen. Sie können beispielsweise eine Scorecard erstellen, um die Datenqualität vor der Anwendung der Datenqualitätsregeln zu beurteilen. Nach dem Anwenden der Datenqualitätsregeln können Sie eine weitere Scorecard erstellen, um die Auswirkung der Regeln auf die Datenqualität zu vergleichen.

Scorecards zeigen die Werthäufigkeit von Spalten als Scores (Punkte). Die Scores spiegeln den Prozentsatz der gültigen Werte in den Spalten wider. Nachdem Sie ein Profil ausgeführt haben, können Sie Spalten aus dem Profil als Metriken zu einer Scorecard hinzufügen. Sie können Metrikgruppen erstellen, so dass Sie in Verbindung stehende Metriken zu einer einzigen Entität gruppieren können. Sie können Schwellenwerte definieren, die den Bereich von ungültigen Daten, welche für Spalten akzeptabel sind, in einem Datensatz angeben, und Metrikgewichtungen für jede Metrik zuweisen. Wenn Sie eine Scorecard ausführen, generiert das Analyst Tool gewichtete Durchschnittswerte für jede Metrikgruppe. Zur weiteren Bewertung der Datenqualität können Sie jeder Metrik auch feste oder variable Kosten zuweisen. Beim Ausführen der

Scorecard berechnet das Analyst Tool die Summe der Kosten für fehlerhafte Daten für jede Metrik und zeigt die Gesamtkosten an.

Wenn Sie eine Scorecard erstellen oder bearbeiten, können Sie basierend auf den Quelldaten Scorecard-Filter erstellen. Mit den Scorecard-Filtern können Sie Metrik-Scores basierend auf der Filterbedingung neu berechnen. Zur Angabe gültiger und ungültiger Datensätze können Sie einen Drilldown für jede Metrik durchführen. Sie können Trenddiagramme verwenden, um die Änderung von Metrik-Scores sowie der Kosten für fehlerhafte Daten in Metriken über einen bestimmten Zeitraum zu verfolgen. Sie können die Profilfilter in einer Scorecard wiederverwenden.

Wenn Sie eine Scorecard erstellen oder bearbeiten, können Sie „Nativ“ oder „Hadoop“ als Laufzeitumgebung auswählen.

Wenn das Versionsverwaltungssystem im Analyst Tool aktiviert ist, können Sie mehrere Versionen einer Scorecard erstellen und die Versionshistorie für eine Scorecard anzeigen. Nach dem Erstellen einer Scorecard ist diese standardmäßig ausgecheckt. Sie müssen die Scorecard einchecken, damit die anderen Benutzer sie bearbeiten können.

Sie können die Scorecard-Details im Scorecard-Dashboard anzeigen.

Sie können E-Mail-Benachrichtigungen für Scorecards in Informatica Analyst konfigurieren und verwalten. Verwenden Sie den E-Mail-Dienst, um die E-Mail-Benachrichtigungen zu verwalten. Der E-Mail-Dienst ist ein Systemdienst, den Sie in Informatica Administrator konfigurieren können.

Informatica Analyst Scorecard-Prozess

Sie können eine Scorecard im Developer Tool und im Analyst Tool erstellen und bearbeiten. Sie können eine Scorecard im Analyst Tool ausführen. Sie können die Scorecard für aktuelle Daten im Datenobjekt oder für im Profiling Warehouse zwischengespeicherte Daten ausführen.

Sie können eine Scorecard im Arbeitsbereich **Scorecards** anzeigen. Nach dem Ausführen der Scorecard können Sie die Scores im Bereich **Scorecard** anzeigen. Sie können das Datenobjekt auswählen und von einem Score auf einer Scorecard zu dem Datenobjekt navigieren. Das Analysetool öffnet das Datenobjekt in einer anderen Registerkarte.

Sie können die folgenden Aufgaben durchführen, wenn Sie mit Scorecards arbeiten:

1. Erstellen Sie im Developer Tool oder im Analyst Tool eine Scorecard und fügen Sie Spalten aus einem Profil hinzu.
2. Öffnen Sie die Scorecard im Analyst Tool.
3. Nachdem Sie ein Profil ausgeführt haben, fügen Sie Profilspalten als Metriken zur Scorecard hinzu.
4. Erstellen Sie basierend auf den Quelldaten optional Scorecard-Filter.
5. Konfigurieren Sie optional für alle Metriken die Kosten für ungültige Daten.
6. Ausführen der Scorecard, um die Scores für Spalten zu generieren.
7. Anzeigen der Scorecard, um die Scores für jede Spalte in einem Datensatz zu sehen.
8. Drilldown an den Spalten für einen Score.
9. Bearbeiten einer Scorecard.
10. Legen Sie Schwellenwerte für jede Metrik in einer Scorecard fest.
11. Erstellen Sie eine Gruppe, um der Scorecard verwandte Metriken hinzuzufügen oder diese in die Scorecard zu verschieben.
12. Bearbeiten oder Löschen einer Gruppe nach Bedarf.

13. Zeigen Sie das score-basierte Trenddiagramm für jeden Score an, um die Änderung des Scores während eines bestimmten Zeitraums zu überwachen.
14. Zeigen Sie optional das kostenbasierte Trenddiagramm für jede Metrik an, um den Wert der Datenqualität zu überwachen.
15. Die Scorecard-Herkunft kann für jede Metrik oder Metrikgruppe angezeigt werden.

Erstellen einer Scorecard in Informatica Analyst

Erstellen Sie eine Scorecard und fügen Sie dieser Spalten aus einem Profil hinzu. Sie müssen ein Profil ausführen, bevor Sie der Scorecard Spalten hinzufügen.

1. Wählen Sie im Arbeitsbereich **Bibliothek** das Projekt bzw. den Ordner aus, in dem das Profil enthalten ist.
2. Klicken Sie auf das Profil, um es zu öffnen.

   Die Profilergebnisse werden in der Zusammenfassungsansicht im Arbeitsbereich **Erkennung** angezeigt.
3. Klicken Sie auf **Aktionen** > **Zu Scorecard hinzufügen**.

   Der Assistent **Zu Scorecard hinzufügen** wird eingeblendet.
4. Im Bildschirm **Zu Scorecard hinzufügen** können Sie eine neue Scorecard erstellen oder eine vorhandene bearbeiten, um die Spalten einer vordefinierten Scorecard hinzuzufügen. Standardmäßig ist die Option **Neue Scorecard** ausgewählt. Klicken Sie auf **Weiter**.
5. Geben Sie im Bildschirm **Schritt 2 von 8** einen Namen für die Scorecard ein. Sie können optional eine Beschreibung für die Scorecard eingeben. Wählen Sie das Projekt und den Ordner aus, in dem die Scorecard gespeichert werden soll. Klicken Sie auf **Weiter**.

   Standardmäßig wählt der Scorecard-Assistent die im Profil definierten Spalten und Regeln aus. Sie können Spalten hinzufügen, die nicht im Profil enthalten sind.
6. Wählen Sie im Bildschirm **Schritt 3 von 8** die Spalten und Regeln aus, die Sie der Scorecard als Metriken hinzufügen möchten. Aktivieren Sie optional das Kontrollkästchen in der linken Spaltenüberschrift, um alle Spalten auszuwählen. Optional können Sie auch **Spaltenname** auswählen, um die Spaltennamen zu sortieren. Klicken Sie auf **Weiter**.
7. Im Bildschirm **Schritt 4 von 8** können Sie der Metrik einen Filter hinzufügen.

   Sie können den Filter, den Sie für das Profil erstellt haben, auf Metriken anwenden oder einen neuen Filter erstellen. Wählen Sie eine Metrik im Bereich **Metrikfilter** aus und klicken Sie auf das Symbol **Filter verwalten**, um das Dialogfeld **Filter bearbeiten Spaltenname** zu öffnen. Im Dialogfeld **Filter bearbeiten: Spaltenname** können Sie eine der folgenden Aufgaben durchführen:

   - Sie können einen Filter auswählen, den Sie für das Profil erstellt haben. Klicken Sie auf **Weiter**.
   - Einen vorhandenen Filter auswählen. Klicken Sie auf das Bearbeitungssymbol, um den Filter im Dialogfeld **Filter bearbeiten** zu bearbeiten. Klicken Sie auf **Weiter**.
   - Auf das Pluszeichen (+) klicken, um im Dialogfeld **Neuer Filter** Filter zu erstellen. **Klicken Sie auf Weiter.**

   Der Filter wird im Bereich **Metrikfilter** angezeigt. Sie können denselben Filter auf alle Metriken in der Scorecard anwenden.
8. Klicken Sie im Bildschirm **Schritt 4 von 8** auf **Weiter**.
9. Wählen Sie im Bildschirm **Schritt 5 von 8** im Bereich **Metriken** die einzelnen Metriken aus und konfigurieren Sie die gültigen Werte in der Liste aller Werte im Bereich **Score verwendet: Werte**. Sie können im Bildschirm **Schritt 5 von 7** die folgenden Aufgaben durchführen:

- Sie können im Bereich **Verfügbare Werte** mehrere Werte auswählen und auf den Pfeil nach rechts klicken, um diese in den Bereich **Gültige Werte** zu verschieben. Die Gesamtzahl der gültigen Werte für eine Metrik wird oben im Bereich **Verfügbare Werte** angezeigt.
- Konfigurieren Sie im Bereich **Metrikschwellenwerte** Metrikschwellenwerte.

  Einstellbar sind Schwellenwerte für die Scores **Gut**, **Akzeptabel** und **Inakzeptabel**.
- Wählen Sie alle Metriken aus und konfigurieren Sie die Kosten für ungültige Daten. Um den Kosten für die Metrik einen konstanten Wert zuzuweisen, wählen Sie die Option **Feste Kosten** aus. Klicken Sie optional auf **Kosteneinheit ändern**, um die Kosteneinheit zu ändern, oder wählen Sie **Keine** aus. Um eine numerische Spalte für variable Kosten an die Metrik anzuhängen, wählen Sie **Variable Kosten** aus und klicken zum Auswählen einer numerischen Spalte auf **Spalte auswählen**.

10. Im Bildschirm **Schritt 6 von 8** können Sie eine Metrikgruppe auswählen, der Sie die Metriken hinzufügen können, oder eine neue Metrikgruppe erstellen. Zum Erstellen einer neuen Metrikgruppe klicken Sie auf das Gruppensymbol. Klicken Sie auf **Weiter**.
11. Geben Sie im Bildschirm **Schritt 7 von 8** die Gewichtungen für die Metriken in der Gruppe sowie Schwellenwerte für die Gruppe an.
12. Wählen Sie im Bildschirm **Schritt 8 von 8** die Option **Nativ** oder **Hadoop** als Laufzeitumgebung zur Ausführung der Scorecard aus.
13. Klicken Sie auf **Speichern**, um die Scorecard zu speichern, oder auf **Speichern und ausführen**, um die Scorecard zu speichern und auszuführen.

    Die Scorecard wird im Arbeitsbereich **Scorecard** angezeigt.

Hinzufügen von Spalten zu einer vorhandenen Scorecard

Nach dem Ausführen eines Profils können Sie einer vorhandenen Scorecard Spalten hinzufügen, die gültigen Werte für die Spalten konfigurieren und die Kosten für ungültige Daten für jede Metrik hinzufügen. Wenn Sie einer Scorecard eine Spalte aus einem Profil hinzufügen und einen anderen Filter oder eine andere Stichprobenoption als **Alle Zeilen** verwenden, spiegeln die Profilergebnisse möglicherweise nicht die Scorecard-Ergebnisse wider.

1. Klicken Sie auf ein Profil, um es zu öffnen.

   Die Profilergebnisse werden in der Zusammenfassungsansicht angezeigt.
2. Wählen Sie eine Spalte aus. Klicken Sie auf **Aktionen** > **Zu Scorecard hinzufügen**.

   Der Assistent **Zu Scorecard hinzufügen** wird eingeblendet.

   Hinweis: Bevor Sie Spalten zu einer Scorecard hinzufügen, beachten Sie bitte folgende Regeln und Richtlinien:

   - Sie können einer Scorecard keine Spalte hinzufügen, wenn Spaltenname und Scorecard-Name übereinstimmen.
   - Das zweimalige Hinzufügen ein- und derselben Spalte zu einer Scorecard ist auch dann nicht möglich, wenn Sie den Spaltennamen ändern.
3. Wählen Sie **Vorhandene Scorecard** aus, um die Spalten einer vordefinierten Scorecard hinzuzufügen. Klicken Sie auf **Weiter**.

4. Wählen Sie im Bildschirm **Schritt 2 von 7** die Scorecard, die Metriken und die Metrikgruppe aus, denen Sie die Spalten hinzufügen möchten. Klicken Sie auf **Weiter**.
5. Wählen Sie im Bildschirm **Schritt 3 von 7** die Spalten und Regeln aus, die Sie der Scorecard als Metriken hinzufügen möchten. Aktivieren Sie optional das Kontrollkästchen in der linken Spaltenüberschrift, um alle Spalten auszuwählen. Klicken Sie auf **Spaltenname**, um die Spaltennamen zu sortieren. Klicken Sie auf **Weiter**.
6. Im Bildschirm **Schritt 4 von 7** können Sie Filter für die Metriken erstellen. Sie können den Filter, den Sie für das Profil erstellt haben, auch auf Metriken anwenden.
7. Im Bildschirm **Schritt 5 von 7** können Sie die folgenden Aufgaben durchführen:
   - Wählen Sie im Bereich **Metriken** die einzelnen Metriken aus und konfigurieren Sie Metrikwerte in den anderen Bereichen.
   - Wählen Sie im Bereich **Score verwendet: Werte** unter **Verfügbare Werte** mehrere Werte aus und klicken Sie auf den Pfeil nach rechts, um diese in den Bereich **Gültige Werte** zu verschieben.
     Die Gesamtzahl der gültigen Werte für eine Metrik wird oben im Bereich **Verfügbare Werte** angezeigt.
   - Im Bereich **Metrikschwellenwerte:** können Sie Schwellenwerte für die Scores **Gut**, **Akzeptabel** und **Inakzeptabel** festlegen.
   - Im Bereich **Kosten für ungültige Daten** können Sie die folgenden Aufgaben durchführen:
     - Wählen Sie alle Metriken aus und konfigurieren Sie die Kosten für ungültige Daten für die jeweilige Metrik.
     - Wählen Sie die Option **Feste Kosten** aus, um den Kosten für die Metrik einen konstanten Wert zuzuweisen. Sie können auf **Kosteneinheit ändern** klicken, um die Kosteneinheit zu ändern.
     - Wählen Sie die Option **Variable Kosten** aus, um eine numerische Spalte für variable Kosten an die Metrik anzuhängen. Sie können auf **Spalte auswählen** klicken, um eine numerische Spalte auszuwählen.
8. Klicken Sie auf **Weiter**.
9. Im Bildschirm **Schritt 6 von 7** können Sie die folgenden Aufgaben durchführen:
   - Wählen Sie die Metrikgruppe aus, der Sie die Metriken hinzufügen möchten.
   - Im Bereich **Standard - Metriken** können Sie auf die standardmäßige Metrikgewichtung 0 doppelklicken, um den Wert zu ändern.
   - Im Bereich **Metrikschwellenwerte:** können Sie Schwellenwerte für die Scores **Gut**, **Akzeptabel** und **Inakzeptabel** festlegen.
10. Klicken Sie auf **Weiter**.
11. Wählen Sie im Bildschirm **Schritt 7 von 7** eine Laufzeitumgebung aus.
12. Klicken Sie auf **Speichern**, um die Scorecard zu speichern, oder auf **Speichern und ausführen**, um die Scorecard zu speichern und auszuführen.

Ausführen einer Scorecard

Führen Sie eine Scorecard aus, um Scores für Spalten zu generieren.

1. Wählen Sie im Bereich **Objekte** die Scorecard aus, die Sie ausführen möchten.
2. Klicken Sie auf die Scorecard, um sie zu öffnen.

   Die Scorecard wird im **Scorecards**-Arbeitsbereich angezeigt.

3. Klicken Sie auf **Aktionen** > **Scorecard ausführen.**
4. Wählen Sie einen Score aus dem Bereich **Metriken** und wählen Sie die Spalten für den Drilldown im Bereich **Spalten** aus.
5. Mit der Option **Drilldown** können Sie wählen, ob der Drilldown an Live-Daten oder zwischengespeicherten Daten durchgeführt werden soll.

   Optimale Leistung erzielen Sie mit einem Drilldown für Live-Daten.
6. Klicken Sie auf **Ausführen**.

Anzeigen einer Scorecard

Führen Sie eine Scorecard aus, um die Scores für jede Metrik anzuzeigen. Eine Scorecard zeigt den Score als Prozentsatz und Balkendiagramm an. Zeigen Sie Daten an, die gültig oder nicht gültig sind. Sie können auch Scorecard-Informationen wie Metrikgewichtung, Metrikgruppen-Score, Score-Trend und Name des Datenobjekts anzeigen.

1. Führen Sie eine Scorecard aus, um die Scores anzuzeigen.
2. Wählen Sie eine Metrik aus, die den anzuzeigenden Score enthält.
3. Klicken Sie auf **Aktionen** > **Drilldown**, um die Zeilen mit gültigen Daten oder Zeilen mit Daten, die für die Spalte nicht gültig sind, anzuzeigen.

   Im Analyst-Tool werden die Zeilen der ungültigen Daten standardmäßig im Abschnitt **Drilldown** angezeigt.

Bearbeiten einer Scorecard

Bearbeiten Sie gültige Werte für Metriken in einer Scorecard. Bevor Sie eine Scorecard bearbeiten können, müssen Sie sie ausführen.

1. Klicken Sie im Arbeitsbereich **Bibliothek** auf die Scorecard, die Sie im Bereich **Objekte** bearbeiten möchten.

   Die Scorecard wird im Arbeitsbereich **Scorecards** angezeigt.
2. Falls das Versionsverwaltungssystem aktiviert ist, klicken Sie auf **Aktionen** > **Auschecken**.
3. Klicken Sie auf **Aktionen** > **Bearbeiten** > **Allgemein**.

   Das Dialogfeld **Scorecard bearbeiten** wird eingeblendet.
4. Auf der Registerkarte **Allgemein** können Sie gegebenenfalls den Namen und die Beschreibung der Scorecard bearbeiten.
5. Klicken Sie auf die Registerkarte **Metriken**.
6. Wählen Sie einen Score im Bereich **Metriken** aus und konfigurieren Sie die gültigen Werte in der Liste aller Werte im Bereich **Score verwendet: Werte**.
7. Nehmen Sie im Bereich **Metrikschwellenwerte** gegebenenfalls Änderungen an den Schwellenwerten des Scores vor.

8. Überprüfen Sie die Kosten für ungültige Daten für jede Metrik und nehmen Sie notwendige Änderungen vor.
9. Klicken Sie auf die Registerkarte **Scorecard-Filter**.
10. Sie können Filter hinzufügen, bearbeiten oder löschen.
11. Klicken Sie auf die Registerkarte **Metrikgruppen**.
12. Sie können die Metrikgruppen erstellen, bearbeiten oder entfernen.

    Sie können darüber hinaus die Metrikgewichtungen und Metrikschwellenwerte auf der Registerkarte **Metrikgruppen** bearbeiten.
13. Klicken Sie auf die Registerkarte **Benachrichtigungen**.
14. Sie können gegebenenfalls Änderungen an den Einstellungen für Scorecard-Benachrichtigungen vornehmen.

    Sie können globale und benutzerdefinierte Einstellungen für Metriken und Metrikgruppen festlegen.
15. Klicken Sie auf die Registerkarte **Laufzeitumgebung**.

    Sie können **Nativ** oder **Hadoop** als Laufzeitumgebung auswählen.
16. Klicken Sie auf **Speichern**, um Änderungen an der Scorecard zu speichern, oder auf **Speichern und Ausführen**, um die Änderungen zu speichern und die Scorecard auszuführen.
17. Klicken Sie auf **Einchecken**.

Maße

Eine Metrik ist eine Spalte einer Datenquelle oder Ausgabe einer Regel, die Teil einer Scorecard ist. Wenn Sie eine Scorecard erstellen, können Sie jeder Metrik eine Gewichtung zuweisen. Erstellen Sie eine Metrikgruppe, um verwandte Metriken in einer Scorecard zu einem Satz zu kategorisieren.

Metrikgewichtungen

Wenn Sie eine Scorecard erstellen, können Sie jeder Metrik eine Gewichtung zuweisen. Der Standardwert für eine Gewichtung lautet 0.

Wenn Sie eine Scorecard ausführen, berechnet das Analyst-Tool den gewichteten Durchschnitt für jede Metrikgruppe basierend auf dem Metrik-Score und der Gewichtung, die Sie jeder Metrik zuweisen.

Beispielsweise weisen Sie eine Gewichtung von W1 zu Metrik M1 und eine Gewichtung von W2 zu Metrik M2 zu. Das Analyst-Tool verwendet die folgende Formel zur Berechnung des gewichteten Durchschnitts:

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

Wert der Datenqualität

Die Bewertung der Datenqualität in den Quelldaten stellt für die Verwaltung der Datenbestände des Unternehmens eine äußerst wichtige Information dar. Die Kosten für ungültige Daten in Metriken, die in einer Scorecard dargestellt werden, unterstützen Unternehmen dabei, die Datenqualität der Quelldaten nutzbringend zu überwachen. Als Datenanalyst möchten Sie einen Wert, wie beispielsweise eine Währungseinheit oder eine beliebige benutzerdefinierte Einheit, mit Metriken und Metrikgruppen verbinden.

Anschließend können Sie die Scorecard ausführen, um die Gesamtkosten für ungültige Daten in den Quelldaten anzuzeigen.

Sie können die Kosten für eine Metrik basierend auf den Geschäftsanforderungen definieren. Darüber hinaus können Sie variable oder feste Kosten für jede Metrik konfigurieren, wenn Sie eine Scorecard erstellen oder bearbeiten.

Feste Kosten

Feste Kosten sind ein konstanter Wert, der einer Metrik in einer Scorecard zugewiesen werden kann. Sie können eine vordefinierte Kosteneinheit auswählen oder eine benutzerdefinierte Kosteneinheit erstellen, die die Geschäftsanforderungen erfüllt.

Variable Kosten

Bei den variablen Kosten handelt es sich um einen Wert, den Sie einer Metrik basierend auf den Werten in einer numerischen Spalte einer Datenquelle zuweisen. Der Datenintegrationsdienst berechnet die variablen Kosten für die Metrik basierend auf der Spalte oder der virtuellen Spalte, die den Kosten zugewiesen wurde.

Beispiel

Als Kreditberater müssen Sie Ihren Kunden Tilgungspläne zur Verfügung stellen, damit diese ihre Darlehenszahlungen leisten können. Sie können eine Scorecard verwenden, um die Genauigkeit der Kundenadressen zu messen und somit die Zustellung der Tilgungspläne zu gewährleisten. Unter Umständen möchten Sie die variablen Kosten der Spalte mit dem monatlichen Zahlungsbetrag für die Metrik „Adressgenauigkeit“ festlegen. Führen Sie die Scorecard aus, um die Gesamtkosten zu berechnen, die dem Kreditinstitut entstehen, wenn die Kunden die monatlichen Zahlungen nicht pünktlich leisten.

Definieren von Schwellenwerten

Für jeden Score einer Scorecard können Schwellenwerte eingestellt werden. Ein Schwellenwert gibt den Bereich als Prozentsatz falscher Daten, die für Spalten in einem Datensatz akzeptabel sind, an. Sie können Schwellenwerte für gute, akzeptable oder inakzeptable Datenbereiche einstellen. Beim Hinzufügen von Spalten zu einer Scorecard oder beim Bearbeiten einer Scorecard können Sie ebenfalls Schwellenwerte für jede Spalte definieren.

Vor dem Definieren der Schwellenwerte für Spalten in einer Scorecard muss eine der folgenden Aufgaben erledigt werden:

- Öffnen Sie ein Profil und fügen Sie im Dialogfeld **Zu Scorecard hinzufügen** Spalten aus dem Profil zur Scorecard hinzu.
- Sie können auch im Arbeitsbereich **Bibliothek** auf eine Scorecard klicken und **Aktionen** > **Bearbeiten** auswählen, um die Scorecard im Dialogfeld **Scorecard bearbeiten** zu bearbeiten.

1. Wählen Sie im Dialogfeld **Zu Scorecard hinzufügen** oder im Dialogfeld **Scorecard bearbeiten** jede Metrik im Bereich **Metriken** aus.
2. Geben Sie im Bereich **Metrikschwellenwerte** die Schwellenwerte ein, mit denen die obere Grenze des inakzeptablen Bereichs und die untere Grenze des guten Bereichs wiedergegeben werden.
3. Klicken Sie auf **Weiter** oder **Speichern**.

Metrikgruppen

Erstellen Sie eine Metrikgruppe, um verwandte Scores in einer Scorecard zu einem Satz zu kategorisieren. Das Analyst-Tool kategorisiert standardmäßig alle Scores in einer Standardmetrikgruppe.

Nachdem Sie eine Metrikgruppe erstellt haben, können Sie Scores aus der Standardmetrikgruppe in eine andere Metrikgruppe verschieben. Sie können eine Metrikgruppe ändern, um deren Namen und Beschreibung, einschließlich der Standardmetrikgruppe, zu ändern. Sie können Metrikgruppen löschen, die Sie nicht länger verwenden. Sie können die Standardmetrikgruppe nicht löschen.

Erstellen einer Metrikgruppe

Erstellen Sie eine Gruppe, um ähnliche Scores in der Scorecard zur Gruppe hinzuzufügen.

1. Klicken Sie im Arbeitsbereich **Bibliothek** auf die Scorecard, die Sie im Bereich **Objekte** bearbeiten möchten.

   Die Scorecard wird im Arbeitsbereich **Scorecards** angezeigt.
2. Klicken Sie auf **Aktionen** > **Bearbeiten**.

   Das Fenster **Scorecard bearbeiten** wird angezeigt.
3. Klicken Sie auf die Registerkarte **Metrikgruppen**.

   Die Standardgruppe wird im Bereich **Metrikgruppen** angezeigt, und die Scores in der Standardgruppe werden im Bereich **Metriken** angezeigt.
4. Klicken Sie auf das Symbol **Neue Gruppe**, um eine Metrikgruppe zu erstellen.

   Das Dialogfeld **Metrikgruppen** wird eingeblendet.
5. Geben Sie einen Namen und eine optionale Beschreibung ein.
6. Klicken Sie auf **OK**.
7. Klicken Sie auf **Speichern**, um die Änderungen an der Scorecard zu speichern.

Verschieben von Scores in eine Metrikgruppe

Nachdem Sie eine Metrikgruppe erstellt haben, können Sie zugehörige Scores in die Metrikgruppe verschieben.

1. Klicken Sie im Arbeitsbereich **Bibliothek** auf die Scorecard, die Sie im Bereich **Objekte** bearbeiten möchten.

   Die Scorecard wird im Arbeitsbereich **Scorecards** angezeigt.
2. Klicken Sie auf **Aktionen** > **Bearbeiten**.

   Das Fenster **Scorecard bearbeiten** wird angezeigt.
3. Klicken Sie auf die Registerkarte **Metrikgruppen**.

   Die Standardgruppe wird im Bereich **Metrikgruppen** angezeigt, und die Scores in der Standardgruppe werden im Bereich **Metriken** angezeigt.
4. Wählen Sie eine Metrik aus dem Bereich **Metriken** aus und klicken Sie auf das Symbol **Metriken verschieben**.

   Das Dialogfeld **Metriken verschieben** wird eingeblendet.

   Hinweis: Um mehrere Scores auszuwählen, halten Sie die Umschalttaste gedrückt.
5. Wählen Sie die Metrikgruppe aus, in die die Scores verschoben werden sollen.

6. Klicken Sie auf **OK**.

Bearbeiten einer Metrikgruppe

Bearbeiten Sie eine Metrikgruppe zum Ändern des Namens und der Beschreibung. Sie können den Namen der Standardmetrikgruppe ändern.

1. Klicken Sie im Arbeitsbereich **Bibliothek** auf die Scorecard, die Sie im Bereich **Objekte** bearbeiten möchten.

   Die Scorecard wird im Arbeitsbereich **Scorecards** angezeigt.
2. Klicken Sie auf **Aktionen** > **Bearbeiten**

   Das Fenster **Scorecard bearbeiten** wird angezeigt.
3. Klicken Sie auf die Registerkarte **Metrikgruppen**.

   Die Standardmetrikgruppe wird im Bereich **Metrikgruppen** und die Metriken in der Standardmetrikgruppe im Bereich **Metriken** angezeigt.
4. Klicken Sie im Bereich **Metrikgruppen** auf das Symbol **Gruppe bearbeiten**.

   Das Dialogfeld **Bearbeiten** wird eingeblendet.
5. Geben Sie den Namen und optional eine Beschreibung ein.
6. Klicken Sie auf **OK**.

Löschen einer Metrikgruppe

Sie können eine Metrikgruppe, die nicht mehr gültig ist, löschen. Beim Löschen einer Metrikgruppe können Sie die Scores in der Metrikgruppe in die Standardmetrikgruppe verschieben. Sie können die Standardmetrikgruppe nicht löschen.

1. Klicken Sie im Arbeitsbereich **Bibliothek** auf die Scorecard, die Sie im Bereich **Objekte** bearbeiten möchten.

   Die Scorecard wird im Arbeitsbereich **Scorecards** angezeigt.
2. Klicken Sie auf **Aktionen** > **Bearbeiten**.

   Das Fenster **Scorecard bearbeiten** wird angezeigt.
3. Klicken Sie auf die Registerkarte **Metrikgruppen**.

   Die Standardmetrikgruppe wird im Bereich **Metrikgruppen** und die Metriken in der Standardmetrikgruppe im Bereich **Metriken** angezeigt.
4. Wählen Sie eine Metrikgruppe im Bereich **Metrikgruppen** und klicken Sie auf das Symbol **Gruppe löschen**.

   Das Dialogfeld **Gruppen löschen** wird eingeblendet.
5. Wählen Sie die Option zum Löschen der Metriken in der Metrikgruppe oder die Option zum Verschieben der Metriken in die Standardmetrikgruppe, bevor die Metrikgruppe gelöscht wird.
6. Klicken Sie auf **OK**.

Spalten-Drilldown

Drilldown für die Spalten eines Score zum Auswählen von Spalten, die beim Einblenden der gültigen oder ungültigen Datenzeilen erscheinen. Die Spalten, die Sie zum Drilldown auswählen, werden im Fenster **Drilldown** angezeigt.

1. Führen Sie eine Scorecard aus, um die Scores anzuzeigen.
2. Wählen Sie eine Spalte, die den anzuzeigenden Score enthält.
3. Klicken Sie auf **Aktionen** > **Zeilen anzeigen**, um die Zeilen der gültigen bzw. ungültigen Daten der Spalte anzuzeigen.
4. Klicken Sie auf **Aktionen** > **Drilldown-Spalten**

   Die Spalten erscheinen im Fenster **Drilldown** für den ausgewählten Score. Im Analyst Tool werden standardmäßig die Zeilen mit gültigen Daten für die Spalten angezeigt. Optional können Sie auf **Ungültig** klicken, um die ungültigen Datenzeilen anzuzeigen.

Trenddiagramme

Verwenden Sie Trenddiagramme, um die Änderung von Metrik-Scores sowie der Kosten für ungültige Daten in Metriken über einen bestimmten Zeitraum zu überwachen.

Die Trenddiagramme enthalten score- und kostenbasierte Diagramme, die die Score- und Kostenwerte auf der vertikalen Achse und alle Scorecard-Ausführungen auf der horizontalen Achse abbilden. Standardmäßig werden in einem Trenddiagramm die Daten aus den letzten 10 Scorecard-Ausführungen angezeigt. Sie können die Gesamtzahl der Zeilen sowie der ungültigen Zeilen für die Metrik im Trenddiagramm anzeigen. Das Trenddiagramm zeigt darüber hinaus, ob die Score- und Kostentrends konstant geblieben sind oder sich basierend auf der letzten Scorecard-Ausführung noch oben oder unten bewegt haben.

Das Analyst-Tool nutzt für jedes Datum historische Scorecard-Ausführungsdaten und für die Berechnung des Scores aktuell gültige Score-Werte. Für die Darstellung der Farbe der Score-Punkte verwendet das Analyst-Tool die aktuellen Schwellenwerteinstellungen im Diagramm. Sie können die Schwellenwerte „Gut“, „Akzeptabel“ und „Inakzeptabel“ für den Score anzeigen. Nach der Bearbeitung der Werte für Scores in der Scorecard ändern sich die Schwellenwerte bei jeder Scorecard-Ausführung. Wenn Sie eine Scorecard exportieren, enthält das Analyst-Tool die Trenddiagramminformationen, einschließlich der Score- und Kosteninformationen in der exportierten Datei.

Score-basiertes Trenddiagramm

In einem score-basierten Trenddiagramm wird die Änderung des Metrik-Scores während mehrerer Profilausführungen grafisch dargestellt. Das score-basierte Trenddiagramm bildet die Werte des Metrik-Scores auf der vertikalen Achse und alle Scorecard-Ausführungen auf der horizontalen Achse ab.

Die folgende Abbildung zeigt ein Beispiel für ein score-basiertes Trenddiagramm:

Beispiel

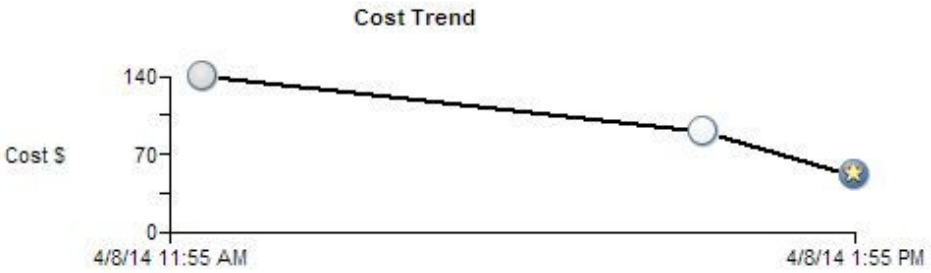
Als Datenanalyst können Sie die Datenqualität überwachen und analysieren, ob die Änderungen an Mappings und anderen Prozessen zu einer Erhöhung des Datenqualitäts-Scores führen. Nachdem Sie die Änderung in der Datenqualität gemessen haben, können Sie die Datenqualitätsänderung zurückmelden, damit sie vom Unternehmen analysiert und verwendet werden kann. Beispiel: Am Ende mehrerer Scorecard-Ausführungen hat sich der Prozentsatz der gültigen Werte in einer Spalte mit Sozialversicherungsnummern unter Umständen von 84 auf 90 geändert. Sie können diese Änderung in der Datenqualität als visuelles Diagramm zur schnellen Analyse bereitstellen.

Kostenbasiertes Trenddiagramm

Bei einem kostenbasierten Trenddiagramm handelt es sich um eine grafische Darstellung, in der die Änderung der Kosten für ungültige Daten in Metriken während mehrerer Profilausführungen abgebildet wird. Mit kostenbasierten Trenddiagrammen kann der Einfluss der Datenqualität in einem Unternehmen gemessen werden. Das kostenbasierte Trenddiagramm bildet die Kostenwerte auf der vertikalen Achse und alle Scorecard-Ausführungen auf der horizontalen Achse ab. Sie können die Gesamtkosten für ungültige Daten und die gültigen Werte für die Metrik in einem Raster unter dem kostenbasierten Trenddiagramm anzeigen.

Mit einem kostenbasierten Trenddiagramm können Sie die Auswirkungen ungültiger Daten auf Datensätze mit hohen Werten verfolgen. Wenn Sie gelegentlich feste Kosten zum Berechnen ungültiger Daten verwenden, entgehen Ihnen unter Umständen die Auswirkungen ungültiger Daten auf Datensätze mit hohen Werten. Dieses Problem tritt auf, weil in den Trenddiagrammen unter Umständen eine Verbesserung im Score und eine Abnahme in den Gesamtkosten während mehrerer Scorecard-Ausführungen angezeigt werden. Je weniger Datenqualitätsprobleme jedoch in der Scorecard auftreten, desto höher ist die Qualität der Datensätze mit hohen Werten.

Die folgende Abbildung zeigt ein Beispiel für ein kostenbasiertes Trenddiagramm:



Beispiel

In einem Kreditinstitut gibt es mehrere Kunden mit hohen Salden sowie hohen Einlagen und Investitionen, wie z. B. 10 Millionen US-Dollar. Es gibt aber auch eine große Anzahl an Kunden mit niedrigen Salden. Das score-basierte Trenddiagramm zeigt unter Umständen eine Verbesserung in Scores über einen bestimmten Zeitraum an. Jedoch kann selbst für wenige Kundenkonten mit hohen Salden die Angabe einer falschen Adresse oder einer falschen Anrede die Beziehungen zu den wichtigsten Kunden des Unternehmens beeinflussen. Sie können die Spalte „Kontostand“ als variable Kostenspalte für die Berechnung ungültiger Daten verwenden. Wenn die Kosten für ungültige Daten aufgrund dieser Spalte hoch sind, können Sie das potenzielle finanzielle Risiko einschätzen und sofort korrigierende Maßnahmen ergreifen.

Anzeigen von Trenddiagrammen

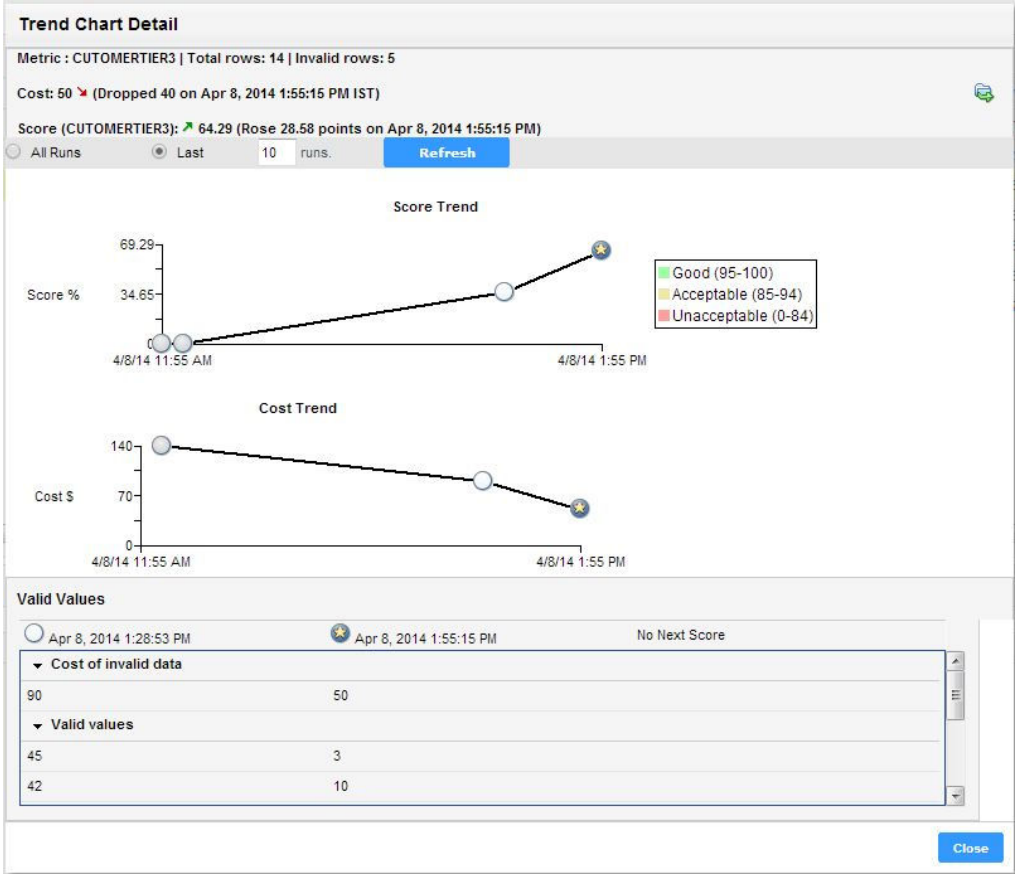
Sie können Trenddiagramme für jede Metrik anzeigen, um die Änderung des Scores oder der Kosten für ungültige Daten über einen bestimmten Zeitraum zu überwachen.

1. Wählen Sie im Arbeitsbereich **Bibliothek** das Projekt bzw. den Ordner aus, in dem die Scorecard enthalten ist.
2. Klicken Sie auf die Scorecard, um sie zu öffnen.

   Die Scorecard wird im **Scorecards**-Arbeitsbereich angezeigt.
3. Wählen Sie in der Ansicht **Scorecard** eine Metrik aus.

4. Klicken Sie auf **Aktionen** > **Trenddiagramm anzeigen**.

   Das Dialogfeld **Trenddiagramm-Detail** wird angezeigt.
   Die folgende Abbildung zeigt das Dialogfeld **Trenddiagramm-Detail**:



   Sie können Score- und Kostenwerte anzeigen, die sich im Laufe der Zeit geändert haben. Im oberen Bereich des Dialogfelds können Sie die Gesamtanzahl der Zeilen und die Anzahl der ungültigen Zeilen anzeigen. Das Analyst Tool nutzt für jedes Datum historische Scorecard-Ausführungsdaten und für die Berechnung des Scores die neuesten Score-Werte. Unter den score- und kostenbasierten Trenddiagrammen können Sie die gültigen Werte für die Metrik und die Kosten für ungültige Daten anzeigen.

Exportieren von Trenddiagrammen

Sie können die score- und kostenbasierten Trenddiagramme in eine XLSX-Datei exportieren, um die Daten in einer Datei anzuzeigen.

1. Öffnen Sie eine Scorecard.
2. Wählen Sie eine Metrik aus und klicken Sie auf **Aktionen** > **Trenddiagramm anzeigen**.

   Das Dialogfeld **Trenddiagramm-Details** wird angezeigt.
3. Klicken Sie auf das Symbol **Daten exportieren**.

   Das Dialogfenster **Daten in Datei exportieren** wird geöffnet.
4. Geben Sie einen Dateinamen ein. Optional können Sie den Standarddateinamen verwenden.

   Das Standarddateiformat ist Microsoft Excel.

5. Wählen Sie die Codepage der Datei aus.
6. Klicken Sie auf **OK**.

Scorecard-Exportdateien in Informatica Analyst

Sie können Scorecard-Ergebnisse in eine Microsoft Excel-Datei exportieren. Das Analyst-Tool speichert die Datei im XLSX-Format.

Die exportierte Datei enthält eine Scorecard-Zusammenfassung, Trenddiagramme, ungültige Zeilen und Scorecard-Eigenschaften.

Scorecard-Exportergebnisse in Microsoft Excel

Beim Exportieren der Scorecard-Ergebnisse speichert das Analyst-Tool die Informationen in mehreren Arbeitsblättern in einer Microsoft Excel-Datei. Das Analyst Tool speichert die Datei im XLSX-Format.

Die folgende Tabelle beschreibt die Informationen, die in jedem Arbeitsblatt in der Exportdatei angezeigt werden.

Tab	Beschreibung
Scorecard-Zusammenfassung	Zusammenfassung der exportierten Scorecard-Ergebnisse. Zu den Informationen gehören der Scorecard-Name, die Gesamtanzahl der Zeilen für jede Spalte, die Anzahl der ungültigen Zeilen, der Score und die Metrikgewichtung.
Trenddiagramm	Trenddiagramme für Scores.
Ungültige Zeilen	Die Details der ungültigen Zeilen für jede Spalte. Das Analyst Tool exportiert maximal 100 Zeilen in das Arbeitsblatt.
Eigenschaften	Scorecard-Eigenschaften, wie z. B. Name, Typ, Beschreibung und Speicherort.

Exportieren von Scorecard-Ergebnissen aus Informatica Analyst

Sie können Scorecard-Ergebnisse in eine XLSX-Datei exportieren, um die Daten in einer Datei anzuzeigen.

1. Öffnen Sie eine Scorecard.
2. Klicken Sie auf **Aktionen** > **Daten exportieren**.

   Das Dialogfeld **Daten in Datei exportieren** wird geöffnet.
3. Geben Sie einen Dateinamen ein. Optional können Sie den Standarddateinamen verwenden.

   Das Standarddateiformat ist Microsoft Excel.
4. Wählen Sie die Codepage der Datei aus.
5. Klicken Sie auf **OK**.

Scorecard-Benachrichtigungen

Konfigurieren Sie Einstellungen für Scorecard-Benachrichtigungen, damit vom Analyst Tool E-Mails versendet werden, sobald bestimmte Metrik-Scores, Metrikgruppen-Scores oder Metrikkosten Schwellenwerte über- oder unterschreiten. Metrik-Scores oder Metrikgruppen-Scores über- oder unterschreiten unter Umständen Schwellenwerte oder verbleiben in bestimmten Score-Bereichen, wie z. B. „Inakzeptabel“, „Akzeptabel“ und „Gut“. Metrikkostenwerte können die von Ihnen festgelegten Kostenschwellenwerte über- und unterschreiten.

Sie können E-Mail-Benachrichtigungen für einzelne Metrik-Scores, Metrikgruppen und Metrikkosten konfigurieren. Bei Verwendung der globalen Einstellungen für Scores sendet das Analyst Tool Benachrichtigungs-E-Mails, wenn bestimmte Metrik-Scores den Schwellenwert aus dem Score-Bereich „Gut“ zu „Akzeptabel“ und „Akzeptabel“ zu „Schlecht“ über- oder unterschreiten. Sie erhalten auch Benachrichtigungs-E-Mails für jede Scorecard-Ausführung, wenn der Score über mehrere aufeinanderfolgende Scorecard-Ausführungen hinweg weiterhin im Score-Bereich „Inakzeptabel“ liegt. Bei Verwendung der globalen Einstellungen für Metrikkosten, sendet das Analyst Tool Benachrichtigungs-E-Mails, wenn die Kosten für ungültige Daten in ausgewählten Metriken die Schwellenwerte über- oder unterschreiten.

Sie können die Benachrichtigungseinstellungen so anpassen, dass Scorecard-Benutzer E-Mail-Benachrichtigungen erhalten, wenn die Scores vom Score-Bereich Inakzeptabel zu Akzeptabel und von Akzeptabel zu Gut wechseln. Sie können E-Mail-Benachrichtigungen senden, wenn ein Metrik-Score oder Metrikkosten für jede Scorecard-Ausführung innerhalb bestimmter Bereiche liegen. Sie können die aktuellen Kosten für ungültige Daten für jede Metrik in den Benachrichtigungseinstellungen anzeigen, auf deren Basis die Kostenschwellenwerte eingerichtet werden können.

Bevor Sie Scorecards zum Senden von E-Mail-Benachrichtigungen konfigurieren, müssen Sie den E-Mail-Dienst im Administrator Tool konfigurieren.

Vorlage für Benachrichtigungs-E-Mail

Sie können den Nachrichtentext und die Struktur der E-Mail-Benachrichtigungen festlegen, die das Analyst-Tool als Teil von Scorecard-Benachrichtigungen an Empfänger sendet. Die E-Mail-Vorlage hat einen optionalen Abschnitt für einführenden Text, einen Hauptabschnitt für Nachrichten mit Lesezugriff und einen optionalen Abschnitt für den abschließenden Text.

Die folgende Tabelle beschreibt die Tags in der E-Mail-Vorlage:

Tag	Beschreibung
ScorecardName	Name der Scorecard
ObjectURL	Ein Hyperlink zur Scorecard. Sie müssen Benutzernamen und Kennwort angeben.
MetricGroupName	Name der Metrikgruppe, zu der die Metrik gehört
CurrentWeightedAverage	Gewichteter Durchschnittswert für die Metrikgruppe im aktuellen Scorecard-Lauf
CurrentRange	Der Score-Bereich, z. B. Inakzeptabel, Akzeptabel und Gut, für die Metrikgruppe im aktuellen Scorecard-Lauf
PreviousWeightedAverage	Gewichteter Durchschnittswert für die Metrikgruppe im vorherigen Scorecard-Lauf

Tag	Beschreibung
PreviousRange	Der Score-Bereich, z. B. Inakzeptabel, Akzeptabel und Gut, für die Metrikgruppe im vorherigen Scorecard-Lauf
MetricName	Name der Metrik.
MetricGroupName	Name der Metrikgruppe.
CurrentScore	Score basierend auf der aktuellen Scorecard-Ausführung.
CurrentRange	Score-Bereich, in dem der aktuelle Score basierend auf der aktuellen Scorecard-Ausführung verbleibt.
PreviousScore	Score basierend auf der vorherigen Scorecard-Ausführung.
PreviousRange	Score-Bereich basierend auf der vorherigen Scorecard-Ausführung.
CurrentCost	Kosten für ungültige Daten in der Metrik basierend auf der aktuellen Scorecard-Ausführung.
PreviousCost	Kosten für ungültige Daten in der Metrik basierend auf der vorherigen Scorecard-Ausführung.
ColumnName	Name der Quellspalte, die der Metrik zugewiesen ist
ColumnType	Typ der Quellspalte
RuleName	Name der Regel
RuleType	Typ der Regel
DataObjectName	Name des Quelldatenobjekts

Einrichten der Scorecard-Benachrichtigungen

Sie können Scorecard-Benachrichtigungen sowohl auf Metrik- als auch auf Metrikgruppenebene einrichten. Globale Benachrichtigungseinstellungen gelten für Metriken und Metrikgruppen, die keine individuellen Benachrichtigungseinstellungen haben.

1. Führen Sie eine Scorecard im Analyst-Tool aus.
2. Klicken Sie auf **Aktionen** > **Bearbeiten**.
3. Klicken Sie auf die Registerkarte **Benachrichtigungen**.
4. Wählen Sie **Benachrichtigungen aktivieren** aus, um mit der Konfiguration von Scorecard-Benachrichtigungen zu beginnen.
5. Wählen Sie eine Metrik oder Metrikgruppe aus.
6. Klicken Sie auf das Kontrollkästchen **Benachrichtigungen**, um die globalen Einstellungen für die Metrik oder Metrikgruppe zu aktivieren.
7. Wählen Sie **Benutzerdefinierte Einstellungen verwenden** aus, um die Einstellungen für die Metrik oder Metrikgruppe zu ändern.

Sie können eine Benachrichtigungs-E-Mail senden, wenn der Score in den Bereichen **Inakzeptabel**, **Akzeptabel** und **Gut** liegt und Schwellenwerte über- oder unterschreitet. Sie können auch eine Benachrichtigungs-E-Mail senden, wenn die Metrikkosten die Schwellenwerte über- oder unterschreiten.

8. Um die globalen Einstellungen für Scorecard-Benachrichtigungen zu bearbeiten, klicken Sie auf das Symbol **Globale Einstellungen bearbeiten**.

   Das Dialogfeld **Globale Einstellungen bearbeiten** wird eingeblendet. Dort können Sie die Einstellungen, einschließlich der E-Mail-Vorlage, bearbeiten.

Konfigurieren von globalen Einstellungen für Scorecard-Benachrichtigungen

Bei Auswahl der globalen Einstellungen für Scorecard-Benachrichtigungen sendet das Analyst-Tool E-Mails an Zielbenutzer, sobald der Score im Bereich **Inakzeptabel** liegt. Darüber hinaus können Sie Benachrichtigungseinstellungen zum Senden von E-Mails konfigurieren, wenn die Metrik-Scores oder Metrikkosten Schwellenwerte über- oder unterschreiten. Sie können die E-Mail-Vorlage, einschließlich der E-Mail-Adressen und des Nachrichtentextes für eine Scorecard, konfigurieren.

1. Führen Sie eine Scorecard im Analyst-Tool aus.
2. Klicken Sie auf **Aktionen** > **Bearbeiten** > **Benachrichtigungen**, um das Dialogfeld **Scorecard bearbeiten** zu öffnen.
3. Wählen Sie **Benachrichtigungen aktivieren** aus, um mit der Konfiguration von Scorecard-Benachrichtigungen zu beginnen.
4. Klicken Sie auf das Symbol **Globale Einstellungen bearbeiten**.

   Das Dialogfeld **Globale Einstellungen bearbeiten** wird eingeblendet. Dort können Sie die Einstellungen einschließlich der E-Mail-Vorlage bearbeiten.
5. Legen Sie über die Kontrollkästchen **Score rein** und **Score-Bewegungen** fest, wann E-Mail-Benachrichtigungen für Metrik-Scores gesendet werden sollen.
6. Verwenden Sie die Kontrollkästchen unter **Kostenabgang**, um den Zeitpunkt für das Senden von E-Mail-Benachrichtigungen für Metrikkosten festzulegen.
7. Geben Sie im Feld **E-Mail an** die E-Mail-ID des Empfängers ein. Verwenden Sie ein Semikolon, um mehrere E-Mail-IDs voneinander zu trennen.

   Die Standard-E-Mail-ID des Absenders ist die in den SMTP-Eigenschaften der Domäne konfigurierte **E-Mail-Adresse des Absenders**.
8. Geben Sie den Text für die Betreffzeile der E-Mail ein.
9. Geben Sie im Feld **Text** den Einleitungs- und Schlusstext der E-Mail-Nachricht ein.
10. Um die globalen Einstellungen anzuwenden, wählen Sie **Einstellungen auf alle Metriken und Metrikgruppen anwenden** aus.
11. Klicken Sie auf **OK**.

Scorecard-Herkunft

Das Scorecard-Herkunftsdiagramm für eine Metrik oder Metrikgruppe zeigt den Ursprung der Daten, beschreibt den Pfad und verdeutlicht den Datenfluss. Mithilfe der Scorecard-Herkunft können Sie die Ursache

einer inakzeptablen Score-Varianz bei Metriken oder Metrikgruppen analysieren. Die Scorecard-Herkunft können Sie im Analyst-Tool anzeigen lassen.

Führen Sie die folgenden Schritte aus, um die Scorecard-Herkunft anzuzeigen:

1. Verknüpfen Sie in Informatica Administrator einen Metadata Manager Service mit dem Analyst Service.
2. Wählen Sie ein Projekt und exportieren Sie die enthaltenen Scorecard-Objekte in einer XML-Datei. Nutzen Sie hierzu die Option „Ressourcendatei für Metadata Manager exportieren“ im Developer-Tool oder den Befehl „infacmd oie exportResources“.
3. Verwenden Sie die exportierte XML-Datei in Metadata Manager, um eine Ressource zu erstellen und zu laden.

   Hinweis: Der Name der Ressourcen-Datei, die Sie in Metadata Manager erstellen und laden, muss der folgenden Benennungskonvention entsprechen: <MRS-Name>_<Projektname>. Weitere Informationen zum Erstellen und Laden einer Ressourcendatei finden Sie im *Benutzerhandbuch für Informatica PowerCenter Metadata Manager*.
4. Öffnen Sie die Scorecard im Analyst-Tool und wählen Sie eine Metrik oder Metrikgruppe.
5. Lassen Sie sich die Scorecard-Herkunft anzeigen.

Anzeigen der Scorecard-Herkunft in Informatica Analyst

Sie können ein Scorecard-Herkunftsdiagramm für eine Metrik oder Metrikgruppe anzeigen lassen. Bevor Sie das Scorecard-Herkunftsdiagramm im Analyst-Tool anzeigen können, müssen Sie die Scorecard-Herkunft und Metadaten in Metadata Manager laden.

1. Klicken Sie im Arbeitsbereich **Bibliothek** auf die Scorecard, die Sie im Bereich **Objekte** anzeigen möchten.

   Die Scorecard wird im Arbeitsbereich **Scorecards** angezeigt.
2. Wählen Sie in der Ansicht **Scorecard** eine Metrik oder Metrikgruppe.
3. Wählen Sie per Rechtsklick die Option **Herkunft anzeigen**.

   Das Scorecard-Herkunftsdiagramm erscheint in einem neuen Fenster.

   Wichtig: Wenn Sie eine Ressource nicht in Metadata Manager mit einer exportierten XML-Datei der Scorecard-Objekte erstellen und laden, erscheint möglicherweise eine Fehlermeldung, die Sie darauf hinweist, dass die Ressource im Katalog nicht verfügbar ist. Weitere Informationen zum Exportieren einer XML-Datei für die Scorecard-Herkunft finden Sie unter “Exportieren einer Ressourcendatei zur Scorecard-Herkunftskontrolle” auf Seite 100.

Teil III: Profiling mit Informatica Developer

Dieser Teil enthält die folgenden Kapitel:

- Datenobjektprofile, 79
- Spaltenprofile in halbstrukturierten Datenquellen, 84
- Regeln in Informatica Developer, 91
- Spaltenprofilergebnisse in Informatica Developer, 93
- Scorecards in Informatica Developer, 99
- Mapplet- und Mapping-Profiling, 102

KAPITEL 8

Datenobjektprofile

Dieses Kapitel umfasst die folgenden Themen:

- Spaltenprofile in Informatica Developer, 79
- Erstellen eines einzelnen Datenobjektprofils Erstellen eines Profils, 81
- Erstellen mehrerer Datenobjektprofile, 82
- Synchronisieren eines Einfachdatei-Datenobjekts, 82
- Synchronisieren eines nicht-relationalen Datenobjekts, 83

Spaltenprofile in Informatica Developer

Verwenden Sie ein Spaltenprofil, um die Eigenschaften von Spalten in einer Datenquelle, wie z. B. Wertprozentsätze und Wertmuster, zu analysieren. Um zu bestimmen, welche Zeilen das Profil während der Laufzeit liest, können Sie Filter hinzufügen. Zeilen, die die Filterkriterien nicht erfüllen, werden vom Profil nicht bearbeitet.

Sie können folgende Typen von Informationen über die Spalten, für die Sie ein Profil ausführen, erkennen:

- Wie häufig ein Wert in einer Spalte erscheint.
- Häufigkeit des Auftretens eines Werts in einer Spalte, ausgedrückt als Prozentsatz oder Anzahl von Zeilen.
- Zeichenmuster der Werte in einer Spalte.
- Statistiken, wie die maximale und minimale Länge der Werte in einer Spalte, und erster und letzter Wert.
- Abgeleitete Datentypen, Häufigkeit, Konformitätskriterien für die Datendomänenerkennung sowie Status der Datentypableitung.

Sie können ein Spaltenprofil für ein Datenobjekt in einer Zuordnung oder einem Mapplet oder einem Objekt im Modellrepository definieren. Das Objekt im Repository kann in einem einzelnen Datenobjektprofil, mehreren Datenobjektprofilen oder einem Enterprise-Erkennungsprofil vorhanden sein.

Sie können Stichprobenoptionen, Drilldown-Optionen und eine Laufzeitumgebung für ein Spaltenprofil auswählen. Sie können Regeln und Filter zu einem Spaltenprofil hinzufügen.

Filteroptionen

Sie können Filter hinzufügen, um die Zeilen zu bestimmen, die ein Spaltenprofil für Profilingoperationen verwendet. Zeilen, die die Filterkriterien nicht erfüllen, werden vom Profil nicht bearbeitet.

1. Erstellen oder öffnen Sie ein Spaltenprofil.

2. Wählen Sie die Ansicht **Filter**.
3. Klicken Sie auf **Hinzufügen**.
4. Wählen Sie einen Filtertyp aus, und klicken Sie auf **Weiter**.
5. Geben Sie einen Namen für den Filter ein. Geben Sie optional einen beschreibenden Text für den Filter ein.
6. Wählen Sie **Als aktiv festlegen**, um den Filter auf das Profil anzuwenden. Klicken Sie auf **Weiter**.
7. Definieren Sie die Filterkriterien.
8. Klicken Sie auf **Fertigstellen**.

Stichprobenoptionen

Konfigurieren Sie die Stichprobenoptionen, um die Anzahl der Zeilen festzulegen, die vom Profil während eines Profiling-Vorgangs gelesen werden.

In der folgenden Tabelle werden die Stichprobenoptionen beschrieben:

Eigenschaft	Beschreibung
Alle Zeilen	Wählt alle Zeilen im Datenobjekt aus.
Erste	Die Anzahl der Zeilen, für die Sie das Profil ausführen möchten. Das Developer Tool wählt die Zeilen aus den ersten Zeilen in der Quelle aus.
Zufällige Stichprobe von	Die Anzahl der Zeilen für eine zufällige Stichprobe, für die das Profil ausgeführt werden soll.
Zufallsstichprobe (auto)	Größe der zufälligen Stichprobe basierend auf der Anzahl der Zeilen im Datenobjekt.
Genehmigte Datentypen und Datendomänen in den nachfolgenden Profilausführungen aus der Datentyp- und Datendomänenableitung ausschließen	Schließt den genehmigten Datentyp oder die genehmigte Datendomäne bei der nächsten Profilausführung aus der Datentyp- und Datendomänenableitung aus.

Eigenschaft	Beschreibung
Alle Zeilen	Liest alle Zeilen aus der Quelle. Standardwert ist „Aktiviert“.
Erste	Liest von der ersten bis zur angegebenen Zeile.

Erstellen eines einzelnen Datenobjektprofils Erstellen eines Profils

Sie können für eine oder mehrere Spalten in einem Datenobjekt ein einzelnes Datenobjektprofil erstellen und das Profilobjekt im Modellrepository speichern. Sie können ein Profil für eine oder mehrere Spalten in einem Datenobjekt erstellen und das Profilobjekt im Modellrepository speichern.

1. Wählen Sie in der Ansicht **Objekt-Explorer** das Datenobjekt aus, das Sie profilieren möchten.
2. Klicken Sie zum Öffnen des Profilassistenten auf **Datei** > **Neu** > **Profil**.
3. Wählen Sie **Profil**, und klicken Sie auf **Weiter**.
4. Geben Sie einen Namen für das Profil ein, und überprüfen Sie den Projektspeicherort. Suchen Sie bei Bedarf nach einem neuen Speicherort.
5. Geben Sie optional eine Textbeschreibung des Profils ein.
6. Überprüfen Sie, dass der Name des von Ihnen ausgewählten Datenobjekts im Abschnitt **Datenobjekte** erscheint.
7. Klicken Sie auf **Weiter**.
8. Konfigurieren Sie die Profilvorgänge, die Sie ausführen möchten. Sie können folgende Vorgänge konfigurieren:
   - Spalten-Profiling
   - Primärschlüsselerkennung
   - Funktionale Abhängigkeitserkennung
   - Datendomänenerkennung

   Hinweis: Wählen Sie zum Aktivieren einer Profiloperation **Als Teil der Aktion "Profil ausführen" aktiviert** für die betreffende Operation. Standardmäßig ist Spalten-Profiling aktiviert.
9. Konfigurieren Sie die Spaltenprofiloptionen, z. B. Spaltenauswahl, Filter, Stichproben- und Drilldown-Optionen.
10. Prüfen Sie die Optionen Ihres Profils.

    Sie können die Spaltenauswahl für alle Profiltypen ändern. Überprüfen Sie die Filter- und Stichprobenoptionen für die Spaltenprofile. Sie können die Interferenzoptionen für den Primärschlüssel, die funktionale Abhängigkeit und die Datendomänenerkennung überprüfen. Sie können ebenfalls die Datendomänenauswahl für die Datendomänenerkennung überprüfen.
11. Überprüfen Sie die Drilldown-Optionen und bearbeiten Sie sie gegebenenfalls. Standardmäßig ist die Option **Zeilen-Drilldown aktivieren** ausgewählt. Sie können Drilldown-Optionen für Spaltenprofile bearbeiten. Die Optionen bestimmen auch, ob Drilldown-Vorgänge aus der Datenquelle oder aus Staging-Daten gelesen werden sollen und ob das Profil Ergebnisdaten aus vorherigen Profildurchläufen speichert.
12. Konfigurieren Sie im Abschnitt **Einstellungen ausführen** die Validierungs- und Laufzeiteinstellungen. Sie können die Verbindung **Nativ**, **Hive** oder **Hadoop** als Laufzeitumgebung auswählen.
13. Klicken Sie auf **Fertig stellen**.

Erstellen mehrerer Datenobjektprofile

Bei mehreren Datenobjekt-Profiling-Operationen werden zum Generieren von Spaltenprofilen für ein oder mehrere Datenobjekte Standardspalten-Profilingoptionen verwendet.

1. Wählen Sie in der Ansicht **Objekt-Explorer** das Datenobjekt aus, das Sie profilieren möchten.
2. Klicken Sie auf **Datei** > **Neu** > **Profil** um den Assistenten **Neues Profil** zu öffnen.
3. Wählen Sie die Option **Mehrfachprofile**, und klicken Sie auf **Weiter**.
4. Wählen Sie den Speicherort, an dem Sie die Profile erstellen möchten. Sie können jedes Profil an demselben Speicherort wie dessen profiliertes Objekt erstellen oder einen gemeinsamen Speicherort für die Profile angeben.
5. Überprüfen Sie, dass die Namen der von Ihnen ausgewählten Datenobjekte im Abschnitt **Datenobjekte** erscheinen.

   Optional können Sie auch auf **Hinzufügen** klicken, um ein anderes Datenobjekt hinzuzufügen.
6. Geben Sie optional die Anzahl der zu profilierenden Zeilen an, und wählen Sie, ob das Profil ausgeführt werden soll, wenn der Assistent beendet ist.
7. Klicken Sie auf **Fertigstellen**.
8. Geben Sie optional die Präfix- und Suffix-Strings ein, die den Profilnamen hinzugefügt werden sollen.
9. Klicken Sie auf **OK**.

Synchronisieren eines Einfachdatei-Datenobjekts

Bei Informatica Developer können Sie die Änderungen an einer externen Einfachdatei-Datenquelle mit deren Datenobjekt synchronisieren. Zum Synchronisieren der Datenobjekte verwenden Sie den Assistenten zum **Synchronisieren von Einfachdateien**.

1. Wählen Sie in der Ansicht **Objekt-Explorer** ein relationales Einfachdatei-Datenobjekt aus.
2. Klicken Sie mit der rechten Maustaste, und wählen Sie **Synchronisieren**.

   Der Assistent zum **Synchronisieren eines Einfachdatei-Datenobjekts** wird angezeigt.
3. Überprüfen Sie den Dateipfad im Feld **Vorhandene Einfachdatei wählen**.
4. Klicken Sie auf **Weiter**.
5. Wählen Sie optional die Codeseite, das Format, die Eigenschaften für das Trennzeichenformat und die Spalteneigenschaften.
6. Klicken Sie auf **Fertig stellen** und anschließend auf **OK**.

Synchronisieren eines nicht-relationalen Datenobjekts

Im Informatica Developer können Sie externe Änderungen einer relationalen Datenquelle mit deren Datenobjekt synchronisieren. Externe Datenquelländerungen sind beispielsweise das Hinzufügen, Ändern und Entfernen von Spalten sowie Änderungen an Regeln.

1. Wählen Sie in der Ansicht **Objekt-Explorer** ein relationales Datenobjekt aus.
2. Klicken Sie mit der rechten Maustaste, und wählen Sie **Synchronisieren**.

   Sie werden aufgefordert, die Aktion zu bestätigen.
3. Um den Synchronisierungsprozess abzuschließen, klicken Sie auf **OK**. Klicken Sie auf **Abbrechen**, um den Vorgang abzubrechen.

   Wenn Sie auf **OK** klicken, wird eine Statusmeldung für den Synchronisierungsprozess angezeigt.
4. Wenn Sie die Meldung **Synchronisierung beendet** sehen, klicken Sie auf **OK**.

   In Verbindung mit der Meldung wird eine Zusammenfassung der an dem Datenobjekt vorgenommenen Metadatenänderungen angezeigt.

KAPITEL 9

Spaltenprofile in halbstrukturierten Datenquellen

Dieses Kapitel umfasst die folgenden Themen:

- Spaltenprofile in halbstrukturierten Datenquellen - Übersicht, 84
- Spaltenprofil in einer JSON- oder XML-Datenquelle, 85
- Spaltenprofil in einer halbstrukturierten Datenquelle in HDFS, 86
- Erstellen einer HDFS-Verbindung, 87
- Erstellen eines komplexen Dateidatenobjekts in einer JSON- oder XML-Datei in HDFS, 87
- Erstellen eines Datenobjekts in einer Avro- oder Parquet-Datenquelle, 88
- Ausführen eines Spaltenprofils in halbstrukturierten Datenquellen, 89

Spaltenprofile in halbstrukturierten Datenquellen - Übersicht

Sie können Datenobjekte aus Avro-, JSON-, Parquet- und XML-Datenquellen erstellen und dann ein Spaltenprofil für die Datenobjekte anlegen.

Bei den Avro-, JSON-, Parquet- und XML-Formaten handelt es sich um halbstrukturierte Datenquellen. Zum Erstellen eines Spaltenprofils mithilfe der halbstrukturierten Datenquellen können Sie folgende Aufgaben durchführen:

1. Erstellen eines physischen Datenobjekts in der halbstrukturierten Datenquelle.
2. Erstellen und Ausführen eines Spaltenprofils im physischen Datenobjekt.

Sie können Einfachdatei-Datenobjekte für JSON- oder XML-Datenquellen erstellen. Sie können komplexe Dateidatenobjekte für Avro-, JSON-, Parquet- und XML-Datenquellen in HDFS (Hadoop Distributed File System) erstellen.

Spaltenprofil in einer JSON- oder XML-Datenquelle

Sie können ein Einfachdatei-Datenobjekt oder ein komplexes Dateidatenobjekt aus einer JSON- oder XML-Datenquelle erstellen. Sie können ein Spaltenprofil im Datenobjekt erstellen und ausführen.

Erstellen Sie eine Textdatei, die den Pfad der JSON- oder XML-Datenquelle enthält, und verwenden Sie die Textdatei als Datenquelle zum Erstellen eines Einfachdatei-Datenobjekts. Sie können auch den Dateipfad für mehrere JSON- oder XML-Datenquellen in die Textdatei einfügen.

Sie können ein komplexes Dateidatenobjekt mithilfe eines komplexen Datei-Readers aus einer JSON- oder XML-Datenquelle erstellen. Der komplexe Datei-Reader stellt Eingabe für eine Datenprozessorumwandlung bereit, die die Datei parst und die Datenquelle in einfache Datensätze mit kommagetrennten Werten konvertiert.

Hinweis: Das Developer Tool bietet keine Unterstützung für eine JSON-Datenquelle mit UTF-8-Codierung.

Erstellen eines Datenobjekts in einer JSON- oder XML-Datenquelle

Sie können ein Einfachdatei-Datenobjekt oder ein komplexes Dateidatenobjekt in einer JSON- oder XML-Datenquelle erstellen.

1. Wählen Sie im Developer Tool in der Ansicht **Object Explorer** das Projekt aus, in dem Sie das Datenobjekt und das Spaltenprofil erstellen möchten.
2. Klicken Sie auf **Datei** > **Neu** > **Datenobjekt**.

   Das Dialogfeld **Neu** wird eingeblendet.
3. Sie können ein Einfachdatei-Datenobjekt und ein komplexes Dateidatenobjekt erstellen.
   - Führen Sie zum Erstellen eines Einfachdatei-Datenobjekts die folgenden Aufgaben durch:
     1. Wählen Sie **Physische Datenobjekte** > **Einfachdatei-Datenobjekt** aus und klicken Sie auf **Weiter**. Das Dialogfeld **Neues Einfachdatei-Datenobjekt** wird eingeblendet.
     2. Wählen Sie **Aus vorhandener Einfachdatei erstellen** aus und klicken Sie auf **Durchsuchen**, um die Textdatei auszuwählen. Klicken Sie auf **Weiter**.
     3. Stellen Sie sicher, dass die Codepage **MS Windows Latin 1 (ANSI), eine Obermenge von Latin 1** und das delimitergetrennte Format verwendet werden. Klicken Sie auf **Weiter**.
     4. Stellen Sie sicher, dass der Delimiter auf **Komma** gesetzt ist. Klicken Sie auf **Fertig stellen**.
   - Führen Sie zum Erstellen eines komplexen Dateidatenobjekts die folgenden Aufgaben durch:
     1. Wählen Sie **Physische Datenobjekte** > **Komplexes Dateidatenobjekt** aus und klicken Sie auf **Weiter**.
        Das Dialogfeld **Neues komplexes Dateidatenobjekt** wird eingeblendet.
     2. Geben Sie einen Namen für das Datenobjekt ein. Wählen Sie **Datei** als Zugriffstyp aus.
     3. Klicken Sie auf **Durchsuchen**, um eine JSON- oder XML-Datei auszuwählen. Klicken Sie auf **Fertigstellen**.
        Wenn der Developer-Server unter Linux ausgeführt wird, müssen Sie den Dateipfad der Datenquelle zum Speicherort auf dem Server aktualisieren. Wählen Sie zum Aktualisieren des Dateipfads das komplexe Datei-Datenobjekt aus, klicken Sie auf der Registerkarte **Datenobjekt-Vorgänge** auf **Lesen** und fügen Sie den Dateipfad auf der Registerkarte **Erweitert** im Bereich **Details zu Datenobjekt-Vorgang** hinzu.

   Das Datenobjekt wird im Projektordner angezeigt.

Spaltenprofil in einer halbstrukturierten Datenquelle in HDFS

Sie können ein Spaltenprofil in einer Avro-, JSON-, Parquet- oder XML-Datei, die HDFS verwendet, erstellen und ausführen. Verwenden Sie zum Lesen der JSON- oder XML-Datei in HDFS einen komplexen Datei-Reader, um die JSON- oder XML-Eingabe an die Datenprozessorumwandlung zu übergeben.

Spaltenprofil in einer JSON- oder XML-Datenquelle in HDFS

Sie können ein komplexes Dateidatenobjekt aus einer JSON- oder XML-Datei erstellen. Sie können ein Spaltenprofil im Datenobjekt erstellen und ausführen.

Erstellen Sie eine Verbindung mit HDFS, bevor Sie die Datenobjekte für JSON- oder XML-Dateien in HDFS anlegen.

Sie können eine der folgenden Methoden verwenden, um ein Datenobjekt aus einer JSON- oder XML-Datei in HDFS zu erstellen:

- Erstellen Sie ein komplexes Dateidatenobjekt in einer JSON- oder XML-Datei.
- Erstellen Sie ein komplexes Dateidatenobjekt in einem Ordner, der mehrere JSON- oder XML-Dateien enthält.

Nach der Erstellung des Datenobjekts können Sie ein Spaltenprofil im Datenobjekt erstellen und ausführen.

Spaltenprofil in einer Avro- oder Parquet-Datenquelle in HDFS

Sie können ein komplexes Dateidatenobjekt aus einer Avro- oder Parquet-Datenquelle in HDFS erstellen. Sie können das Datenobjekt verwenden, um ein Spaltenprofil zu erstellen und auszuführen.

Sie können ein komplexes Dateidatenobjekt in einer Avro- oder Parquet-Datei bzw. einem Ordner erstellen, der mehrere Avro- oder Parquet-Dateien enthält. Sie können ein komplexes Dateidatenobjekt in einer Avro- oder Parquet-Datenquelle mit dem Zugriffstyp „Datei“ oder „Verbindung“ und folgendem Ressourcenformat erstellen: Binär, Avro oder Parquet. Sie müssen eine HDFS-Verbindung erstellen, bevor Sie ein komplexes Dateidatenobjekt in den Avro- und Parquet-Datenquellen anlegen.

Hinweis: Sie können das Ressourcenformat **Avro** oder **Parquet** nur für Avro- und Parquet-Datenquellen mit flacher Struktur auswählen.

Sie können eine der folgenden Optionen auswählen, wenn Sie ein Datenobjekt aus Avro- und Parquet-Dateien in HDFS erstellen:

- Wählen Sie „Datei“ als Zugriffstyp und „Binär“ als Ressourcenformat aus.
- Wählen Sie „Datei“ als Zugriffstyp und „Avro“ oder „Parquet“ als Ressourcenformat aus.
- Wählen Sie „Verbindung“ als Zugriffstyp und „Avro“ oder „Parquet“ als Ressourcenformat aus.

Erstellen einer HDFS-Verbindung

Konfigurieren Sie die HDFS-Verbindung in Informatica Developer, um ein Spaltenprofil in Avro-, JSON-, Parquet- und XML-Datenquellen in HDFS zu erstellen. Sie können ein komplexes Dateidatenobjekt erstellen, nachdem Sie eine HDFS-Verbindung eingerichtet haben.

1. Klicken Sie auf **Fenster** > **Einstellungen**.
2. Wählen Sie **Informatica** > **Verbindungen** aus.
3. Erweitern Sie die Domäne.
4. Wählen Sie den Verbindungstyp **Dateisysteme** > **Hadoop-Dateisystem** aus und klicken Sie auf **Hinzufügen**.
5. Geben Sie einen Verbindungsnamen ein.
6. Geben Sie optional eine Beschreibung der Verbindung ein.
7. Klicken Sie auf **Weiter**.
8. Konfigurieren Sie die Verbindungseigenschaften.
9. Klicken Sie auf **Verbindung testen**, um die Verbindung zu HDFS zu überprüfen.
10. Klicken Sie auf **Fertigstellen**.

Erstellen eines komplexen Dateidatenobjekts in einer JSON- oder XML-Datei in HDFS

Sie können ein komplexes physisches Datenobjekt für die JSON- oder XML-Dateiquelle erstellen, die HDFS verwendet, und ein Spaltenprofil im physischen Datenobjekt anlegen.

1. Wählen Sie im Developer Tool in der Ansicht **Object Explorer** das Projekt aus, in dem Sie ein physisches Datenobjekt und ein Spaltenprofil erstellen möchten.
2. Klicken Sie auf **Datei** > **Neu** > **Datenobjekt**.

   Das Dialogfeld **Neu** wird eingeblendet.
3. Wählen Sie **Physische Datenobjekte** > **Komplexes Dateidatenobjekt** aus und klicken Sie auf **Weiter**.

   Das Dialogfeld **Neues komplexes Dateidatenobjekt** wird eingeblendet.
4. Geben Sie einen Namen für das Datenobjekt ein. Wählen Sie **Verbindung** als Zugriffstyp aus.
5. Sie können ein Datenobjekt in einer JSON- oder XML-Datei oder in einem Ordner erstellen, der mehrere JSON- und XML-Dateien enthält.
   - Führen Sie die folgenden Aufgaben durch, um mit der Erstellung eines komplexen Dateidatenobjekts in einer JSON- oder XML-Datei fortzufahren:
     1. Klicken Sie auf **Durchsuchen**, um eine Verbindung auszuwählen. Klicken Sie im Dialogfeld **Ressource hinzufügen** auf **Hinzufügen**, um eine JSON- oder XML-Datei auszuwählen. Klicken Sie auf **Fertig stellen**.
     2. Das Datenobjekt wird im Projektordner angezeigt.

- Führen Sie die folgenden Aufgaben durch, um mit der Erstellung eines komplexen Dateidatenobjekts in einem Ordner fortzufahren, der mehrere JSON- oder XML-Dateien enthält:
  1. Klicken Sie auf **Durchsuchen**, um eine Verbindung auszuwählen. Klicken Sie im Dialogfeld **Ressource hinzufügen** auf **Hinzufügen**, um eine JSON- oder XML-Datei im Ordner auszuwählen. Klicken Sie auf **Fertig stellen**.
  2. Das Datenobjekt wird im Projektordner angezeigt.
  3. Wählen Sie das Datenobjekt im Projektordner aus und klicken Sie auf **Erweitert** > **Laufzeit: Lesen** > **Quelldateiverzeichnis**.
  4. Entfernen Sie den Dateinamen und behalten Sie den Ordnernamen im Dateipfad bei.

Erstellen eines Datenobjekts in einer Avro- oder Parquet-Datenquelle

Sie können ein komplexes physisches Datenobjekt für eine Avro- oder Parquet-Datenquelle mit **Datei** oder **Verbindung** als Zugriffstyp auswählen. Sie können ein Spaltenprofil im physischen Datenobjekt erstellen.

1. Wählen Sie in der Ansicht **Objekt-Explorer** ein Projekt aus.
2. Klicken Sie auf **Datei** > **Neu** > **Datenobjekt**.

   Das Dialogfeld **Neu** wird eingeblendet.
3. Wählen Sie **Physische Datenobjekte** > **Komplexes Dateidatenobjekt** aus und klicken Sie auf **Weiter**.

   Das Dialogfeld **Neues komplexes Dateidatenobjekt** wird eingeblendet.
4. Geben Sie einen Namen für das Datenobjekt ein.
5. Als Zugriffstyp können Sie **Verbindung** oder **Datei** auswählen.
   - Wenn Sie **Verbindung** als Zugriffstyp auswählen, führen Sie folgende Aufgaben durch:
     1. Klicken Sie auf **Durchsuchen**, um eine HDFS-Verbindung auszuwählen.
        Das Dialogfeld **Verbindung auswählen** wird geöffnet.
     2. Wählen Sie im Dialogfeld **Verbindung auswählen** eine Datenquelle aus und klicken Sie auf **OK**.
     3. Klicken Sie im Dialogfeld **Neues komplexes Dateidatenobjekt** auf **Fertigstellen**.
        Das Datenobjekt wird im Projektordner angezeigt.
   - Wenn Sie **Datei** als Zugriffstyp und **Binär** als Ressourcenformat auswählen, führen Sie die folgenden Aufgaben durch:
     1. Klicken Sie auf **Durchsuchen**, um eine Avro- oder Parquet-Datei auf dem lokalen Computer auszuwählen.
     2. Klicken Sie im Dialogfeld **Neues komplexes Dateidatenobjekt** auf **Fertigstellen**.
        Das Datenobjekt wird im Projektordner angezeigt.
     3. Wählen Sie das Datenobjekt im Projektordner aus und klicken Sie auf die Ansicht **Datenobjektvorgänge**.
     4. Klicken Sie in der Ansicht **Datenobjektvorgänge** auf **Lesen** > **Erweitert** (Registerkarte).
     5. Geben Sie auf der Registerkarte **Erweitert** im Feld **Dateipfad** den Dateipfad der Datenquelle auf dem Linux- oder Windows-Computer ein.
     6. Geben Sie **Benutzerdefinierte Eingabe** als Dateiformat ein.

7. Geben Sie **com.informatica.avro.AvroToXML** im Feld **Eingabeformat** für Avro-Datenquellen und **com.informatica.parquet.ParquetToXML** im Feld **Eingabeformat** für Parquet-Datenquellen ein. Wenn Sie das Eingabeformat hinzufügen, verarbeitet die Datenprozessorumwandlung die Datenquellen im Avro- oder Parquet-Format und wandelt sie zur Laufzeit in Datenquellen im XML-Format um.

- Wenn Sie **Datei** als Zugriffstyp und **Avro** oder **Parquet** als Ressourcenformat auswählen, führen Sie die folgenden Aufgaben durch:

  1. Klicken Sie auf **Durchsuchen**, um eine Avro- oder Parquet-Datei auf dem lokalen Computer auszuwählen.
  2. Klicken Sie im Dialogfeld **Neues komplexes Dateidatenobjekt** auf **Fertigstellen**.
     Das Datenobjekt wird im Projektordner angezeigt.
  3. Navigieren Sie nach dem Erstellen des Datenobjekts zu **Datenobjektvorgänge** > **Lesen** > **Erweitert** (Registerkarte) und stellen Sie sicher, dass der Dateipfad im Feld **Dateipfad** der Datenquelle auf dem Linux- oder Windows-Computer entspricht.

  Hinweis: Sie können das Ressourcenformat **Avro** oder **Parquet** nur für Avro- und Parquet-Datenquellen mit flacher Struktur auswählen.

  Sie können einen Ordner mit mehreren Avro- oder Parquet-Dateien auswählen, um ein Datenobjekt zu erstellen. Navigieren Sie nach dem Erstellen des Datenobjekts zu **Datenobjektvorgänge** > **Lesen** > **Erweitert** (Registerkarte) und stellen Sie sicher, dass der Dateipfad im Feld **Dateipfad** dem Ordner der Datenquellen auf dem Linux- oder Windows-Computer entspricht.

Ausführen eines Spaltenprofils in halbstrukturierten Datenquellen

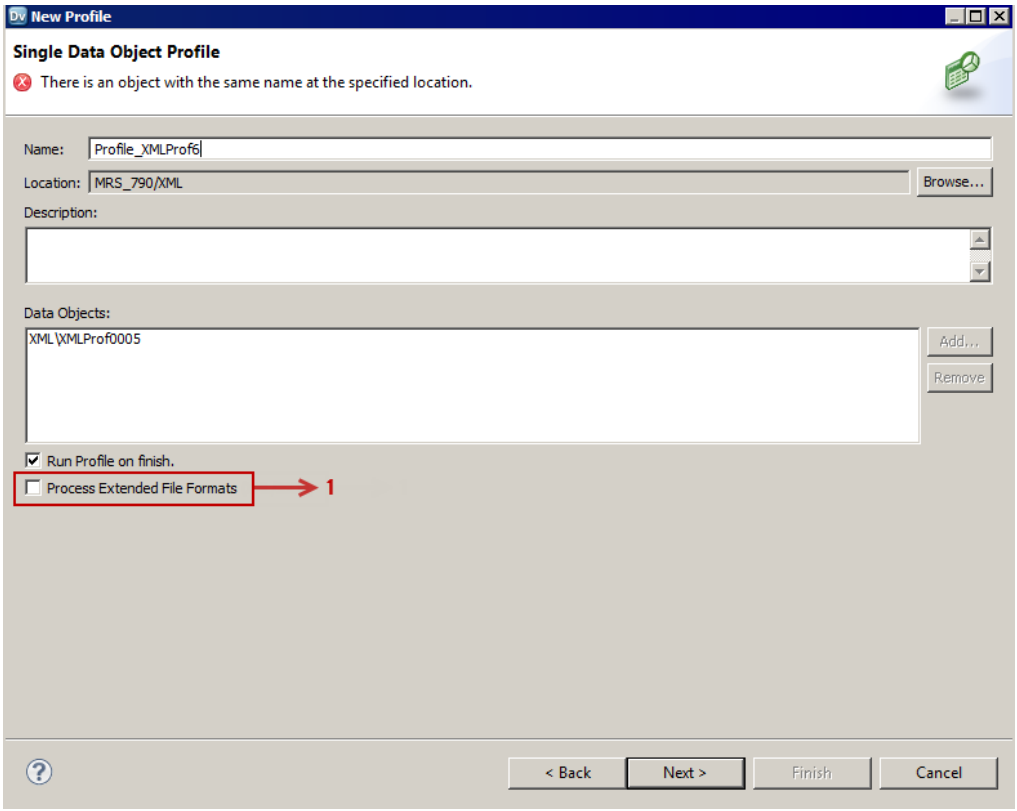
Nach der Erstellung eines Einfachdatei-Datenobjekts oder eines komplexen Dateidatenobjekts aus Avro-, JSON-, Parquet- oder XML-Datenquellen können Sie ein Spaltenprofil im Datenobjekt erstellen und ausführen.

1. Wählen Sie in der Ansicht **Objekt-Explorer** das physische Datenobjekt für die Avro-, JSON-, Parquet- oder XML-Datei aus.
2. Klicken Sie auf **Datei** > **Neu** > **Profil**.

   Das Dialogfeld **Neu** wird eingeblendet.
3. Wählen Sie **Profil** aus. Klicken Sie auf **Weiter**.

   Das Dialogfeld **Neues Profil** wird geöffnet.
4. Fügen Sie im Dialogfeld **Neues Profil** einen Namen für das Profil und eine optionale Beschreibung hinzu.
5. Wählen Sie die Option **Erweiterte Dateiformate verarbeiten** aus. Klicken Sie auf **Weiter**.

   Die folgende Abbildung zeigt den Assistenten **Neues Profil** mit der Option **Erweiterte Dateiformate verarbeiten**:



1. Erweiterte Dateiformate verarbeiten. Wählen Sie diese Option aus, um Datenquellen im Avro-, JSON, Parquet- oder XML-Format zu verarbeiten.

Hinweis: Die Option **Erweiterte Dateiformate verarbeiten** wird für Avro- und Parquet-Datenquellen nicht angezeigt, wenn Sie **Avro** oder **Parquet** als Ressourcenformat ausgewählt haben.

6. Wählen Sie auf der Seite **Einzelnes Datenobjektprofil** die Spalten und Optionen unter **Spaltenauswahl** und **Datendomänenerkennung** nach Bedarf aus. Klicken Sie auf **Fertig stellen**.

   Hinweis: Wenn das Developer Tool auf einem Linux-Computer installiert wurde und es sich bei dem physischen JSON- oder XML-Datenobjekt um ein Einfachdatei-Datenobjekt mit einer Textdatei handelt, führen Sie die folgenden Aufgaben durch:

   1. Aktualisieren Sie auf der Registerkarte **Übersicht** den Wert **Gesamtstellenanzahl**, um die Anzahl der Zeichen im Dateipfad der Datenquelle auf dem Server zu berücksichtigen.
   2. den Dateipfad der Datenquelle zum Speicherort auf dem Server aktualisieren, nachdem Sie ein Profil für das Einfachdateidatenobjekt erstellt haben. Klicken Sie zum Aktualisieren des Dateipfads auf der Registerkarte **Erweitert** auf **Laufzeit: Lesen** > **Quelldateiverzeichnis** und fügen Sie den Dateipfad hinzu.

7. Klicken Sie mit der rechten Maustaste auf das Profil und wählen Sie **Profil ausführen** aus.

   Die Profilergebnisse werden angezeigt.

KAPITEL 10

Regeln in Informatica Developer

Dieses Kapitel umfasst die folgenden Themen:

- Regeln in Informatica Developer - Übersicht Richtlinien für Regeln, 91
- Erstellen einer Regel in Informatica Developer, 92
- Anwenden einer Regel in Informatica Developer Eine Regel anwenden, 92

Regeln in Informatica Developer - Übersicht Richtlinien für Regeln

Eine Regel ist eine Geschäftslogik, die Bedingungen definiert, welche auf Quelldaten angewendet werden, wenn ein Profil ausgeführt wird. Sie können im Developer-Tool wiederverwendbare Regeln aus Mapplets erstellen. Sie können diese Regeln in Analyst-Tool-Profilen wiederverwenden, um Quelldaten zu validieren. Sie können wiederverwendbare Regeln aus Mapplets im Developer-Tool erstellen. Sie können Sie diese Regeln in Profilen verwenden, um Quelldaten zu validieren.

Erstellen Sie ein Mapplet, und validieren Sie es als Regel. Diese Regel erscheint im Analysetool als wiederverwendbare Regel. Sie können die Regel im Developer-Tool oder Analysetool auf ein Spaltenprofil anwenden.

Eine Regel muss folgende Anforderungen erfüllen:

- Sie muss eine Eingabe- und eine Ausgabe-Umwandlung enthalten. In einer Regel können Sie keine Datenquellen verwenden.
- Sie kann Expressions-Umwandlungen, Lookup-Umwandlungen und passive Datenqualitäts-Umwandlungen enthalten. Andere Umwandlungstypen darf sie nicht enthalten. So darf eine Regel beispielsweise keine Match-Umwandlung enthalten, da es sich hierbei um eine aktive Umwandlung handelt.
- Sie gibt keine Kardinalität zwischen Eingabegruppen an.

Eine Regel muss folgende Anforderungen erfüllen:

- Sie muss eine Eingabe- und eine Ausgabe-Umwandlung enthalten. In einer Regel können Sie keine Datenquellen verwenden.
- Sie kann Ausdrucksumwandlungen, Lookup-Umwandlungen und passive Umwandlungen enthalten. Andere Umwandlungstypen darf sie nicht enthalten. So darf eine Regel beispielsweise keine Sortierer-Umwandlung enthalten, da es sich hierbei um eine aktive Umwandlung handelt.
- Sie gibt keine Kardinalität zwischen Eingabegruppen an.

Erstellen einer Regel in Informatica Developer

Sie müssen ein Mapplet als Regel validieren, um im Developer-Tool eine Regel zu erstellen

Erstellen eines Mapplets im Developer-Tool.

1. Klicken Sie mit der rechten Maustaste auf den Mapplet-Editor.
2. Wählen Sie **Validieren als** > **Regel**.

Anwenden einer Regel in Informatica Developer Eine Regel anwenden

Sie können eine Regel zu einem gespeicherten Spaltenprofil hinzufügen. Sie können einem Profil, das für die Join-Analyse konfiguriert wurde, keine Regel hinzufügen.Sie können eine Regel zu einem gespeicherten Spaltenprofil hinzufügen.

1. Durchsuchen Sie die Ansicht **Objekt-Explorer** nach dem benötigten Profil.
2. Klicken Sie mit der rechten Maustaste auf das Profil, und wählen Sie **Öffnen** aus.

   Das Profil wird im Editor geöffnet.
3. Klicken Sie auf die Registerkarte **Definition** und wählen Sie Regeln aus.
4. Klicken Sie auf **Hinzufügen**.

   Das Dialogfenster **Regel anwenden** wird geöffnet.
5. Klicken Sie auf **Durchsuchen**, um die Regel zu finden, die Sie anwenden möchten.

   Wählen Sie eine Regel aus einem Repository-Projekt, und klicken Sie auf **OK**.
6. Klicken Sie auf die Spalte **Wert** unter **Eingabewerte**, um einen Eingabeport für die Regel auszuwählen.
7. Optional klicken Sie auf die Spalte **Wert** unter **Ausgabewerte** zum Bearbeiten des Namens des Regelausgabeports.

   Die Regel wird nun auf der Registerkarte **Definition** angezeigt.

KAPITEL 11

Spaltenprofilergebnisse in Informatica Developer

Dieses Kapitel umfasst die folgenden Themen:

- Spaltenprofilergebnisse in Informatica Developer Spaltenprofilergebnisse, 93
- Spaltenwert-Eigenschaften, 94
- Spaltenmuster-Eigenschaften, 95
- Spaltenstatistik-Eigenschaften, 95
- Eigenschaften des Spaltendatentyps, 96
- Wiederherstellung im Informatica Developer Wiederherstellung in Informatica Developer, 97
- Exportieren von Profilergebnissen aus Informatica Developer, 98

Spaltenprofilergebnisse in Informatica Developer Spaltenprofilergebnisse

Die Spaltenprofilanalyse liefert Informationen über die Datenqualität, indem Werthäufigkeiten, Muster und Datenstatistiken hervorgehoben werden.

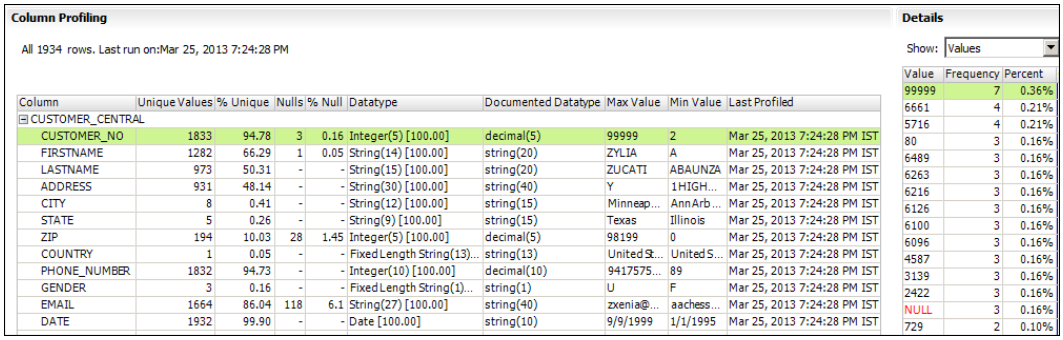
Die folgende Tabelle beschreibt die Profilergebnisse für jeden Analysetyp:

Die Spalten-Profiling-Analyse generiert die folgenden Profilergebnisse:

Profiltyp	Profilergebnisse
Spaltenprofil	- Prozentuale und Zählstatistiken für eindeutige und Nullwerte - Abgeleitete Datentypen - Der Datentyp, den die Datenquelle für die Daten deklariert - Die Höchst- und Mindestwerte - Datum und Uhrzeit des zuletzt ausgeführten Profildurchlaufs - Prozentuale und Zählstatistik für jedes eindeutige Datenelement in einer Spalte - Prozentuale und Zählstatistik für jedes eindeutige Zeichenmuster in einer Spalte
Primärschlüsselprofil	- Gefolgerte Primärschlüssel - Schlüssel-Verstöße
Funktionales Abhängigkeitsprofil	- Gefolgerte Funktionsabhängigkeiten - Funktionale Abhängigkeitsverstöße

- Prozentuale und Zählstatistiken für eindeutige und Nullwerte
- Abgeleitete Datentypen
- Der Datentyp, den die Datenquelle für die Daten deklariert
- Die Höchst- und Mindestwerte
- Datum und Uhrzeit des zuletzt ausgeführten Profildurchlaufs
- Prozentuale und Zählstatistik für jedes eindeutige Datenelement in einer Spalte
- Prozentuale und Zählstatistik für jedes eindeutige Zeichenmuster in einer Spalte

Die folgende Abbildung zeigt die Spaltenprofilergebnisse:



Spaltenwert-Eigenschaften

Die Spaltenwert-Eigenschaften zeigen die Werte in den profilierten Spalten und die Häufigkeit, mit der die einzelnen Werte in jeder Spalte auftauchen. Die Häufigkeiten werden als Zahlenwert, prozentual und als Balkendiagramm angezeigt.

Um die Spaltenwert-Eigenschaften anzuzeigen, wählen Sie „Werte“ aus der Liste **Anzeigen**. Doppelklicken Sie auf einen Spaltenwert, um zu den Zeilen zu gelangen, die den Wert enthalten.

Die folgende Tabelle beschreibt die Eigenschaften des Spaltenwertes:

Eigenschaft	Beschreibung
Werte	Liste aller Werte für die Spalte im Profil.
Häufigkeit	Häufigkeit, mit der ein Wert in einer Spalte vorkommt.
Prozent	Häufigkeit, mit der ein Wert in einer Spalte vorkommt, ausgedrückt als Prozentsatz aller Werte in der Spalte.
Diagramm	Balkendiagramm für prozentuale Angabe.

Spaltenmuster-Eigenschaften

Spaltenmuster-Eigenschaften zeigen die Muster der Daten in den profilierten Spalten und die Häufigkeit, mit der die Muster in jeder Spalte auftauchen. Die Muster werden in Form von Zahlen, prozentual und als Balkendiagramm dargestellt.

Um Musterinformationen anzuzeigen, wählen Sie „Muster“ aus der Liste **Anzeigen**. Doppelklicken Sie auf ein Muster, um auch die Zeilen anzuzeigen, die die Muster enthalten.

Die folgende Tabelle beschreibt die Eigenschaften für Spaltenwertmuster:

Eigenschaft	Beschreibung
Muster	Muster für die ausgewählte Spalte.
Häufigkeit	Gibt an, wie oft ein Muster in einer Spalte auftaucht.
Prozent	Gibt prozentual an, wie oft ein Muster in einer Spalte auftaucht.
Diagramm	Balkendiagramm für prozentuale Angabe.

Spaltenstatistik-Eigenschaften

Die Spaltenstatistik umfasst Eigenschaften wie die maximale und minimale Länge der Werte sowie die ersten und letzten Werte.

Um die Statistikinformationen anzuzeigen, wählen Sie „Statistik“ aus der Liste **Anzeigen**.

Die folgende Tabelle beschreibt die Spaltenstatistikeigenschaften:

Eigenschaft	Beschreibung
Maximale Länge	Länge des längsten Werts in der Spalte.
Minimale Länge	Länge des kürzesten Werts in der Spalte.
Unterster Wert	Letzte fünf Werte in der Spalte.
Oberster Wert	Erste fünf Werte in der Spalte.
Summe	Summe aller Werte in der Spalte mit einem numerischen Datentyp.

Hinweis: Das Profil enthält außerdem die Mittelwert- und Standardabweichungs-Statistik für Spalten des Typs Ganzzahl.

Eigenschaften des Spaltendatentyps

Zu den Spaltendatentypen gehören alle abgeleiteten Datentypen für jede Spalte in den Profilergebnissen.

Wählen Sie zur Anzeige von Datentypinformationen die Option **Datentypen** in der Liste **Anzeigen** aus. Doppelklicken Sie auf einen Datentyp, um einen Drilldown zu den Zeilen mit dem entsprechenden Datentyp durchzuführen.

In der folgenden Tabelle werden die Eigenschaften für die Spaltendatentypen beschrieben:

Eigenschaft	Beschreibung
Datentyp	Liste aller abgeleiteten Datentypen für die Spalte im Profil.
Häufigkeit	Zeigt an, wie oft ein Datentyp für eine Spalte angezeigt wird, ausgedrückt als Zahl.
% Konformität	Anteiliger Prozentsatz eines Datentyps pro Spalte.
Status	Gibt den Status des Datentyps an. Die Statusangaben lauten: „Abgeleitet“, „Genehmigt“ oder „Abgelehnt“. **Abgeleitet** Gibt den Datentyp der Spalte an, der vom Developer Tool abgeleitet wurde. **Genehmigt** Gibt einen genehmigten Datentyp für die Spalte an. Wenn Sie einen Datentyp genehmigen, binden Sie den Datentyp an das Modellrepository. **Abgelehnt** Gibt einen abgelehnten Datentyp für die Spalte an.

Wiederherstellung im Informatica Developer Wiederherstellung in Informatica Developer

Wiederherstellung ist der Prozess des Validierens und Verwaltens erkannter Metadaten einer Datenquelle, um die Metadaten für ihre Verwendung und das Reporting anzupassen. Wenn Sie Metadaten in Informatica Developer wiederherstellen, können Sie die abgeleiteten Datentypen oder Datendomänen in Profilergebnissen genehmigen, ablehnen und zurücksetzen.

Sie können einen Datentyp oder eine Datendomäne für eine Spalte genehmigen. Sie können die abgelehnten Datentypen und Datendomänen für eine Spalte ausblenden. Nachdem Sie einen abgeleiteten Datentyp oder eine abgeleitete Datendomäne genehmigt oder abgelehnt haben, können Sie den Datentyp oder die Datendomäne zurücksetzen, um den Ableitungsstatus wiederherzustellen.

Genehmigen von DatentypenGenehmigen von Datentypen in Informatica Developer

Die Profilergebnisse beinhalten die abgeleiteten Datentypen, die Häufigkeit, den Konformitätsprozentsatz und den Ableitungsstatus für jede Spalte in der Datenquelle. Sie können einen einzelnen Datentyp für jede Spalte auswählen und genehmigen.

1. Wählen und öffnen Sie ein Profil in der Ansicht **Objekt-Explorer**.
2. Stellen Sie sicher, dass Sie auf der Registerkarte **Ergebnisse** sind.
3. Wählen Sie in der Ansicht **Spalten-Profiling** im rechten Bereich eine Spalte zum Anzeigen der Werthäufigkeiten, Muster, Datentypen und Statistiken aus.
4. Wählen Sie im Bereich **Details** in der Liste **Anzeigen** die Option **Datentypen** aus.

   Die abgeleiteten Datentypen für die Spalte werden angezeigt.
5. Klicken Sie mit der rechten Maustaste auf die Spalte, die Sie genehmigen wollen, und klicken Sie auf **Genehmigen**.

   Der Status des Datentyps ändert sich in **Genehmigt**.
6. Klicken Sie zum Wiederherstellen des abgeleiteten Status des Datentyps mit der rechten Maustaste auf den Datentyp und klicken Sie auf **Zurücksetzen**.

Ablehnen von DatentypenAblehnen von Datentypen in Informatica Developer

Informatica Developer zeigt abgeleitete Datentypen standardmäßig in den Profilergebnissen an. Sie können abgeleitete oder genehmigte Datentypen ablehnen. Sie haben die Möglichkeit, die abgelehnten Datentypen ein- oder auszublenden.

1. Wählen Sie in der Ansicht **Objekt-Explorer** ein Profil aus.
2. Doppelklicken Sie auf das Profil, um es zu öffnen.

   Das Profil wird auf einer Registerkarte geöffnet.
3. Wählen Sie in der Ansicht **Spalten-Profiling** eine Zeile aus.
4. Wählen Sie zum Ablehnen abgeleiteter Spaltendatentypen die Ansicht **Datentypen** im rechten Bereich aus. Wählen Sie den abgeleiteten Datentyp aus, den Sie ablehnen möchten, klicken Sie mit der rechten Maustaste auf die Zeile und wählen Sie **Ablehnen** aus.

   Informatica Developer blendet den abgelehnten Datentyp in der Liste der Datentypen aus.

5. Klicken Sie zum Ausblenden der abgelehnten Datentypen mit der rechten Maustaste auf die Zeile und wählen Sie **Abgelehnte ausblenden** aus.
6. Klicken Sie zur Anzeige der abgelehnten Datentypen mit der rechten Maustaste auf eine der Zeilen und wählen Sie **Abgelehnte einblenden** aus.

Exportieren von Profilergebnissen aus Informatica Developer

Sie können Spaltenprofilergebnisse in eine CSV-Datei oder Microsoft Excel-Datei exportieren. Beim Exportieren der Profilergebnisse in eine Microsoft Excel-Datei speichert das Developer Tool die Informationen als XLSX-Datei.

1. Öffnen Sie in der Ansicht **Objekt-Explorer** ein Profil.
2. Optional führen Sie das Profil aus, um die Profilergebnisse zu aktualisieren.
3. Wählen Sie die Ansicht **Ergebnisse**.
4. Wählen Sie eine Spalte aus.
5. Wählen Sie unter **Details** die Optionen **Werte**, **Muster** oder **Datentypen** aus und klicken Sie auf das Symbol **Exportieren**.

   Das Dialogfenster **Daten in Datei exportieren** wird geöffnet.
6. Akzeptieren oder ändern Sie den Standarddateinamen.
7. Wählen Sie den Typ der zu exportierenden Daten. Sie können die Optionen **Werte für die ausgewählte Spalte**, **Muster für die ausgewählte Spalte**, **Datentypen für die ausgewählte Spalte** oder **Alle (Zusammenfassung, Werte, Muster, Datentypen, Statistiken, Eigenschaften)** auswählen.
8. Klicken Sie auf **Durchsuchen**, um einen Speicherort auszuwählen, und speichern Sie die Datei lokal auf Ihrem Computer.
9. Wenn Sie Feldnamen nicht als erste Zeile exportieren möchten, deaktivieren Sie das Kontrollkästchen **Feldnamen als erste Zeile exportieren**.
10. Klicken Sie auf **OK**.

KAPITEL 12

Scorecards in Informatica Developer

Dieses Kapitel umfasst die folgenden Themen:

- Scorecards in Informatica Developer - Übersicht, 99
- Erstellen einer Scorecard, 99
- Exportieren einer Ressourcendatei zur Scorecard-Herkunftskontrolle, 100
- Anzeigen der Scorecard-Herkunft aus Informatica Developer, 100

Scorecards in Informatica Developer - Übersicht

Eine Scorecard ist eine grafische Darstellung der Qualitätsmessungen in einem Profil. Sie können Scorecards im Developer-Tool ansehen. Nachdem Sie eine Scorecard im Developer-Tool erstellt haben, können Sie sich mit dem Analyst-Tool verbinden und die Scorecard zur Bearbeitung öffnen. Führen Sie die Scorecard für aktuelle Daten im Datenobjekt oder für im Profiling Warehouse zwischengespeicherte Daten aus.

Sie können eine Scorecard bearbeiten, sie ausführen und im Analyst-Tool die Scorecard-Herkunft für eine Metrik oder Metrikgruppe anzeigen.

Erstellen einer Scorecard

Erstellen Sie eine Scorecard, und fügen Sie Spalten aus einem Profil zur Scorecard hinzu. Sie müssen ein Profil ausführen, bevor Sie Spalten zur Scorecard hinzufügen.

1. Wählen Sie im **Objekt-Explorer** das Projekt oder den Ordner, in dem Sie die Scorecard erstellen möchten.
2. Klicken Sie auf **Datei** > **Neu** > **Scorecard**.

   Das Dialogfenster **Neue Scorecard** wird aufgerufen.
3. Klicken Sie auf **Hinzufügen**.

   Das Dialogfenster **Profil auswählen** wird angezeigt. Wählen Sie das Profil aus, das die Spalten enthält, die Sie hinzufügen möchten.
4. Klicken Sie auf **OK** und anschließend auf **Weiter**.
5. Wählen Sie die Spalten aus, die Sie zu der Scorecard hinzufügen möchten.

Standardmäßig wählt der Scorecard-Assistent die im Profil definierten Spalten und Regeln aus. Sie können Spalten hinzufügen, die nicht im Profil enthalten sind.

6. Klicken Sie auf **Fertigstellen**.

   Das Developer Tool erstellt die Scorecard.

7. Klicken Sie optional auf **Mit Informatica Analyst öffnen**, um eine Verbindung zum Analyst Tool herzustellen und die Scorecard im Analyst Tool zu öffnen.

Exportieren einer Ressourcendatei zur Scorecard-Herkunftskontrolle

Sie können ein Projekt, das Scorecards und abhängige Objekte enthält, als Ressourcendatei für Metadata Manager exportieren. Verwenden Sie die exportierte Ressourcendatei im XML-Format, um eine Ressource zur Scorecard-Herkunftskontrolle in Metadata Manager zu erstellen und zu laden.

1. Um den **Exportassistenten** zu öffnen, klicken Sie auf **Datei** > **Exportieren**.
2. Wählen Sie **Informatica** > **Ressourcendatei für Metadata Manager**.
3. Klicken Sie auf **Weiter**.
4. Klicken Sie auf **Durchsuchen**, um das Projekt auszuwählen, das die zu exportierenden Scorecard-Objekte und Herkunftsdaten enthält.
5. Klicken Sie auf **Weiter**.
6. Wählen Sie die zu exportierenden Scorecard-Objekte.
7. Geben Sie den Namen und den Speicherort der Exportdatei ein.
8. Um die abhängigen Objekte anzuzeigen, die der **Exportassistent** mit den von Ihnen gewählten Objekten exportiert, klicken Sie auf **Weiter**.

   Der **Export**-Assistent zeigt die abhängigen Objekte an.

9. Klicken Sie auf **Fertigstellen**.

   Das Developer-Tool exportiert die Objekte in die XML-Datei.

Anzeigen der Scorecard-Herkunft aus Informatica Developer

Um die Scorecard-Herkunft für eine Metrik oder Metrikgruppe aus dem Developer-Tool anzuzeigen, starten Sie das Analyst-Tool.

1. Wählen Sie in der Ansicht **Objekt-Explorer** das Projekt oder den Ordner, in dem sich die Scorecard befindet.
2. Doppelklicken Sie auf die Scorecard, um sie zu öffnen.

   Die Scorecard wird auf einer Registerkarte angezeigt.

3. Klicken Sie auf **Mit Informatica Analyst öffnen**.

Das Analyst-Tool wird im Browserfenster geöffnet.

4. Wählen Sie in der Ansicht **Scorecard** im Analyst-Tool eine Metrik oder Metrikgruppe.
5. Wählen Sie per Rechtsklick die Option **Herkunft anzeigen**.

   Das Scorecard-Herkunftsdiagramm erscheint in einem Dialogfenster.

KAPITEL 13

Mapplet- und Mapping-Profiling

Dieses Kapitel umfasst die folgenden Themen:

- Mapplet- und Mapping-Profiling - Übersicht Mapplet- und Mapping-Profile, 102
- Ausführen eines Profils für ein Mapplet- oder Mapping-Objekt, 102
- Vergleichen von Profilen für Mapping- oder Mapplet-Objekte, 103
- Generieren eines Mapping aus einem Profil, 103

Mapplet- und Mapping-Profiling - Übersicht Mapplet- und Mapping-Profile

Sie können ein Spaltenprofil für ein Objekt in einem Mapplet oder Mapping definieren. Führen Sie ein Profil an einem Mapplet- oder Mapping-Objekt aus, wenn Sie den Aufbau des Mappings oder Mapplets verifizieren möchten, ohne die Profilergebnisse zu speichern. Sie können aus einem Profil auch ein Mapping erstellen.

Ausführen eines Profils für ein Mapplet- oder Mapping-Objekt

Beim Ausführen eines Profils für ein Mapplet- oder Mapping-Objekt wird das Profil für alle Datenspalten ausgeführt und ermöglicht Drilldown-Vorgänge für die Daten, die für das Datenobjekt zwischengespeichert wurden. Sie können ein Profil für ein Mapplet oder Mapping-Objekt mit mehreren Ausgabeports ausführen. Beim Ausführen eines Profils für ein Mapplet- oder Mapping-Objekt wird das Profil für alle Datenspalten ausgeführt und ermöglicht Drilldown-Vorgänge für die Daten. Sie können ein Profil für ein Mapplet oder Mapping-Objekt mit mehreren Ausgabeports ausführen.

Das Profil verfolgt die Quellendaten durch das Mapping zu den Ausgabeports des ausgewählten Objekts. Es analysiert die Daten, die bei Ausführung des Mappings auf diesen Ports erscheinen würden.

1. Öffnen Sie ein Mapplet oder Mapping.
2. Vergewissern Sie sich, dass das Mapplet oder Mapping gültig ist.
3. Klicken Sie mit der rechten Maustaste auf ein Datenobjekt oder eine Umwandlung, und wählen Sie **Jetzt profilieren**.

Wenn die Umwandlung mehrere Ausgabegruppen hat, wird das Dialogfeld **Ausgabegruppe auswählen** eingeblendet. Wenn die Umwandlung eine einzige Ausgabegruppe hat, werden die Profilergebnisse auf der Registerkarte **Ergebnisse** des Profils eingeblendet.

4. Wenn die Umwandlung mehrere Ausgabegruppen enthält, wählen Sie nach Bedarf die Ausgabegruppen.
5. Klicken Sie auf **OK**.

   Die Profilergebnisse werden auf der Registerkarte **Ergebnisse** des Profils angezeigt.

Vergleichen von Profilen für Mapping- oder Mapplet-Objekte

Sie können ein Profil erstellen, das zwei Objekte in einem Mapplet oder Mapping analysiert und die Ergebnisse der Spaltenprofile für diese Objekte vergleicht.

Wie Profile einzelner Mapping- oder Mapplet-Objekte laufen Vergleiche auf allen Datenspalten und ermöglichen Drilldown-Operationen mit den Staging-Daten für die Datenobjekte.

Wie die Profile einzelner Mapping- oder Mapplet-Objekte werden Profilvergleiche auf allen Datenspalten ausgeführt.

1. Öffnen Sie ein Mapplet oder Mapping.
2. Vergewissern Sie sich, dass das Mapplet oder Mapping gültig ist.
3. Drücken Sie die Taste **STRG**, und klicken Sie zwei Objekte im Editor an.
4. Klicken Sie mit der rechten Maustaste auf eines der Objekte, und wählen Sie **Profile vergleichen**.
5. Konfigurieren Sie optional den Profilvergleich zum Abgleichen der Spalten von einem Objekt zum anderen.
6. Gleichen Sie optional Spalten ab, indem Sie eine Spalte in einem Objekt anklicken und sie in eine Spalte im anderen Objekt ziehen.
7. Wählen Sie optional, ob das Profil alle Spalten oder nur abgeglichene Spalten analysiert.
8. Klicken Sie auf **OK**.

Generieren eines Mapping aus einem Profil

Sie können aus einem Profil ein Mapping-Objekt erstellen. Verwenden Sie das von Ihnen erstellte Mapping-Objekt, um ein gültiges Mapping zu generieren. Das von Ihnen erstellte Mapping verfügt über eine Datenquelle, die auf dem profilierten Objekt basiert und Umwandlungen enthalten kann, die auf der Profilregellogik basiert. Nachdem Sie das Mapping erstellt haben, müssen Sie Objekte hinzufügen, um es fertig zu stellen.

1. Suchen Sie in der Ansicht **Objekt-Explorer** das Profil, für das Sie das Mapping erstellt haben.
2. Klicken Sie mit der rechten Maustaste auf den Profilnamen, und wählen Sie **Mapping generieren**.

   Das Dialogfenster **Mapping generieren** wird angezeigt.
3. Geben Sie einen Mapping-Namen ein. Optional können Sie eine Beschreibung für das Mapping eingeben.
4. Bestätigen Sie den Ordnerspeicherplatz für das Mapping.

Standardmäßig erstellt das Developer-Tool das Mapping im Ordner **Mappings** in demselben Projekt wie das Profil. Klicken Sie auf **Durchsuchen**, um einen anderen Speicherort für das Mapping auszuwählen.

5. Bestätigen Sie die Profildefinition, die das Developer-Tool zum Erstellen des Mappings verwendet. Um ein anderes Profil zu verwenden, klicken Sie auf **Profil auswählen**.
6. Klicken Sie auf **Fertigstellen**.

   Das Mapping wird im **Objekt-Explorer**angezeigt.

Fügen Sie Objekte zu dem Mapping hinzu, um es fertigzustellen.

INDEX

D

Datenobjektprofile
- Erstellen eines einzelnen Profils 81
- Erstellen mehrerer Kundenprofile 82

E

Einfachdatei-Datenobjekt
- Synchronisieren 29

Erstellen einer Expressionsregel
- Regeln 34

Erstellen eines Spaltenprofils
- Profile 26

Exportieren
- Scorecard-Herkunft in XML 100

Expressionsregeln
- Verarbeiten 33

I

Informatica Analyst
- Regeln 31
- Spaltenprofile - Übersicht 22, 50
- Spaltenprofilergebnisse 36, 48
- Verwaltung von Sperren und Versionen 26

Informatica Developer
- Regeln 91

M

Mapping-Objekt
- Ausführen eines Profils 102

Mapplet- und Mapping-Profile
- Übersicht 102

Mapplet- und Mapping-Profiling
- Übersicht 102

P

Profil
- Avro- oder Parquet-Formate 86
- Komponenten 14
- XML- und JSON-Formate 85, 86

Profile
- Ausführen 28, 49, 50
- Bearbeiten eines Spaltenprofils 28
- Einführung 13
- Erstellen eines Spaltenprofils 26

Profilergebnisse
- Ablehnen von Datentypen 56
- Ablehnen von Datentypen im Developer Tool 97

Profilergebnisse *(Fortsetzung)*
- Detailansicht 40
- Drilldown 55
- Excel 57
- Exportieren 57
- Exportieren aus Informatica Analyst 58
- Exportieren in Informatica Developer 98
- Genehmigen von Datentypen 56
- Genehmigen von Datentypen in Informatica Developer 97
- Spaltendatentypen 42, 96
- Spaltenmuster 44
- Spaltenwerte 45
- Zusammenfassung 52, 54
- Zusammenfassungsansicht 38

Profiling
- Prozess 14
- Tools 14
- Verwaltung von Sperren und Versionen 17

R

Regeln
- Anwenden einer vordefinierten Regel 32
- Anwenden in Informatica Developer 92
- Anwenden in PowerCenter Express 92
- Erstellen einer Expressionsregel 34
- Erstellen in Informatica Developer 92
- Expression 33
- Voraussetzungen 91
- vordefiniert 32

S

Scorecard
- Konfigurieren von Benachrichtigungen 75
- Konfigurieren von globalen Benachrichtigungseinstellungen 76

Scorecard-Ergebnisse
- Exportieren 73
- Exportieren aus Informatica Analyst 73
- Exportieren nach Excel 73

Scorecard-Herkunft
- Anzeigen aus Informatica Developer 100
- Anzeigen in Informatica Analyst 77

Scorecards
- anzeigen 64
- Ausführen 63
- bearbeiten 64
- Bearbeiten einer Metrikgruppe 68
- Benachrichtigungen 74
- Definieren von Schwellenwerten 66
- Drilldown 69
- Erstellen einer Metrikgruppe 67
- Feste Kosten 66
- Hinzufügen von Spalten zu einer Scorecard 62

Scorecards *(Fortsetzung)*
Informatica Analyst 59
Informatica Analyst-Prozess 60
Informatica Developer 99
Kosten für ungültige Daten 66
Löschen einer Metrikgruppe 68
Metriken 65
Metrikgewichtungen 65
Metrikgruppen 67
Trenddiagramm 69
Übersicht 18
Variable Kosten 66
Verschieben von Scores 67
Sonderfall
Erkennen 44
Spaltenprofil
Betriebssystemprofil 25
Drilldown 55
Informatica Developer 79
Optionen 17
Übersicht 16
Verarbeiten 23
Spaltenprofilergebnisse
Informatica Developer 93
Spaltenprofil 93
Spaltenprofilergebnisse im Analyst Tool
Schnittstelle 39, 51, 53
Spaltendetails 39, 54
Spaltenprofilergebnisse im Analyst Tool *(Fortsetzung)*
Zusammenfassung 37

T

Tabellendatenobjekt
Synchronisieren 30
Trenddiagramme
anzeigen 71
Exportieren aus Informatica Analyst 72
Kostenbasiert 70
Score 70

V

Vordefinierte Regeln
Prozess 32

W

Wiederherstellung
Aufgaben 20
Informatica Analyst 56
Informatica Developer 97
Konzepte 19