



Informatica®

9.6.1 HotFix 2

Guide de la découverte de données

© Copyright Informatica LLC 1998, 2018

Ce logiciel et sa documentation contiennent des informations appartenant à Informatica Corporation, protégées par la loi sur le droit d'auteur et fournies dans le cadre d'un accord de licence prévoyant des restrictions d'utilisation et de divulgation. Toute ingénierie inverse du logiciel est interdite. Il est interdit de reproduire ou transmettre sous quelque forme et par quelque moyen que ce soit (électronique, photocopie, enregistrement ou autre) tout ou partie de ce document sans le consentement préalable d'Informatica Corporation. Ce logiciel peut être protégé par des brevets américains et/ou internationaux, ainsi que par d'autres brevets en attente.

L'utilisation, la duplication ou la divulgation du Logiciel par le gouvernement américain est sujette aux restrictions décrites dans l'accord de licence applicable du logiciel conformément aux documents DFARS 227.7202-1(a) et 227.7702-3(a) (1995), DFARS 252.227-7013[©](1)(ii) (OCT 1988), FAR 12.212(a) (1995), FAR 52.227-19 ou FAR 52.227-14 (ALT III) le cas échéant.

Les informations dans ce produit ou cette documentation sont sujettes à modification sans préavis. Si vous rencontrez des problèmes dans ce produit ou la documentation, veuillez nous en informer par écrit.

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 et Informatica Master Data Management sont des marques de commerce ou des marques déposées d'Informatica Corporation aux États-Unis et dans d'autres juridictions du monde. Tous les autres noms de société ou de produit peuvent être des marques de commerce ou des marques déposées de leurs détenteurs respectifs.

Des portions de ce logiciel et/ou de la documentation sont sujettes au copyright détenu par des tierces parties, dont Copyright DataDirect Technologies. Tous droits réservés. Copyright © Sun Microsystems. Tous droits réservés. Copyright © RSA Security Inc. Tous droits réservés. Copyright © Ordinal Technology Corp. Tous droits réservés. Copyright © Aandacht c.v. Tous droits réservés. Copyright Genivia, Inc. Tous droits réservés. Copyright Isomorphic Software. Tous droits réservés. Copyright © Meta Integration Technology, Inc. Tous droits réservés. Copyright © Intalio. Tous droits réservés. Copyright © Oracle. Tous droits réservés. Copyright © Adobe Systems Incorporated. Tous droits réservés. Copyright © DataArt, Inc. Tous droits réservés. Copyright © ComponentSource. Tous droits réservés. Copyright © Microsoft Corporation. Tous droits réservés. Copyright © Rogue Wave Software, Inc. Tous droits réservés. Copyright © Teradata Corporation. Tous droits réservés. Copyright © Yahoo! Inc. Tous droits réservés. Copyright © Glyph & Cog, LLC. Tous droits réservés. Copyright © Thinkmap, Inc. Tous droits réservés. Copyright © Clearpace Software Limited. Tous droits réservés. Copyright © Information Builders, Inc. Tous droits réservés. Copyright © OSS Nokalva, Inc. Tous droits réservés. Copyright Edifecs, Inc. Tous droits réservés. Copyright Cleo Communications, Inc. Tous droits réservés. Copyright © International Organization for Standardization 1986. Tous droits réservés. Copyright © ej-technologies GmbH. Tous droits réservés. Copyright © Jaspersoft Corporation. Tous droits réservés. Copyright © International Business Machines Corporation. Tous droits réservés. Copyright © yWorks GmbH. Tous droits réservés. Copyright © Lucent Technologies. Tous droits réservés. Copyright © Université de Toronto. Tous droits réservés. Copyright © Daniel Veillard. Tous droits réservés. Copyright © Uniconic, Inc. Copyright IBM Corp. Tous droits réservés. Copyright © MicroQuill Software Publishing, Inc. Tous droits réservés. Copyright © PassMark Software Pty Ltd. Tous droits réservés. Copyright © LogiXML, Inc. Tous droits réservés. Copyright © 2003-2010 Lorenzi Davide. Tous droits réservés. Copyright © Red Hat, Inc. Tous droits réservés. Copyright © The Board of Trustees of the Leland Stanford Junior University. Tous droits réservés. Copyright © EMC Corporation. Tous droits réservés. Copyright © Flexera Software. Tous droits réservés. Copyright © Jinfonet Software. Tous droits réservés. Copyright © Apple Inc. Tous droits réservés. Copyright © Telerik Inc. Tous droits réservés. Copyright © BEA Systems. Tous droits réservés. Copyright © PDFlib GmbH. Tous droits réservés. Copyright © Orientation in Objects GmbH. Tous droits réservés. Copyright © Tanuki Software, Ltd. Tous droits réservés. Copyright © Ricebridge. Tous droits réservés. Copyright © Sencha, Inc. Tous droits réservés. Copyright © Scalable Systems, Inc. Tous droits réservés. Copyright © jQWidgets. Tous droits réservés.

Ce produit inclut des logiciels développés par Apache Software Foundation (<http://www.apache.org/>), et/ou d'autres logiciels sous licence et sous diverses versions Apache License (la « Licence »). Vous pouvez obtenir une copie de ces licences à l'adresse suivante : <http://www.apache.org/licenses/>. Sauf dispositions contraires de la loi en vigueur ou accord écrit, le logiciel distribué sous cette licence est livré « EN L'ÉTAT », SANS GARANTIE NI CONDITION D'AUCUNE SORTE, expresse ou implicite. Se reporter aux Licences pour la langue spécifique régissant les droits et limitations dans le cadre des Licences.

Ce produit inclut des logiciels développés par Mozilla (<http://www.mozilla.org/>), copyright de logiciel The JBoss Group, LLC, tous droits réservés ; copyright de logiciel © 1999-2006 de Bruno Lowagie et Paulo Soares et d'autres logiciels sous licence et sous diverses versions du GNU Lesser General Public License Agreement, accessible sur <http://www.gnu.org/licenses/lgpl.html>. Les matériaux sont fournis gratuitement par Informatica, « en l'état », sans garantie d'aucune sorte, expresse ou implicite, notamment les garanties implicites de conformité légale et d'usage normal.

Le produit inclut les logiciels ACE(TM) et TAO(TM), copyright Douglas C. Schmidt et son groupe de recherche à Washington University, University of California, Irvine, et Vanderbilt University, Copyright (©) 1993-2006, tous droits réservés.

Ce produit inclut des logiciels développés par OpenSSL Project pour une utilisation dans OpenSSL Toolkit (copyright The OpenSSL Project. Tous droits réservés) et la redistribution de ce logiciel est sujette aux termes publiés sur <http://www.openssl.org> et <http://www.openssl.org/source/license.html>.

Ce produit inclut le logiciel Curl, copyright 1996-2013, Daniel Stenberg, <daniel@haxx.se>. Tous Droits Réservés. Les autorisations et limitations concernant ce logiciel sont sujettes aux conditions publiées sur <http://curl.haxx.se/docs/copyright.html>. L'autorisation d'utiliser, copier, modifier et distribuer ce logiciel à toute fin, avec ou sans rémunération, est accordée par les présentes, à la condition que la notification de copyright ci-dessus et cette notification d'autorisation apparaissent dans toutes les copies.

Le produit inclut des logiciels sous copyright 2001-2005 (©) MetaStuff, Ltd. Tous droits réservés. Les autorisations et limitations concernant ce logiciel sont sujettes aux conditions publiées sur <http://www.dom4j.org/license.html>.

Le produit inclut des logiciels sous copyright © 2004-2007, The Dojo Foundation. Tous Droits Réservés. Les autorisations et limitations concernant ce logiciel sont sujettes aux conditions publiées sur <http://dojotoolkit.org/license>.

Ce produit inclut le logiciel ICU sous copyright de International Business Machines Corporation et autres. Tous Droits Réservés. Les autorisations et limitations concernant ce logiciel sont sujettes aux conditions publiées sur <http://source.icu-project.org/repos/icu/icu/trunk/license.html>.

Ce produit inclut des logiciels sous copyright © 1996-2006 Per Bothner. Tous Droits Réservés. Votre droit à utiliser de tels matériels est défini dans la licence qui peut être consultée sur <http://www.gnu.org/software/kawa/Software-License.html>.

Ce produit inclut le logiciel OSSP UUID sous copyright © 2002 Ralf S. Engelschall, copyright © 2002 The OSSP Project Copyright © 2002 Cable & Wireless Deutschland. Les autorisations et limitations concernant ce logiciel sont sujettes aux conditions publiées sur <http://www.opensource.org/licenses/mit-license.php>.

Ce produit inclut des logiciels développés par Boost (<http://www.boost.org/>) ou sous licence de logiciel Boost. Les autorisations et limitations concernant ce logiciel sont sujettes aux conditions publiées sur http://www.boost.org/LICENSE_1_0.txt.

Ce produit inclut des logiciels sous copyright © 1997-2007 University of Cambridge. Les autorisations et limitations concernant ce logiciel sont sujettes aux conditions publiées sur <http://www.pcre.org/license.txt>.

Ce produit inclut des logiciels sous copyright © 2007 The Eclipse Foundation. Tous Droits Réservés. Les autorisations et limitations concernant ce logiciel sont sujettes aux conditions publiées sur <http://www.eclipse.org/org/documents/epl-v10.php> et <http://www.eclipse.org/org/documents/edl-v10.php>.

Ce produit comprend des logiciels sous licence dont les conditions se trouvent aux adresses : <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/hsqldbLicense.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/javaxservice/>, <http://www.postgresql.org/about/license.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>, et <https://code.google.com/p/lz4/>.

Ce produit inclut un logiciel sous licence Academic Free License (<http://www.opensource.org/licenses/afl-3.0.php>), licence Common Development Distribution License (<http://www.opensource.org/licenses/cddl1.php>) licence Common Public License (<http://www.opensource.org/licenses/cpl1.0.php>), licence Sun Binary Code License Agreement Supplemental License Terms, licence BSD (<http://www.opensource.org/licenses/bsd-license.php>), la nouvelle licence BSD License (<http://opensource.org/licenses/BSD-3-Clause>), la licence MIT (<http://www.opensource.org/licenses/mit-license.php>), la licence Artistic License (<http://www.opensource.org/licenses/artistic-license-1.0>) et la licence publique du développeur initial Version 1.0 (<http://www.firebirdsql.org/en/initial-developer-s-public-license-version-1-0/>).

Ce produit inclut des logiciels sous copyright © 2003-2006 Joe Walnes, 2006-2007 XStream Committers. Tous Droits Réservés. Les autorisations et limitations concernant ce logiciel sont sujettes aux conditions publiées sur <http://xstream.codehaus.org/license.html>. Ce produit inclut des logiciels développés par Indiana University Extreme! Lab. Pour plus d'informations, veuillez vous rendre sur <http://www.extreme.indiana.edu/>.

Ce produit inclut des logiciels sous copyright © 2013 Frank Balluffi et Markus Moeller. Tous droits réservés. Les autorisations et limitations concernant ce logiciel sont sujettes aux conditions de la licence MIT.

Ce logiciel est protégé par des brevets américains (5,794,246; 6,014,670; 6,016,501; 6,029,178; 6,032,158; 6,035,307; 6,044,374; 6,092,086; 6,208,990; 6,339,775; 6,640,226; 6,789,096; 6,823,373; 6,850,947; 6,895,471; 7,117,215; 7,162,643; 7,243,110; 7,254,590; 7,281,001; 7,421,458; 7,496,588; 7,523,121; 7,584,422; 7,676,516; 7,720,842; 7,721,270; 7,774,791; 8,065,266; 8,150,803; 8,166,048; 8,166,071; 8,200,622; 8,224,873; 8,271,477; 8,327,419; 8,386,435; 8,392,460; 8,453,159; 8,458,230; 8,707,336; 8,886,617 et RE44,478), des brevets internationaux ainsi que par d'autres brevets en attente.

EXCLUSION DE RESPONSABILITÉ : Informatica Corporation fournit cette documentation « en l'état », sans garantie d'aucune sorte, explicite ou implicite, notamment les garanties implicites de non-infraction, de conformité légale ou d'usage normal. Informatica Corporation ne garantit pas que ce logiciel et cette documentation sont exempts d'erreurs. Les informations fournies dans ce logiciel ou cette documentation peuvent inclure des inexactitudes techniques ou des erreurs typographiques. Les informations contenues dans ce logiciel et sa documentation sont sujettes à modification à tout moment sans préavis.

AVIS

Ce produit Informatica (le « Logiciel ») inclut certains pilotes (les « Pilotes DataDirect ») de DataDirect Technologies, une société de Progress Software Corporation (« DataDirect ») qui sont sujets aux conditions suivantes :

1. LES PILOTES DATADIRECT SONT FOURNIS « EN L'ÉTAT », SANS GARANTIE D'AUCUNE SORTE, EXPRESSE OU IMPLICITE, NOTAMMENT LES GARANTIES IMPLICITES DE CONFORMITÉ LÉGALE, D'USAGE NORMAL ET DE NON-INFRACTION.
2. DATADIRECT OU SES FOURNISSEURS TIERS NE POURRONT EN AUCUN CAS ÊTRE TENUS RESPONSABLES ENVERS LE CLIENT UTILISATEUR FINAL DE TOUT DOMMAGE DIRECT, ACCESSOIRE, INDIRECT, SPÉCIAL, CONSÉCUTIF OU AUTRE RÉSULTANT DE L'UTILISATION DES PILOTES ODBC, QU'ILS SOIENT INFORMÉS OU NON À L'AVANCE DE LA POSSIBILITÉ DE TELS DOMMAGES. CES LIMITATIONS S'APPLIQUENT À TOUTES LES CAUSES D'ACTION, NOTAMMENT TOUTE INFRACTION AU CONTRAT, INFRACTION À LA GARANTIE, NÉGLIGENCE, RESPONSABILITÉ STRICTE, REPRÉSENTATION INCORRECTE ET AUTRES TORTS.

Date de publication: 2018-06-18

Sommaire

Préface.....	13
Ressources Informatica.	13
Portail Mon support Informatica.	13
Documentation Informatica.	13
Matrices de disponibilité de produit Informatica.	13
Site Web Informatica.	14
Bibliothèque de procédures Informatica.	14
Base de connaissances Informatica.	14
Canal YouTube du support Informatica.	14
Informatica Marketplace.	14
Informatica Velocity.	14
Support client international Informatica.	15
 Partie I: Introduction à Découverte de données.....	 16
 Chapitre 1: Introduction à Profilage.	 17
Présentation du profilage.	17
Architecture de profilage.	18
Processus de découverte de données.	20
 Chapitre 2: Découverte de données.	 22
Présentation de la découverte de données.	22
Types de profil et d'analyse.	22
Composants de profilage.	24
Résultats de profil.	24
 Chapitre 3: Concepts de profil de colonne.	 26
Présentation des concepts de profil de colonnePrésentation des profils.	26
Options de profil de colonne.	27
Règles.	27
Fiches d'évaluation.	28
 Chapitre 4: Concepts de découverte des domaines de données.	 29
Présentation des concepts de découverte des domaines de données.	29
Domaines de données.	30
Groupes de domaines de données.	30
Glossaire de domaine de données.	31
Processus de découverte des domaines de données.	31

Chapitre 5: Concepts de conservation.	32
Présentation des concepts de conservation.	32
Conservation pour les analystes et les développeurs.	32
Tâches de conservation.	33
Partie II: Découverte de données avec Informatica Analyst.	35
Chapitre 6: Profils de colonne dans Informatica Analyst.	36
Présentation des profils de colonne dans Informatica Analyst.	36
Processus de profilage de colonne.	37
Options de profil.	37
Option de résultats de profils.	38
Options d'échantillonnage.	38
Options de développement.	39
Création d'un profil de colonne dans l'outil Analyst.	39
Modification d'un profil de colonne.	40
Exécution d'un profil.	41
Création d'un filtre.	41
Gestion des filtres.	42
Synchronisation d'un objet de données fichier plat.	42
Synchronisation d'un objet de données relationnel.	43
Chapitre 7: Résultats de profil de colonne dans Informatica Analyst.	44
Présentation des résultats de profil de colonne dans Informatica Analyst.	44
Résumé de profil.	45
Valeurs de colonne.	46
Formes de colonne.	47
Types de données de colonne.	48
Statistiques de colonne.	49
Développement de profil de colonne.	50
Développement des données de ligne.	50
Application de filtres aux données de développement.	50
Conservation dans l'outil Analyst.	51
Approbation des types de données.	51
Rejet des types de données.	51
Fichiers d'exportation de profil de colonne dans Informatica Analyst.	52
Résultats de l'exportation de profil dans un fichier CSV.	52
Résultats de l'exportation de profil dans Microsoft Excel.	52
Exportation des résultats de profil depuis Informatica Analyst.	53
Chapitre 8: Règles dans Informatica Analyst.	54
Présentation des règles dans Informatica Analyst.	54

Règles prédéfinies.	55
Processus de règles prédéfinies.	55
Application d'une règle prédéfinie.	55
Règles d'expression.	56
Processus de règles d'expression.	56
Création d'une règle d'expression.	57
Chapitre 9: Fiches d'évaluation dans Informatica Analyst.	58
Présentation des fiches d'évaluation dans Informatica Analyst.	58
Processus de fiche d'évaluation de Informatica Analyst.	59
Mesures.	60
Poids métriques.	60
Valeur de la qualité des données.	60
Ajout de colonnes à une fiche d'évaluation.	61
Exécution d'une fiche d'évaluation.	62
Affichage d'une fiche d'évaluation.	62
Édition d'une fiche d'évaluation.	62
Définition des seuils.	63
Groupes de mesures.	63
Création d'un groupe de mesures.	64
Déplacement des scores vers un groupe de mesures.	64
Modification d'un groupe de mesures.	64
Suppression d'un groupe de mesures.	65
Développement dans les colonnes.	65
Graphes de tendance.	66
Graphe de tendance de score.	66
Graphe de tendance de coût.	67
Affichage des graphes de tendances.	67
Exportation des graphes de tendance.	68
Fichier d'exportation de fiche d'évaluation dans Informatica Analyst.	69
Résultats de l'exportation de fiche d'évaluation dans Microsoft Excel.	69
Exportation des résultats de fiche d'évaluation à partir d'Informatica Analyst.	69
Notifications de la fiche d'évaluation.	70
Modèle de notification de message électronique.	70
Configuration des notifications de la fiche d'évaluation.	71
Configuration des paramètres généraux des notifications de la fiche d'évaluation.	72
Intégration des fiches d'évaluation à des applications externes.	72
Consultation d'une fiche d'évaluation dans des applications externes.	73
Lignage des fiches d'évaluation.	74
Affichage du lignage des fiches d'évaluation dans Informatica Analyst.	74

Chapitre 10: Découverte des domaines de données dans Informatica Analyst. 75

Présentation de la découverte des domaines de données dans Informatica Analyst.	75
Glossaire de domaine de données dans Informatica Analyst.	75
Création d'un groupe de domaines de données dans Informatica Analyst.	76
Création d'un domaine de données dans Informatica Analyst.	76
Création d'un domaine de données depuis les résultats de profil dans Informatica Analyst. . . .	77
Rechercher des domaines de données et des groupes de domaines de données dans Informatica Analyst.	77
Options de découverte des domaines de données dans Informatica Analyst.	78
Sélection de colonne de domaine de données dans Informatica Analyst.	78
Sélection des domaines de données dans Informatica Analyst.	79
Options d'inférence des domaines de données dans Informatica Analyst.	79
Création d'un profil pour découvrir des domaines de données dans Informatica Analyst.	80
Modification d'un profil dans Informatica Analyst.	81
Exécution d'un profil pour effectuer la découverte des domaines de données.	81
Résultats de la découverte de domaines de données dans Informatica Analyst.	81
Affichage par groupes de domaines de données dans Informatica Analyst.	82
Affichage par colonnes dans Informatica Analyst.	83
Approbation des domaines de données.	83
Rejet des domaines de données.	83
Fichiers d'exportation de la découverte des domaines de données dans Informatica Analyst.	84
Résultats de la découverte des domaines de données dans Microsoft Excel.	84
Exportation des résultats de la découverte des domaines de données depuis Informatica Analyst.	84

Chapitre 11: Découverte des données d'entreprise dans Informatica Analyst. 86

Présentation de la découverte des données d'entreprise dans Informatica Analyst.	86
Processus de découverte des données d'entreprise dans l'outil Analyst.	87
Options de configuration pour la découverte des données d'entreprise.	87
Paramètres de découverte de domaines de données.	87
Options d'échantillonnage de profil de colonne.	88
Exécution d'un profil de découverte des données d'entreprise des données dans l'outil Analyst. . . .	89
Modification des options de découverte des données d'entreprise.	90

Chapitre 12: Résultats de la découverte des données d'entreprise dans Informatica Analyst. 91

Présentation des résultats de la découverte des données d'entreprise dans l'outil Analyst.	91
Vue Résumé.	91
Résultats de profil de la vue Résumé.	92
Affichage des résultats de la découverte de domaines de données dans l'outil Analyst.	93
Affichage des résultats de profil de colonne.	93
Conflit de type de données.	94

Affichage des conflits de type de données.	94
Vue Profils.	94
Affichage des propriétés de profil.	95
Chapitre 13: Recherche de découverte dans Informatica Analyst.	96
Recherche de découverte dans Informatica Analyst - Présentation.	96
Prérequis de la recherche de découverte.	97
Processus de recherche de découverte dans Informatica Analyst.	98
Options de recherche de découverte.	98
Critères de recherche de découverte.	99
Recherche d'une ressource.	99
Résultats de la recherche de découverte dans Informatica Analyst.	100
Panneau de résultats de la recherche de découverte.	100
Filtrage des résultats de la recherche de découverte.	101
Types de correspondances.	102
Correspondance directe.	102
Correspondance indirecte.	102
Affichage des informations de correspondance.	102
Ouverture des ressources à partir des résultats de la recherche de découverte.	103
Ressources liées.	103
Ressources liées pour chaque type de ressource.	104
Affichage des ressources liées.	104
Forum Aux Questions (FAQ).	105
Chapitre 14: Bureau de glossaire métier dans Informatica Analyst.	106
Termes métier.	106
Gestion des termes métier dans le glossaire métier du gestionnaire de métadonnées.	107
Recherche d'un terme métier dans le Bureau de glossaire métier.	107
Partie III: Découverte de données avec Informatica Developer.	108
Chapitre 15: Profils de Informatica Developer.	109
Présentation des profils de Informatica Developer.	109
Vues des profils d'Informatica DeveloperVues des profils.	111
Chapitre 16: Profils d'objet de données.	112
Présentation de profils d'objet de données.	112
Profils de colonne dans Informatica Developer.	113
Options de filtrage.	113
Propriétés d'échantillonnage.	114
Découverte de clé primaire.	114
Propriétés d'inférence de clés primaires.	115
Propriétés de clés primaires inférées.	115

Propriétés de violations de clés.	116
Découverte de dépendance fonctionnelle.	116
Propriétés d'inférence de dépendances fonctionnelles.	116
Propriétés des dépendances fonctionnelles inférées.	117
Propriétés de violations de dépendances fonctionnelles.	117
Création d'un profil d'objet de données uniqueCréation d'un profil.	118
Création de plusieurs profils d'objets de données.	119
Synchronisation d'un objet de données fichier plat.	119
Synchronisation d'un objet de données relationnel.	119

Chapitre 17: Résultats de profil de colonne dans Informatica Developer. . . . 121

Résultats de profil de colonne dans Informatica DeveloperRésultats de profil de colonne.	121
Propriétés des valeurs de colonne.	122
Propriétés des formes de colonne.	123
Propriétés des statistiques de colonne.	123
Propriétés des types de données de colonne.	124
Conservation dans l'outil DeveloperConservation dans Informatica Developer.	125
Approbation des types de données dans l'outil DeveloperApprobation des types de données dans Informatica Developer.	125
Rejet des types de données dans l'outil DeveloperRejet des types de données dans Informatica Developer.	125
Exportation des résultats de profil depuis Informatica Developer.	126

Chapitre 18: Règles dans Informatica Developer. 127

Présentation des règles dans Informatica DeveloperInstructions pour les règles.	127
Création d'une règle dans Informatica Developer.	128
Application d'une règle dans Informatica DeveloperApplication d'une règle.	128

Chapitre 19: Fiches d'évaluation dans Informatica Developer. 129

Présentation des fiches d'évaluation dans Informatica Developer.	129
Création d'une fiche d'évaluation.	129
Exportation d'un fichier de ressource pour le lignage des fiches d'évaluation.	130
Affichage du lignage des fiches d'évaluation dans Informatica Developer.	130

Chapitre 20: Profilage de mapplet et de mappage. 132

Présentation du profilage de mapplet et de mappageProfils de mapplet et de mappage.	132
Exécution d'un profil sur un mapplet ou un objet de mappage.	132
Comparaison des profils pour les objets de mappage ou de mapplet.	133
Génération d'un mappage à partir d'un profil.	133

Chapitre 21: Découverte des domaines de données dans Informatica Developer. 135

Présentation de la découverte de domaines de données dans Informatica Developer.	135
Glossaire de domaine de données dans Informatica Developer.	136

Création d'un groupe de domaines de données dans Informatica Developer.	136
Création d'un domaine de données dans Informatica Developer.	137
Création d'un domaine de données depuis les résultats de profil dans Informatica Developer.	137
Rechercher des domaines de données dans Informatica Developer.	138
Importation des domaines de données.	138
Exportation des domaines de données.	139
Options de découverte des domaines de données dans Informatica Developer.	140
Sélection de domaines de données dans Informatica Developer.	140
Sélection de colonne de domaine de données dans Informatica Developer.	141
Options d'inférence des domaines de données dans Informatica Developer.	141
Création d'un profil pour découvrir des domaines de données dans Informatica Developer.	142
Modification d'un profil dans Informatica Developer.	142
Exécution d'un profil pour effectuer une découverte des domaines de données dans Informatica Developer.	143
Résultats de la découverte des domaines de données dans Informatica Developer.	143
Affichage par groupes de domaines de données dans Informatica Developer.	144
Affichage par colonnes dans Informatica Developer.	144
Vérification des résultats dans Informatica Developer.	145
Approbation des domaines de données dans l'outil Developer.	145
Rejet des domaines de données dans l'outil Developer.	146
Exportation des résultats de la découverte des domaines de données depuis Informatica Developer.	146

Chapitre 22: Découverte des données d'entreprise dans Informatica Developer. 147

Présentation de la découverte des données d'entreprise dans Informatica Developer.	147
Processus de découverte des données d'entreprise.	148
Options de profil pour la découverte des données d'entreprise.	149
Sélection de domaine de données pour la découverte des données d'entreprise.	149
Options d'échantillonnage de profil de colonne pour la découverte des données d'entreprise.	150
Options d'inférence de clé principale pour la découverte des données d'entreprise.	150
Options d'inférence de clé étrangère pour la découverte des données d'entreprise.	150
Création d'un profil de découverte des données d'entreprise dans Informatica Developer.	152
Modification d'un profil.	153
Exécution d'un profil de découverte des données d'entreprise.	153
Découverte de clé étrangère.	154
Définition des relations entre les objets parent et enfant.	155
Découverte de relations de clés étrangères entre les objets de données.	155
Résultats d'analyse de clé étrangère.	155
Analyse de jointure.	156
Création d'un profil de jointure.	156
Résultats d'analyse de jointure.	157
Exportation des résultats de profil de jointure dans un fichier.	158

Découverte du chevauchement.	158
Résultats de découverte du chevauchement.	158
Découverte de données de chevauchement.	159
Fichiers de script DDL.	160
Création de scripts DDL à partir d'un profil de découverte des données d'entreprise.	160

Chapitre 23: Résultats de la découverte des données d'entreprise. 161

Présentation des résultats de la découverte des données d'entreprise.	161
Vue Relations.	162
Rechercher un objet de données.	162
Navigation dans la vue Profilage de clé étrangère.	163
Vue Profilage de clé étrangère.	163
Affichage des relations d'objets de données.	163
Zoom avant et zoom arrière dans la vue.	164
Recherche d'un objet de données.	164
Affichage des relations de colonne.	165
Enregistrement du diagramme de la relation de l'entité sous la forme d'une image.	165
Affichage des résultats de profil d'objet de données depuis la vue Profilage de clé étrangère.	165
Vue tabulaire.	166
Volet Détails de la table.	166
Vérification des résultats de la découverte des données d'entreprise.	166
Conservation des relations de colonne dans l'outil Developer.	167
Validation des résultats dans le référentiel modèle.	167
Vue Domaines de données.	167
Affichage des résultats de la découverte des domaines de données.	168
Vérification des résultats de la découverte des domaines de données.	168
Développement des lignes.	168
Affichage des résultats de profil d'objet de données depuis la vue Domaines de données.	169
Vue Profil de colonne.	169
Affichage des résultats de profil d'objet de données.	169
Affichage des résultats de profil de colonne lors de l'exécution de la découverte des données d'entreprise.	170
Affichage des résultats de la découverte des domaines de données lors de l'exécution de la découverte des données d'entreprise.	170
Affichage du statut d'exécution de la découverte des données d'entreprise.	170
Fichiers d'exportation de la découverte des données d'entreprise.	171
Exportation des résultats de la découverte des données d'entreprise.	171

Chapitre 24: Bureau de glossaire métier dans Informatica Developer. 172

Recherche dans le glossaire métier.	172
Recherche d'un terme métier.	173
Personnalisation des touches de raccourci pour rechercher un terme métier.	173

Annexe A: Glossaire.....	174
Index.	178

Préface

Le *Guide de découverte de données* s'adresse aux utilisateurs d'Informatica Analyst et d'Informatica Developer. Il contient des informations sur la manière d'utiliser les profils pour analyser le contenu et la structure des données.

Utilisez les profils pour découvrir les problèmes de qualité des données dans un ensemble de données et pour comprendre les relations entre les colonnes dans un ou plusieurs ensembles de données.

Ressources Informatica

Portail Mon support Informatica

En tant que client Informatica, vous pouvez accéder au portail Mon support Informatica sur <http://mysupport.informatica.com>.

Ce site contient des informations sur les produits et les groupes d'utilisateurs, des bulletins d'information, un lien vers le système de gestion des dossiers d'assistance à la clientèle d'Informatica (ATLAS), une bibliothèque de procédures Informatica, une base de connaissances Informatica, ainsi que la documentation nécessaire sur les produits Informatica et l'accès à sa communauté d'utilisateurs.

Ce site contient des informations sur les produits et les groupes d'utilisateurs, des bulletins d'information, un lien vers la bibliothèque de procédures Informatica, une base de connaissances Informatica, ainsi que la documentation nécessaire sur les produits Informatica et l'accès à sa communauté d'utilisateurs.

Documentation Informatica

L'équipe Documentation d'Informatica s'efforce de fournir une documentation précise et utilisable. N'hésitez pas à contacter l'équipe Documentation d'Informatica par courriel à l'adresse infa_documentation@informatica.com pour lui faire part de vos questions, commentaires ou suggestions concernant cette documentation. Ces commentaires et suggestions nous permettront d'améliorer notre documentation. Veuillez préciser si vous acceptez d'être contacté au sujet de ces commentaires.

L'équipe Documentation met à jour la documentation chaque fois que nécessaire. Pour obtenir la toute dernière version de la documentation concernant votre produit, consultez la Documentation de produit sur <http://mysupport.informatica.com>.

Matrices de disponibilité de produit Informatica

Les matrices de disponibilité de produit (PAM) indiquent les versions des systèmes d'exploitation, les bases de données et les autres types de sources et cibles de données pris en charge par une version d'un produit.

Vous pouvez consulter les PAM sur le portail Mon Support Informatica à l'adresse <https://mysupport.informatica.com/community/my-support/product-availability-matrices>.

Site Web Informatica

Vous pouvez accéder au site Web d'entreprise Informatica sur <http://www.informatica.com>. Le site contient des informations sur Informatica, son expertise, les événements à venir et les bureaux de vente. Vous y trouverez aussi des informations sur ses produits et ses partenaires. Les rubriques de service du site fournissent des informations importantes sur le support technique, la formation et l'éducation, ainsi que les services d'implémentation.

Bibliothèque de procédures Informatica

En tant que client Informatica, vous avez accès à la bibliothèque de procédures Informatica sur <http://mysupport.informatica.com>. La bibliothèque de procédures Informatica est une collection de ressources destinée à vous familiariser avec les produits Informatica et leurs fonctionnalités. Elle regroupe des articles et des démonstrations interactives qui permettent de résoudre des problèmes courants et de comparer les fonctionnalités et les comportements, et qui vous guident lors de la réalisation de tâches concrètes spécifiques.

Base de connaissances Informatica

En tant que client Informatica, vous avez accès à la base de connaissances Informatica sur <http://mysupport.informatica.com>. Utilisez la base de connaissances pour rechercher des solutions documentées aux problèmes techniques connus concernant les produits Informatica. Vous y trouverez également la réponse aux questions les plus fréquentes, des livres blancs et des conseils techniques. N'hésitez pas à contacter l'équipe Base de connaissances Informatica par courriel à l'adresse KB_Feedback@informatica.com pour lui faire part de vos questions, commentaires et suggestions concernant la base de connaissances.

Canal YouTube du support Informatica

Vous pouvez accéder au canal YouTube du support Informatica sur <http://www.youtube.com/user/INFASupport>. Le canal YouTube du support Informatica contient des vidéos concernant les solutions qui vous guident dans l'exécution de tâches spécifiques. Si vous avez des questions, commentaires ou suggestions concernant le canal YouTube du support Informatica, contactez l'équipe de support YouTube par courriel à l'adresse supportvideos@informatica.com ou envoyez un tweet à @INFASupport.

Informatica Marketplace

Informatica Marketplace est un forum où développeurs et partenaires peuvent partager des solutions qui permettent d'augmenter, d'étendre ou d'améliorer les implémentations d'intégration de données. En tirant profit des centaines de solutions disponibles sur Marketplace, vous pouvez améliorer votre productivité et accélérer le temps d'implémentation de vos projets. Vous pouvez accéder à Informatica Marketplace à l'adresse <http://www.informaticamarketplace.com>.

Informatica Velocity

Vous pouvez accéder à Informatica Velocity à l'adresse <http://mysupport.informatica.com>. Développé à partir de l'expérience concrète de centaines de projets de gestion de données, Informatica Velocity représente le savoir collectif de nos consultants, qui ont travaillé avec des entreprises du monde entier pour

planifier, développer, déployer et tenir à jour des solutions de gestion des données efficaces. Si vous avez des questions, des commentaires et des suggestions sur Informatica Velocity, contactez le support des services professionnels Informatica à l'adresse ips@informatica.com.

Support client international Informatica

Vous pouvez contacter un centre de support client par téléphone ou via l'assistance en ligne.

L'assistance en ligne requiert un nom d'utilisateur et un mot de passe. Vous pouvez demander un nom d'utilisateur et un mot de passe sur <http://mysupport.informatica.com>.

Les numéros de téléphone du support client international Informatica sont disponibles sur le site Web Informatica à l'adresse <http://www.informatica.com/us/services-and-training/support-services/global-support-centers/>.

Partie I : Introduction à Découverte de données

Cette partie contient les chapitres suivants :

- [Introduction à Profilage, 17](#)
- [Découverte de données, 22](#)
- [Concepts de profil de colonne, 26](#)
- [Concepts de découverte des domaines de données, 29](#)
- [Concepts de conservation, 32](#)

CHAPITRE 1

Introduction à Profilage

Ce chapitre comprend les rubriques suivantes :

- [Présentation du profilage, 17](#)
- [Architecture de profilage, 18](#)
- [Processus de découverte de données, 20](#)

Présentation du profilage

Les profils s'utilisent pour rechercher le contenu, la qualité et la structure des sources de données d'une application, d'un schéma ou d'une entreprise. Le contenu de la source de données comprend les fréquences de valeurs et les types de données. La structure de sources de données comprend les clés et les dépendances fonctionnelles.

Vous pouvez créer et exécuter des profils comme partie du processus de découverte. Un profil est un objet de référentiel qui recherche et analyse toutes les irrégularités de données dans les sources de données de l'entreprise et les problèmes de données cachés qui mettent en péril les projets de données. L'exécution d'un profil sur une source de données quelconque de l'entreprise vous aide à bien comprendre les points forts et les points faibles de ses données et métadonnées.

Vous pouvez utiliser les outils Analyst et Developer pour analyser les données source et les métadonnées. Les analystes et les développeurs peuvent utiliser ces outils pour collaborer, identifier les problèmes de qualité de données et analyser les relations de données. Selon votre fonction, vous pouvez utiliser les fonctionnalités de l'outil Analyst ou de l'outil Developer. Le degré de profilage que vous pouvez effectuer varie selon l'outil que vous utilisez.

Vous pouvez effectuer les tâches suivantes dans l'outil Developer et l'outil Analyst :

- Effectuer le profilage de colonnes. Ce processus comprend la découverte du nombre de valeurs uniques, de valeurs Null et de modèles de données dans une colonne.
- Effectuez la découverte de domaines de données. Vous pouvez découvrir des caractéristiques de données critiques au sein d'une entreprise.
- Conserver les résultats de profil, y compris les types de données, les domaines de données, les clés primaires et les clés étrangères.
- Créer des fiches d'évaluation pour contrôler la qualité des données.
- Afficher le lignage des fiches d'évaluation pour chaque score de fiche d'évaluation et groupe de scores.
- Créer et affecter des balises aux objets de données.

- Rechercher la signification d'un nom d'objet comme un terme métier dans le bureau de glossaire métier. Par exemple, vous pouvez rechercher la signification d'un nom de colonne ou d'un nom de profil pour comprendre sa spécification dans l'entreprise et son implémentation actuelle.

Vous pouvez effectuer les tâches suivantes dans l'outil Developer :

- Découvrir le degré de jointures potentielles entre deux colonnes de données dans une source de données.
- Déterminer le pourcentage de chevauchement des données entre des paires de colonnes d'une source de données ou de plusieurs sources de données.
- Comparer les résultats du profilage de colonnes.
- Générer un objet de mappage depuis un profil.
- Découvrir des clés principale dans une source de données.
- Découvrir des clés étrangères dans un ensemble d'une ou plusieurs sources de données.
- Découvrir la dépendance fonctionnelle entre les colonnes d'une source de données.
- Exécuter des tâches de découverte de données sur un grand nombre de sources de données sur plusieurs connexions. Les tâches de découverte des données incluent le profil de colonne, l'inférence des relations de clé principale à clé étrangère, la découverte des domaines de données et la génération d'un résumé graphique consolidé des relations de données.

Vous pouvez effectuer les tâches suivantes dans l'outil Analyst :

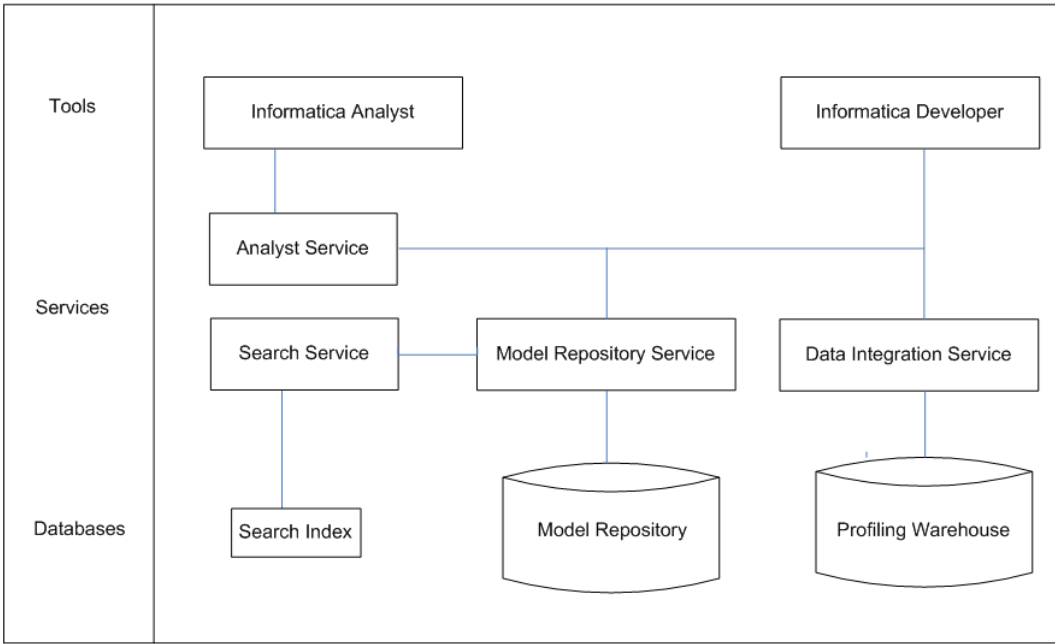
- Effectuer la découverte des données d'entreprise sur un grand nombre de sources de données sur plusieurs connexions. Vous pouvez afficher un résumé consolidé des résultats de découverte de métadonnées de colonne et de domaines de données.
- Effectuer une recherche de découverte pour trouver où les données et métadonnées existent dans l'entreprise. Vous pouvez rechercher des ressources spécifiques, telles que des objets de données, des règles et des profils. La découverte de recherche trouve les ressources et identifie les relations avec d'autres ressources des bases de données et des schémas de l'entreprise.

Architecture de profilage

L'architecture de profilage comprend des outils, des services et des bases de données. Le composant Outils comprend les applications clients. Le composant Services comprend des services d'applications nécessaires

pour gérer les outils, effectuer les tâches d'intégration de données et gérer les métadonnées des objets de profils. Le composant Base de données comprend le référentiel modèle et l'entrepôt de profilage.

La figure suivante montre les composants de l'architecture pour le profilage :



Lors de l'exécution d'un profil, le service Analyst ou l'outil Developer reçoit la définition du profil depuis le Service de Référentiel Modèle. Puis, le service Analyst ou l'outil Developer appelle le plug-in de profilage dans le Data Integration Service. Ensuite, le plug-in de profilage traite la tâche de profilage et l'envoie au Data Integration Service. Le Data Integration Service génère les résultats du profilage. Le service d'intégration de données écrit ensuite les résultats de profilage dans l'entrepôt de profilage.

La recherche de découverte utilise le service de recherche. Le service de recherche effectue chaque recherche sur un index de recherche au lieu du référentiel modèle ou de l'entrepôt de profilage. Le service de recherche génère l'index de recherche en fonction du contenu du référentiel modèle et de l'entrepôt de profilage. Il contient des extracteurs qui permettent d'extraire le contenu de chaque référentiel.

Le tableau suivant décrit les composants de l'architecture :

Composant	Description
Informatica Analyst	Une application client Internet que vous pouvez utiliser pour découvrir, analyser les données et métadonnées des sources de données et établir le rapport correspondant.
Informatica Developer	Application client que vous utilisez pour effectuer une découverte de données avancée, telles que la découverte de clé primaire, la découverte de clé étrangère et la découverte des données d'entreprise.
Service Analyst	Un service d'application qui exécute l'outil Analyst et gère les connexions entre les composants de service et les utilisateurs de l'outil Analyst.
Service de recherche	Service d'application qui gère la recherche dans l'outil Analyst. Par défaut, le service de recherche renvoie les résultats de recherche d'un référentiel modèle, tels que des objets de données, des profils, des spécifications de mappage, des tables de référence, des règles et des fiches d'évaluation.

Composant	Description
Index de recherche	Système de fichiers dans un répertoire personnalisé qui stocke du contenu indexé que le service de recherche extrait depuis le référentiel modèle et l'entrepôt de profilage.
Service de Référentiel Modèle	Service d'application qui gère le référentiel modèle.
Datenintegrationsdienst	Un service d'application qui effectue des tâches d'intégration de données pour l'outil Analyst, l'outil Developer et les clients externes.
Référentiel modèle	Une base de données relationnelle qui stocke les métadonnées pour les projets créés dans l'outil Analyst ou Developer.
Entrepôt de profilage	Une base de données qui stocke les informations de profilage, telles que les résultats de profils et de fiches d'évaluation.

Processus de découverte de données

Lorsque vous commencez un projet d'intégration de données, le profilage est souvent la première étape. Vous pouvez créer des profils pour analyser le contenu, la qualité et la structure des sources de données. Dans le cadre du processus de profilage, vous devez découvrir les métadonnées des sources de données.

Vous devez utiliser différents profils pour différents types d'analyse de données, tels qu'un profil de colonne, la découverte de clé primaire, la découverte de clé étrangère et la découverte de domaines de données. Vous devez découvrir et documenter les problèmes de qualité des données. Effectuez les tâches suivantes pour réaliser la découverte de données :

1. Rechercher et analyser le contenu des données dans les sources de données. Comprend les types de données, la fréquence des valeurs, la fréquence des formes et les statistiques de données, telles que les valeurs minimum et maximum.
2. Découvrir la structure des données. Comprend des clés, des dépendances fonctionnelles et des clés étrangères.
3. Vérifier et valider les résultats de profil.
4. Développer les résultats de profil.
5. Conserver les résultats de profil.
6. Créer des données de référence.
7. Documenter les problèmes de données.
8. Créer et exécuter les règles.
9. Créer des fiches d'évaluation pour contrôler la qualité des données.

Vous pouvez utiliser les outils suivants pour gérer le processus de découverte :

Informatica Administrator

Gérez les utilisateurs, groupes, privilèges et rôles. Vous pouvez administrer le service Analyst et gérer les autorisations pour les projets et objets dans Informatica Analyst. Vous pouvez contrôler les autorisations d'accès dans Informatica Developer à l'aide de cet outil.

Informatica Developer

Créez et exécutez les profils dans cet outil pour rechercher et analyser les métadonnées d'une ou plusieurs sources de données, y compris la découverte des relations entre les colonnes. Vous devez créer des profils à l'aide d'un assistant.

Informatica Analyst

Vous pouvez exécuter un profil de colonne, effectuer la découverte de domaines de données et procéder à la découverte des données d'entreprise sur les objets de données dans l'outil Analyst. Après avoir exécuté un profil, vous pouvez développer les lignes de données dans une source de données.

CHAPITRE 2

Découverte de données

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la découverte de données, 22](#)
- [Types de profil et d'analyse, 22](#)
- [Composants de profilage, 24](#)
- [Résultats de profil, 24](#)

Présentation de la découverte de données

La découverte de données est le processus de découverte des métadonnées des systèmes sources qui incluent du contenu et de la structure. Le contenu désigne les valeurs de données, les fréquences et les types de données. La structure comprend les clés candidates, les clés principales, les clés étrangères et les dépendances fonctionnelles. Vous pouvez créer et exécuter des profils pour découvrir le contenu et la structure des sources de données.

Vous pouvez définir un profil pour analyser des données dans un seul objet de données ou dans plusieurs objets de données. Ajoutez des commentaires aux profils pour pouvoir suivre efficacement le processus de profilage.

Exécutez un profil pour évaluer la structure des données et pour vérifier que les colonnes de données contiennent les types d'informations que vous attendez. Vous pouvez développer des lignes de données dans les données profilées. Si les résultats de profils révèlent des problèmes dans les données, vous pouvez appliquer des règles pour corriger l'ensemble de résultats. Vous pouvez créer des fiches d'évaluation pour suivre et mesurer la qualité des données avant et après avoir appliqué les règles. Si les métadonnées de la source externe d'un profil ou d'une fiche d'évaluation changent, vous pouvez synchroniser les modifications avec son objet de données.

Types de profil et d'analyse

Créez un profil selon le type d'analyse à effectuer. Le type de profil que vous créez correspond au type d'analyse que vous effectuez. Par exemple, créez un profil de clé principale pour effectuer une analyse de clé principale.

Vous pouvez créer les profils suivants pour effectuer l'analyse et la découverte de données :

Profil de colonne

Analyse la qualité des données dans des colonnes sélectionnées d'une table ou d'un fichier. Vous pouvez définir des profils pour l'analyse des colonnes dans l'outil Analyst et l'outil Developer.

Découverte de domaines de données

Découvre les caractéristiques techniques des données critiques d'une entreprise. La découverte des domaines de données identifie tous les domaines de données associés à une colonne en fonction de la valeur ou du nom de la colonne. Dans le cadre du processus de découverte, vous pouvez créer manuellement les règles de données et de noms de colonne pour vérifier si une valeur ou un nom de colonne appartient à un domaine de données. Vous pouvez ensuite associer ces règles lorsque vous créez un domaine de données. Vous pouvez également créer les domaines de données à partir des valeurs et des modèles des résultats du profil de colonne.

Profil de clé principale

Découvre les relations de clés principales entre des colonnes d'une table ou d'un fichier. Vous pouvez définir des profils pour l'analyse de clé principale dans l'outil Developer.

Profil de dépendance fonctionnelle

Découvre les dépendances fonctionnelles entre les colonnes d'une table ou d'un fichier. Vous pouvez définir des profils pour l'analyse de dépendance fonctionnelle dans l'outil Developer.

Profil de clé étrangère

Découvre les relations de clés étrangères entre les colonnes de plusieurs tables ou de plusieurs fichiers. Vous pouvez définir des profils pour l'analyse de clé étrangère dans l'outil Developer.

Profil de jointure

Détermine le degré de jointures potentielles entre des colonnes dans une source de données ou dans plusieurs sources de données. Vous pouvez définir des profils pour l'analyse de jointure dans l'outil Developer. Les résultats s'affichent dans le diagramme de Venn.

Découverte du chevauchement

Détermine le pourcentage de chevauchement des données entre des paires de colonnes d'une source de données ou de plusieurs sources de données. Vous pouvez exécuter la tâche de découverte de chevauchement à partir de l'éditeur dans l'outil Developer. Vous pouvez valider les résultats et les afficher dans un diagramme de Venn.

Découverte des données d'entreprise

Exécute plusieurs tâches de découverte des données sur un grand nombre de sources de données et génère un récapitulatif consolidé des résultats de profil. Inclut l'exécution d'un profil de colonne, une découverte de domaines de données et la découverte des relations de clé principale et de clé étrangère. La découverte des données d'entreprise automatise le traitement du profil pour un grand nombre de sources de données.

Remarque: Les modifications apportées aux profils dans l'outil Analyst ne s'affichent pas dans l'outil Developer avant l'actualisation de la connexion de l'outil Developer au référentiel. Déconnectez-vous du référentiel dans l'outil Developer, puis reconnectez-vous pour actualiser la connexion.

Composants de profilage

Un profil comprend plusieurs composants que vous pouvez utiliser pour analyser efficacement le contenu et la structure des sources de données.

Un profil comprend les composants suivants :

Filtre

Crée un sous-ensemble de source de données d'origine qui répond aux critères spécifiques. Vous pouvez ensuite profiler les exemples de données.

Règle

Logique métier qui définit les conditions appliquées aux données lorsque vous exécutez un profil. Ajoutez une règle au profil pour nettoyer, modifier ou valider les données.

Balise

Métadonnées qui définissent un objet dans le référentiel Modèle en fonction de l'utilisation professionnelle. Créez des balises pour grouper les objets en fonction de leur utilisation professionnelle.

Commentaire

Description concernant le profil. Utilisez des commentaires pour partager des informations sur les profils avec d'autres utilisateurs des outils Analyst et Developer.

Fiche d'évaluation

Représentation graphique de valeurs valides pour une colonne ou la sortie d'une règle dans les résultats de profil. Utilisez des fiches d'évaluation pour mesurer la progression de la qualité des données.

Résultats de profil

Vous pouvez afficher les résultats de profil après avoir exécuté un profil. Vous pouvez afficher un résumé, des valeurs, des formes et des statistiques pour les colonnes et les règles dans le profil. Vous pouvez afficher les propriétés pour les colonnes et les règles dans le profil. Vous pouvez prévisualiser les données de profil.

La table suivante décrit les résultats de profil pour chaque type de profil :

Type de profil	Résultats
Profil de colonne	<ul style="list-style-type: none">- Nombre et pourcentage de valeurs uniques et de valeurs Null dans les colonnes ainsi que les types de données inférés pour les valeurs de colonnes.- Modèles de fréquence et de caractère des valeurs de données dans une colonne sélectionnée et un résumé des statistiques pour la colonne.- Graphiques en secteurs qui représentent les fréquences de valeur et les fréquences de forme.- Types de données inférés par l'analyse des données de colonne.- Type de données documenté pour les données.- Valeurs maximum et minimum.- Date et heure de dernière exécution du profil.
Profil de clé principale	<ul style="list-style-type: none">- Nombre et pourcentage des valeurs uniques, dupliquées et Null pour les candidats de clés principales inférées.- Nombre de violations de clés dans les candidats de clés principales inférées.

Type de profil	Résultats
Profil de dépendance fonctionnelle	<ul style="list-style-type: none"> - Dépendances fonctionnelles inférées. - Nombre de violations de dépendances fonctionnelles.
Profil de clé étrangère	<ul style="list-style-type: none"> - Colonnes de clés principales et étrangères qui répondent aux critères d'inférence principale-étrangère que vous définissez. - Nombre de valeurs de données qui correspondent entre les clés principales et étrangères, exprimé en pourcentage. - Type de relation défini pour les colonnes de clés principales et étrangères avant d'exécuter le profil.
Profil de jointure	<ul style="list-style-type: none"> - Diagramme de Venn qui illustre les relations entre les colonnes. - Nombre et pourcentage des valeurs orphelines, Null et jointes dans les colonnes.
Découverte du chevauchement	<ul style="list-style-type: none"> - Pourcentage de chevauchement entre deux colonnes. - Diagramme de Venn qui illustre les relations entre les colonnes.
Découverte de domaines de données	<ul style="list-style-type: none"> - Nom et données des colonnes qui correspondent aux domaines de données prédéfinis, exprimés sous forme de pourcentage. - Groupe de domaines de données auquel appartient la colonne et son type de données.
Découverte des données d'entreprise	<ul style="list-style-type: none"> - Résultats de profil de colonne. - Résultats de la découverte des domaines de données. - Résultats de la découverte de clé principale. - Résultats de profil de clé étrangère dans les vues graphique et tabulaire.

Vous pouvez utiliser les outils de rapport tiers pour lire les résultats de profil depuis l'entrepôt de profils. Informatica fournit un ensemble de vues de profil que vous pouvez personnaliser pour les statistiques de profil que vous voulez lire. Ces vues sont basées sur les types communs de statistiques de profil et l'analyse des résultats de profil.

CHAPITRE 3

Concepts de profil de colonne

Ce chapitre comprend les rubriques suivantes :

- [Présentation des concepts de profil de colonne](#)[Présentation des profils, 26](#)
- [Options de profil de colonne, 27](#)
- [Règles, 27](#)
- [Fiches d'évaluation, 28](#)

Présentation des concepts de profil de colonne

Présentation des profils

Un profil de colonne détermine les caractéristiques des colonnes dans une source de données, telles que la fréquence des valeurs, les pourcentages et les formes.

Le profilage de colonnes détecte les points suivants concernant les données :

- Le nombre de valeurs uniques et null dans chaque colonne, exprimé sous forme de nombre et de pourcentage.
- La forme des données dans chaque colonne et la fréquence d'exécution de ces valeurs.
- Les statistiques sur les valeurs de colonne, telles que les longueurs maximum et minimum des valeurs et les premières et dernières valeurs dans chaque colonne.

Utilisez les options de profil de colonne pour sélectionner les colonnes dans lesquelles exécuter un profil et définir les options d'échantillonnage et de développement des données lors de la création du profil.

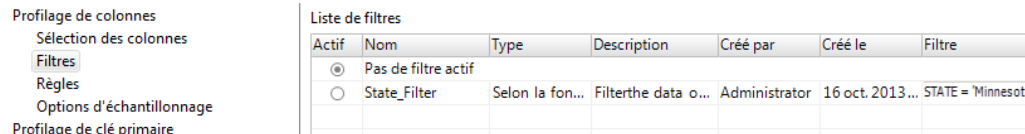
Une règle est une logique métier qui définit les conditions appliquées aux données source lorsque vous exécutez un profil. Vous pouvez ajouter une règle au profil pour valider les données.

Créez des fiches d'évaluation pour examiner périodiquement la qualité des données. Vous devez créer des fiches d'évaluation avant et après d'avoir appliqué des règles aux profils pour pouvoir afficher une représentation graphique des valeurs valides pour les colonnes.

Options de profil de colonne

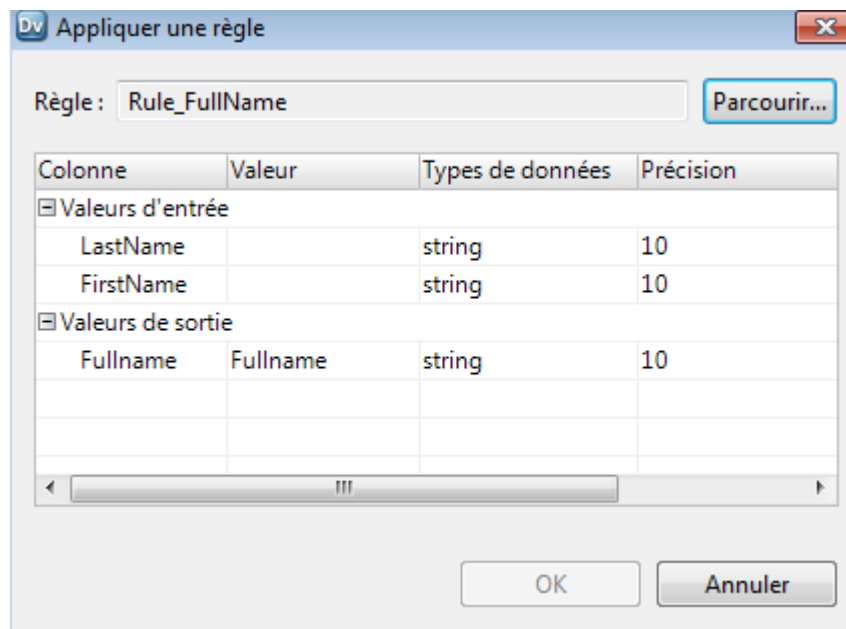
Quand vous créez un profil avec l'option **Profilage de colonnes**, vous pouvez utiliser l'assistant Création de profil pour définir les options de filtrage et d'échantillonnage. Ces options déterminent la manière dont le profil lit les lignes à partir de l'ensemble de données.

La figure suivante montre un exemple de définition de filtre dans un profil :



Après avoir effectué les étapes de l'assistant Création de profil, vous pouvez ajouter une règle au profil. La règle peut comporter la logique métier permettant d'effectuer les opérations de transformation de données avant le profilage de colonnes.

La figure suivante montre une règle intitulée Rule_FullName qui fusionne les colonnes LastName et FirstName dans la colonne FullName :



Règles

Créez et appliquez des règles dans les profils. Une règle est une logique métier qui définit les conditions appliquées aux données lorsque vous exécutez un profil. Utilisez les règles pour valider les données dans un profil et mesurer la progression de la qualité des données.

Vous pouvez ajouter une règle après avoir créé un profil. Vous pouvez réutiliser les règles créées dans l'outil Analyst ou Developer indifféremment dans les deux outils. Ajoutez des règles à un profil en sélectionnant une règle réutilisable ou créez une règle d'expression. Une règle d'expression utilise les fonctions d'expression et

les colonnes pour définir la logique de règle. Après avoir créé une règle d'expression, vous pouvez la rendre réutilisable.

Vous pouvez ajouter une règle après avoir créé un profil. Ajoutez des règles à un profil en sélectionnant une règle réutilisable. Pour créer une règle, validez un mapplet en tant que règle.

Créez des règles d'expression dans l'outil Analyst. Dans l'outil Developer, vous pouvez créer une mapplet et la valider en tant que règle. Vous pouvez exécuter les règles depuis les outils Analyst et Developer.

Fiches d'évaluation

Une fiche d'évaluation est la représentation graphique des valeurs valides d'une colonne ou la sortie d'une règle dans les résultats de profil. Utilisez des fiches d'évaluation pour mesurer la progression de la qualité des données. Vous pouvez créer une fiche d'évaluation depuis un profil et surveiller la progression de la qualité des données au fil du temps.

Une fiche d'évaluation comprend plusieurs composants, tels que des mesures, des groupes de mesures et des seuils. Après avoir exécuté un profil, vous pouvez ajouter des colonnes source comme mesures à une fiche d'évaluation et configurer les valeurs valides des mesures. Les fiches d'évaluation aident l'entreprise à mesurer la valeur de la qualité des données en suivant le coût des données incorrectes au niveau de la mesure et de la fiche d'évaluation. Pour mesurer le coût des données incorrectes pour chaque mesure, affectez une unité coût à la mesure et définissez un coût fixe ou variable. Lorsque vous exécutez la fiche d'évaluation, les résultats de la fiche d'évaluation incluent le coût des données incorrectes pour chaque mesure et la valeur du coût total pour toutes les mesures.

Utilisez un groupe de mesures pour classer les mesures associées d'une fiche d'évaluation dans un ensemble. Un seuil détermine la plage, en pourcentage, de données incorrectes acceptables pour les colonnes dans un enregistrement. Vous pouvez définir des seuils pour des plages de données satisfaisantes, acceptables ou inacceptables.

Lorsque vous exécutez une fiche d'évaluation, indiquez si vous souhaitez développer les mesures de score sur les données en direct ou les données stockées. Après avoir exécuté une fiche d'évaluation et affiché les scores, vous pouvez développer chaque mesure pour identifier des enregistrements de données valides et des enregistrements non valides. Vous pouvez également afficher le lignage des fiches d'évaluation de chaque métrologie ou groupe de métrologies dans une fiche d'évaluation. Pour suivre la qualité des données de manière efficace, vous pouvez utiliser des graphes de tendance de score et des graphes de tendance de coût. Ces graphes surveillent la manière dont évoluent les scores et le coût des données incorrectes sur une période.

L'entrepôt de profilage stocke les statistiques des fiches d'évaluation et les informations de configuration. Vous pouvez configurer une application tierce pour récupérer les résultats de la fiche d'évaluation et pour exécuter des rapports. Vous pouvez également afficher les résultats des fiches d'évaluation dans une application Web, un portail ou un rapport tel que le rapport d'aide à la décision.

CHAPITRE 4

Concepts de découverte des domaines de données

Ce chapitre comprend les rubriques suivantes :

- [Présentation des concepts de découverte des domaines de données, 29](#)
- [Domaines de données, 30](#)
- [Groupes de domaines de données, 30](#)
- [Glossaire de domaine de données, 31](#)
- [Processus de découverte des domaines de données, 31](#)

Présentation des concepts de découverte des domaines de données

Vous devez identifier et comprendre la signification des données source critiques pour être en mesure de travailler efficacement. La découverte des domaines de données est le processus de découverte de la signification fonctionnelle des données dans les sources de données en fonction de la sémantique des données.

Créez un profil pour effectuer une découverte des domaines de données et identifiez les caractéristiques des données critiques au sein d'une entreprise. Vous pouvez ensuite appliquer aux données d'autres règles de gestion, comme la qualité des données ou le masquage des données. Par exemple, découvrez les codes ou descriptions des produits afin d'analyser les règles de normalisation ou d'analyse de la qualité des données qui vous sont nécessaires pour rendre les données utiles et fiables. Autre exemple : vous pouvez rechercher les données sensibles des clients comme les numéros de carte de crédit, les identifiants d'e-mail et les numéros de téléphone. Vous pouvez ensuite vouloir masquer ces informations pour les protéger.

Vous pouvez créer et exécuter un profil pour effectuer la découverte des domaines de données dans les outils Analyst et Developer. Vous pouvez définir un profil pour effectuer la découverte des domaines de données en fonction des règles suivantes :

- Règle de données. Recherche les colonnes dont les données correspondent à une logique spécifique définie dans la règle.
- Règle de nom de colonne. Recherche les colonnes qui correspondent à une logique de nom de colonne spécifique définie dans la règle.

Vous pouvez créer des domaines de données à partir des valeurs et des modèles dans les résultats de profil de colonne. Vous pouvez ensuite utiliser ces domaines de données pour découvrir des données essentielles parmi plusieurs systèmes de données ou au sein de l'entreprise.

Domaines de données

Un domaine de données est un objet du référentiel modèle défini par l'utilisateur ou prédéfini basé sur les sémantiques des données de colonne ou d'un nom de colonne. Par exemple, le numéro de sécurité sociale, le numéro de carte de crédit, l'identifiant de messagerie et le numéro de téléphone peuvent être des domaines de données individuels.

Un domaine de données vous permet de rechercher des données importantes qui n'ont pas découvertes dans une source de données. Par exemple, vous pouvez avoir des systèmes de données héritées contenant des numéros de sécurité sociale dans un champ Commentaires. Vous devez rechercher ces informations et les protéger avant de les déplacer dans de nouveaux systèmes de données.

Vous pouvez regrouper des domaines de données logiques dans plusieurs groupes de domaines de données. Un glossaire de domaine de données répertorie tous les domaines de données et les groupes de domaines de données. Utilisez le menu Préférences dans l'outil Developer pour importer et exporter les domaines de données depuis et vers le glossaire de domaine de données.

Vous utilisez des règles pour définir les modèles de données et de nom de colonne qui correspondent aux données source et aux métadonnées. Lorsque vous créez un domaine de données, l'outil Analyst ou l'outil Developer copie les règles associées et autres objets dépendants dans le glossaire de domaine de données. Utilisez l'outil Developer pour gérer les domaines de données, y compris l'importation et l'exportation des domaines de données depuis et vers le glossaire de domaine de données. Vous pouvez également utiliser l'outil Developer pour gérer la logique de règle des domaines de données.

Remarque: Vous pouvez vouloir enregistrer toutes les règles de domaine de données dans un projet ou dossier unique. Cette étape est utile lorsque vous avez exporté les domaines de données, et que vous devez modifier les règles et autres objets de données associés.

Groupes de domaines de données

Les groupes de domaines de données vous permettent de classer les domaines de données dans des groupes spécifiques. Par exemple, vous pouvez regrouper les domaines de données first_name, last_name et account_number dans le groupe de domaines de données Informations médicales personnelles (PHI).

Vous pouvez créer un groupe de domaines de données Informations personnelles identifiables (PII) qui inclut le numéro de sécurité sociale, le prénom et le nom. Un domaine de données peut faire partie de plusieurs groupes de domaines de données. Par exemple, le numéro de sécurité sociale peut appartenir aux groupes de domaines de données Industrie des cartes de paiement (PCI) et PII. Les groupes de domaines de données peuvent contenir des domaines de données et non d'autres groupes de domaines de données.

Remarque: Si vous importez le fichier de domaine de données `Informatica_IDE_DataDomain.xml` après l'installation, le glossaire de domaine de données affiche les groupes de domaines de données et les domaines de données. Vous pouvez ensuite créer des groupes de domaines de données supplémentaires si nécessaire. Pour afficher et modifier les règles associées aux domaines de données, importez le fichier `Informatica_IDE_DataDomainRule.xml`.

Glossaire de domaine de données

Le glossaire de domaine de données est un conteneur pour tous les groupes de domaines et les domaines de données. Vous pouvez utiliser le glossaire de domaine de données pour créer, gérer et supprimer les domaines de données et groupes de domaines de données.

Vous pouvez rechercher des domaines et des groupes de domaines dans le glossaire de domaine de données. Vous pouvez également exporter des domaines de données dans un fichier XML et importer des domaines de données depuis un fichier XML dans le glossaire de domaine de données. Le glossaire de domaine de données contient des règles copiées et toutes les données de référence associées aux domaines de données. Vous ne pouvez pas modifier les règles dans le glossaire de domaine de données.

Vous pouvez afficher le glossaire de domaine de données depuis le menu Préférences de l'outil Developer et depuis le menu Gérer de l'outil Analyst. Utilisez les droits **Gérer les domaines de données** du Model Repository Service pour déterminer qui crée, modifie et supprime les domaines de données et les groupes de domaines de données.

Processus de découverte des domaines de données

Vous pouvez définir et exécuter un profil pour exécuter la découverte des domaines de données dans l'outil Analyst ou Developer en fonction de votre rôle de travail. Après avoir configuré les options de découverte des domaines de données et exécuté le profil, vous pouvez vérifier et développer les résultats. Si vous exécutez la découverte de domaines de données depuis l'éditeur, vous pouvez ajouter les résultats à un modèle de données.

Procédez comme suit pour effectuer la découverte des domaines de données :

1. Créez ou importez des domaines de données et des groupes de domaines.
2. Éventuellement, consolidez les domaines de données dans les groupes de domaines appropriés.
3. Créez un profil pour effectuer la découverte des domaines de données. Vous commencez par choisir si vous voulez exécuter un profil de colonne avec la découverte des domaines de données ou seulement la découverte des domaines de données.
4. Sélectionnez les colonnes, les domaines et les options d'échantillonnage appropriées.
5. Exécutez le profil.
6. Vérifiez, développez les résultats de profil et ajoutez-les à un modèle de données si nécessaire.

CHAPITRE 5

Concepts de conservation

Ce chapitre comprend les rubriques suivantes :

- [Présentation des concepts de conservation, 32](#)
- [Conservation pour les analystes et les développeurs, 32](#)
- [Tâches de conservation, 33](#)

Présentation des concepts de conservation

La conservation est le processus de validation et de gestion des métadonnées d'une source de données découvertes, de sorte à les préparer afin d'être utilisées et de figurer dans des rapports.

Vous pouvez conserver les résultats de profil inférés suivants :

- Type de données
- Domaines de données
- Clés primaires
- Clés étrangères

Conserver les résultats de profil inférés vous permet de rendre les métadonnées concernant les colonnes, les domaines de données et les relations d'objet de données plus précises dans les bases de données et les schémas. Vous trouverez ainsi les métadonnées les plus pertinentes lorsque vous lancerez une recherche de découverte sur plusieurs référentiels ou lorsque vous afficherez le diagramme des relations de clé étrangère dans les résultats de découverte des données d'entreprise.

Vous pouvez conserver des inférences de métadonnées spécifiques générées lors de l'exécution d'un profil. Par exemple, vous pouvez approuver ou rejeter les types de données inférés dans les résultats de profil de colonne et dans ceux de la découverte de domaines de données. Vous pouvez également approuver ou rejeter les clés primaires et étrangères inférées dans les résultats de la découverte des données d'entreprise.

Conservation pour les analystes et les développeurs

Si vous êtes analyste ou gestionnaire de données, vous pouvez conserver les résultats de profil de colonne et dans ceux de la découverte de domaines de données dans l'outil Analyst. La conservation des résultats de

profil permet de disposer d'informations de profil précises pour la recherche de découverte et la validation des ressources de données.

En tant que développeur ou architecte de données, vous pouvez conserver les résultats de profil de colonne, de découverte de domaines de données, de découverte de clé primaire et de découverte de clé étrangère dans l'outil Developer.

Exemples de conservation

Lorsque vous lancez la découverte des données d'entreprise en tant que développeur, l'outil Developer traite les domaines de données sélectionnés pour la totalité de l'ensemble de données. Cette action peut entraîner des inférences de domaines de données multiples, par exemple, entre le domaine du numéro de téléphone et celui du numéro de sécurité sociale. Celles-ci ont lieu lorsque certaines données d'une colonne correspondent à différents domaines de données. Par exemple, un numéro de téléphone à 10 chiffres auquel il manquerait un chiffre peut être assimilé au modèle de numéro de sécurité sociale. Cela peut indiquer des problèmes de qualité des données présentes dans une colonne ou un modèle analogue dans plusieurs domaines de données. Dans ce cas, l'outil Developer peut inférer les domaines de données de numéro de téléphone et de numéro de sécurité sociale. Vous pouvez conserver les résultats de profil de façon à sélectionner et à approuver le domaine de données le plus approprié. Dans l'exemple précédent, le domaine de données du numéro de téléphone est correct, car l'inférence du domaine de données du numéro de sécurité sociale est due à un problème de qualité des données.

Lorsque vous exécutez la découverte des données d'entreprise, l'outil Developer peut inférer plusieurs types de données (Date, String, Varchar, etc.) pour une colonne de date. En tant qu'architecte de données, vous souhaitez peut-être choisir et approuver le type de données Date, qui est le plus pertinent pour une colonne de date.

Il peut arriver que le processus de découverte des données d'entreprise dans l'outil Developer infère toutes les relations d'objet de données à partir des données de la colonne. Elles peuvent inclure des relations indésirables dans les clés candidates découvertes. Par exemple, il est possible que l'outil Developer infère des colonnes représentant une séquence en tant que clés éventuelles et découvre des relations avec d'autres tables présentant des colonnes similaires. Ces relations d'objet de données peuvent ne pas constituer des relations valides dans la base de données. Dans ce cas, vous pouvez évaluer, vérifier et approuver les résultats de profil inféré les plus appropriés dans le cadre de la conservation.

Tâches de conservation

Vous pouvez conserver les résultats de profil après exécution de ce dernier. Vous pouvez également annuler une décision en matière de conservation prise au moment de l'exécution du profil.

Vous pouvez effectuer les tâches de conservation suivantes dans l'outil Analyst :

- Approuver ou rejeter les types de données inférés pour plusieurs colonnes et domaines de données.
- Restaurer le statut inféré de types de données approuvés ou rejetés.
- Restaurer le statut inféré de domaines de données approuvés ou rejetés.
- Afficher ou masquer les lignes de résultat rejetées.
- Exclure des colonnes des exécutions de profil en fonction des préférences de métadonnées, comme des types de données ou domaines de données approuvés.

Vous pouvez effectuer les tâches de conservation suivantes dans l'outil Developer :

- Approuver ou rejeter les types de données inférés pour plusieurs colonnes.
- Restaurer le statut inféré de types de données approuvés ou rejetés.

- Restaurer le statut inféré de domaines de données approuvés ou rejetés.
- Afficher ou masquer les lignes de résultat rejetées.
- Approuver ou rejeter les objets de données dans les résultats de découverte de clé primaire.
- Approuver ou rejeter les résultats de découverte des données d'entreprise, y compris les résultats de découverte de clé étrangère.
- Exclure des colonnes des exécutions de profil en fonction des préférences de métadonnées, comme des types de données ou domaines de données approuvés.

Partie II : Découverte de données avec Informatica Analyst

Cette partie contient les chapitres suivants :

- [Profils de colonne dans Informatica Analyst, 36](#)
- [Résultats de profil de colonne dans Informatica Analyst, 44](#)
- [Règles dans Informatica Analyst, 54](#)
- [Fiches d'évaluation dans Informatica Analyst, 58](#)
- [Découverte des domaines de données dans Informatica Analyst, 75](#)
- [Découverte des données d'entreprise dans Informatica Analyst, 86](#)
- [Résultats de la découverte des données d'entreprise dans Informatica Analyst, 91](#)
- [Recherche de découverte dans Informatica Analyst, 96](#)
- [Bureau de glossaire métier dans Informatica Analyst, 106](#)

CHAPITRE 6

Profils de colonne dans Informatica Analyst

Ce chapitre comprend les rubriques suivantes :

- [Présentation des profils de colonne dans Informatica Analyst, 36](#)
- [Processus de profilage de colonne, 37](#)
- [Options de profil, 37](#)
- [Création d'un profil de colonne dans l'outil Analyst, 39](#)
- [Modification d'un profil de colonne, 40](#)
- [Exécution d'un profil, 41](#)
- [Création d'un filtre, 41](#)
- [Gestion des filtres, 42](#)
- [Synchronisation d'un objet de données fichier plat, 42](#)
- [Synchronisation d'un objet de données relationnel, 43](#)

Présentation des profils de colonne dans Informatica Analyst

Lors de la création d'un profil, sélectionnez les colonnes dans l'objet de données pour lequel profiler les données. Vous pouvez définir ou configurer des options d'échantillonnage et de développement pour un profilage plus rapide. Une fois le profil exécuté, vous pouvez consulter les statistiques de profilage pour comprendre les données.

Vous pouvez profiler des tables larges et des fichiers plats comprenant un grand nombre de colonnes. Vous pouvez profiler des tables avec plus de 30 colonnes et des fichiers plats avec plus de 100 colonnes. Quand vous créez ou exécutez un profil, vous pouvez choisir de sélectionner toutes les colonnes ou sélectionner chaque colonne à inclure pour le profilage. L'outil Analyst affiche les 30 premières colonnes dans l'aperçu des données. Vous pouvez sélectionner toutes les colonnes pour le développement et afficher les fréquences de valeur pour ces colonnes. Vous pouvez utiliser des règles qui comprennent plus de 50 champs de sortie et inclure les colonnes de règle pour le profilage lors de la réexécution du profil.

Processus de profilage de colonne

Dans le cadre du processus de profilage de colonnes, vous pouvez choisir d'inclure toutes les colonnes source pour le profilage ou de sélectionner des colonnes spécifiques. Vous pouvez également accepter les options de profil par défaut ou configurer les résultats de profil, l'échantillonnage et les options de développement.

Les étapes suivantes décrivent le processus de profilage de colonne :

1. Sélectionnez l'objet de données sur lequel vous souhaitez exécuter un profil.
2. Sélectionnez une source externe ou un objet de données importé sur lequel vous souhaitez exécuter le profil.
3. Déterminez si vous souhaitez créer un profil avec les options par défaut ou modifier les options de profil par défaut.
4. Choisissez l'emplacement d'enregistrement souhaité pour le profil.
5. Sélectionnez les colonnes sur lesquelles vous souhaitez exécuter le profil.
6. Sélectionnez l'option des résultats de profil.
7. Choisissez les options d'échantillonnage.
8. Sélectionnez les options d'exploration et choisissez l'environnement natif ou Hive pour exécuter le profil.
9. Vous pouvez éventuellement ajouter des règles et des filtres au profil.
10. Exécutez le profil.

Remarque: Tenez compte des règles et instructions suivantes pour les noms de colonnes et le profilage des données multilingues et Unicode :

- Vous ne pouvez pas ajouter une colonne à un profil si le nom de la colonne et le nom du profil sont identiques. Vous ne pouvez pas ajouter deux fois la même colonne à un profil même si vous changez le nom de la colonne.
- Vous pouvez profiler des données multilingues provenant de différentes sources et afficher les résultats de profil selon les paramètres régionaux dans le navigateur. L'outil Analyst change les types de données date/heure, numériques et décimales selon les paramètres régionaux de votre navigateur.
- Tri des données multilingues. Vous pouvez trier des données multilingues. L'outil Analyst affiche l'ordre de tri selon les paramètres régionaux du navigateur.
- Pour profiler les données Unicode dans une base de données DB2, définissez la variable d'environnement de base de données DB2CODEPAGE dans la base de données et redémarrez le service d'intégration de données.

Options de profil

Les options de profil comprennent l'option de résultats de profil, les options d'échantillonnage et de développement de données. Vous pouvez configurer ces options lors de la création d'un profil de colonne pour un objet de données.

Vous devez utiliser l'assistant **Nouveau profil** pour configurer les options de profils. Vous pouvez choisir de créer un profil avec les options par défaut pour les options de colonnes, d'échantillonnage et de développement. Lors de la création d'un profil pour plusieurs sources de données, l'outil Analyst utilise les options de profilage de colonnes par défaut.

Option de résultats de profils

Vous pouvez choisir d'ignorer les résultats de profils précédents ou d'afficher les résultats pour les exécutions de profils précédents.

Le tableau suivant décrit l'option de résultats de profils :

Option	Description
Afficher les résultats des colonnes et des règles que vous avez sélectionnées dans l'exécution du profil actuel	Ignore les résultats de profil des colonnes dans lesquelles vous avez exécuté un profil et affiche les résultats des colonnes et des règles sélectionnées pour la dernière exécution du profil. Ne sélectionnez pas cette option si vous souhaitez que l'outil Analyst affiche les résultats de profil dans les colonnes dans lesquelles vous avez exécuté un profil.

Options d'échantillonnage

Les options d'échantillonnage déterminent le nombre de lignes sur lesquels l'outil Analyst choisit d'exécuter un profil. Vous pouvez configurer les options d'échantillonnage lorsque vous définissez ou exécutez un profil.

Le tableau suivant décrit les options d'échantillonnage pour un profil :

Option	Description
Toutes les lignes	Choisit toutes les lignes dans l'objet de données.
<nombre> première(s) ligne(s)	Le nombre de lignes pour lesquelles vous souhaitez exécuter le profil. Analyst Tool choisit les lignes parmi les premières lignes dans la source.
Échantillon aléatoire de <nombre> lignes	Le nombre de lignes d'un échantillon aléatoire pour lequel vous souhaitez exécuter le profil. L'échantillonnage aléatoire force Analyst Tool à effectuer un développement dans les données stockées. Notez que ceci peut affecter les performances du développement.
Échantillon aléatoire	Taille de l'échantillon aléatoire basée sur le nombre de lignes dans l'objet de données. L'échantillonnage aléatoire force Analyst Tool à effectuer un développement dans les données stockées. Notez que ceci peut affecter les performances du développement.
Exclure l'inférence du type de données pour les colonnes avec un type de données approuvé	Exclut les colonnes qui comprennent un type de données approuvé de l'inférence du type de données de l'exécution du profil.

Options de développement

Vous pouvez configurer les options de développement lorsque vous définissez ou exécutez un profil.

La table suivante décrit les options de développement pour un profil :

Options	Description
Activer le développement de la ligne	Développe les données de la ligne dans les résultats de profil. Cette option est sélectionnée par défaut.
Sélectionner colonnes	Identifie les colonnes pour le développement que vous n'avez pas sélectionnées pour le profilage.
Développement sur données en direct	Développe les données actives pour lire les données actuelles dans la source de données.
Développement sur données stockées	Développe les données stockées pour lire les données de profil stockées dans l'entrepôt de profilage.

Création d'un profil de colonne dans l'outil Analyst

Vous pouvez créer un profil personnalisé ou un profil par défaut. Lors de la création d'un profil personnalisé, vous pouvez configurer les colonnes, les lignes à échantillonner et les options d'exploration. Le profil par défaut est exécuté avec toutes les options par défaut.

1. Dans le panneau **Accueil de la découverte**, cliquez sur **Profil d'objet de données** ou sélectionnez **Nouveau > Profil d'objet de données** n'importe où dans l'outil Analyst.

L'assistant **Nouveau profil** s'affiche. L'option **Profilage de colonnes** est sélectionnée par défaut.

2. Cliquez sur **Suivant**.

3. Dans le volet **Sources**, sélectionnez un objet de données.

4. Choisissez de créer un profil par défaut ou un profil personnalisé.

- Pour créer un profil par défaut, cliquez sur **Enregistrer** ou sur **Enregistrer et exécuter**.
- Pour créer un profil personnalisé, cliquez sur **Suivant**.

5. Entrez un nom et une description facultative pour le profil.

6. Dans le volet **Dossiers**, sélectionnez le projet ou le dossier dans lequel créer le profil.

L'outil Analyst affiche le projet que vous avez sélectionné et les projets partagés qui contiennent les dossiers où vous pouvez créer le profil. Les profils contenus dans le dossier s'affichent dans le volet de droite.

7. Cliquez sur **Suivant**.

8. Dans le volet **Colonnes**, sélectionnez les colonnes dans lesquelles vous souhaitez exécuter un profil. Les colonnes comprennent toutes les règles que vous avez appliquées au profil. L'outil Analyst indique les propriétés de colonne, telles que le nom, le type de données, la précision et l'échelle, pour chaque colonne.

Sélectionnez éventuellement **Nom** pour sélectionner toutes les colonnes.

9. Dans le volet **Options d'échantillonnage**, configurez les options correspondantes.

10. Dans le volet **Options d'exploration**, configurez les options correspondantes.
Cliquez éventuellement sur **Sélectionner les colonnes** pour sélectionner les colonnes dans lesquelles développer. Dans la boîte de dialogue **Colonnes de développement**, sélectionnez les colonnes à développer, puis cliquez sur **OK**.
11. Acceptez l'option par défaut dans le volet **Option de résultats de profils**.
Lors de la première exécution du profil, l'outil Analyst affiche les résultats de profil pour toutes les colonnes sélectionnées pour le profilage.
12. Cliquez sur **Suivant**.
13. Définissez éventuellement un filtre pour le profil.
14. Cliquez sur **Suivant** pour vérifier les paramètres d'exploration de la ligne, notamment les colonnes d'aperçu à explorer.
15. Pour exécuter le profil dans l'environnement Hive, sélectionnez **Hive**, puis sélectionnez une connexion Hive. La connexion Hive permet au service d'intégration de données de communiquer avec la grappe Hadoop pour forcer l'exécution du profil depuis le service d'intégration de données vers la grappe Hadoop.
16. Cliquez sur **Enregistrer** pour créer le profil ou sur **Enregistrer et exécuter** pour créer le profil puis l'exécuter.

Modification d'un profil de colonne

Vous pouvez modifier un profil de colonne après l'avoir exécuté.

1. Dans l'espace de travail **Bibliothèque**, sélectionnez le projet qui contient le profil que vous voulez modifier.
2. Cliquez sur le profil pour l'ouvrir.
Le profil s'ouvre dans l'espace de travail **Découverte**.
3. Cliquez sur **Actions > Éditer**.
Un menu contextuel s'affiche à l'écran.
4. Sélectionnez l'une des options de menu suivantes en fonction des modifications que vous voulez effectuer :
 - **Général**. Modifiez les propriétés de base, telles que le nom, la description et le type de profil. Vous pouvez également modifier l'environnement d'exécution en **Natifou Hive**.
 - **Source de données**. Sélectionnez une autre source de données correspondante.
 - **Profilage de colonnes**. Sélectionnez les colonnes dans lesquelles vous voulez exécuter le profil, puis configurez les options d'échantillonnage et de développement requises.
 - **Filtres de profilage de colonnes**. Créez, modifiez et supprimez des filtres.
 - **Règles de profilage de colonnes**. Créez des règles ou modifiez les règles actuelles.
 - **Découverte de domaines de données**. Configurez les options de découverte de domaines de données.
5. Cliquez sur **Enregistrer** pour enregistrer les modifications ou sur **Enregistrer et exécuter** pour enregistrer les modifications puis exécuter le profil.

Exécution d'un profil

Exécutez un profil pour analyser le contenu et la structure d'une source de données et sélectionnez les colonnes et règles pour le développement. Vous pouvez développer les données actives ou stockées dans les colonnes et règles. Vous pouvez exécuter un profil dans une colonne ou une règle sans profiler à nouveau toutes les colonnes source après avoir exécuté le profil.

1. Dans le navigateur, sélectionnez le projet ou le dossier qui contient le profil à exécuter.
2. Cliquez sur le profil pour l'ouvrir.
Le profil s'affiche dans un onglet. Vérifiez les options de profil avant d'exécuter le profil.
3. Cliquez sur **Actions > Exécuter le profil**.
L'outil Analyst affiche les résultats de profil.

Création d'un filtre

Vous pouvez créer un filtre pour réaliser un sous-ensemble de la source de données d'origine qui correspond aux critères de filtre. Vous pouvez ensuite exécuter un profil sur cet échantillon de données.

1. Ouvrez un profil.
2. Cliquez sur **Actions > Éditer > Filtres de profilage de colonne** pour ouvrir la boîte de dialogue **Éditer le profil**.
Les filtres actuels s'affichent dans le panneau **Filtres**.
3. Cliquez sur **Nouveau**.
4. Entrez un nom et une description facultative.
5. Sélectionnez un type de filtre simple, avancé ou SQL.
 - **Simple**. Utilisez des opérateurs conditionnels, par exemple <, >, =, BETWEEN, et ISNULL pour chaque colonne que vous voulez filtrer.
 - **Avancé**. Utilisez les catégories de fonction, par exemple Caractère, Consolidation, Conversion, Financier, Numérique, et Nettoyage de données.
Cliquez sur le nom de la fonction dans le panneau **Fonctions** pour afficher le type de renvoi, la description et les paramètres de cette dernière. Pour inclure la fonction dans le filtre, cliquez sur le bouton flèche droite (>), et vous pouvez spécifier les paramètres dans la boîte de dialogue **Fonction**.
Remarque: Pour un filtre simple ou avancé sur une colonne de date, indiquez la condition au format AAAA/MM/JJ HH:MM:SS.
 - **SQL**. Crée des requêtes SQL. Vous pouvez créer un filtre SQL pour les sources de données relationnelles. Entrez l'expression de clause WHERE pour générer le filtre SQL. Par exemple, pour filtrer les enregistrements de la société dans la région Europe à partir d'une table Société comprenant une colonne Région, saisissez

```
Region = 'Europe'
```


dans l'éditeur.
6. Cliquez sur **Valider** pour vérifier l'expression SQL.

Gestion des filtres

Vous pouvez créer, modifier et supprimer des filtres.

1. Dans l'espace de travail **Bibliothèque**, sélectionnez le projet qui contient le profil que vous voulez filtrer.
2. Ouvrez le profil.
Le profil s'ouvre dans l'espace de travail **Découverte**.
3. Cliquez sur **Actions > Modifier > Filtres de profilage de colonnes** pour ouvrir la boîte de dialogue **Éditer le profil**.
Les filtres actuels s'affichent dans le volet **Filtres**.
4. C'est ici que vous pouvez créer, modifier ou supprimer un filtre.
 - Cliquez sur **Nouveau** pour créer un filtre.
 - Sélectionnez un filtre et cliquez sur **Modifier** pour modifier les paramètres du filtre.
 - Sélectionnez un filtre et cliquez sur **Supprimer** pour retirer le filtre.

Synchronisation d'un objet de données fichier plat

Vous pouvez synchroniser les modifications apportées à une source de données de fichier plat externe avec son objet de données dans l'outil Analyst. Utilisez l'assistant **Synchronisation d'un fichier plat** pour synchroniser les objets de données.

1. Permet d'ouvrir l'espace de travail **Bibliothèque**.
2. Dans la section **Projets**, sélectionnez un objet de données fichier plat à partir d'un projet.
L'outil Analyst affiche l'aperçu des données du fichier plat dans l'onglet **Aperçu des données**.
3. Cliquez sur l'onglet **Propriétés**.
4. Dans le menu Actions, cliquez sur **Synchroniser**.
L'assistant **Synchroniser le fichier plat** s'affiche.
5. Naviguez vers un emplacement ou entrez un chemin réseau pour importer le fichier plat.
 - Pour naviguer vers un emplacement, cliquez sur **Choisir un fichier** pour sélectionner le fichier plat depuis un répertoire auquel votre machine à accès.
 - Pour entrer un chemin réseau, sélectionnez **Entrer un chemin réseau** et configurez le chemin d'accès et le nom du fichier.
6. Cliquez sur **Suivant**.
7. Vous pouvez importer un fichier plat délimité ou de largeur fixe.
 - Pour importer un fichier plat délimité, acceptez l'option **Délimité**.
 - Pour importer un fichier plat de largeur fixe, sélectionnez l'option **Largeur fixe**.
8. Cliquez sur **Suivant**.
9. Configurez les options de fichiers plats pour le fichier plat délimité ou de largeur fixe.
10. Cliquez sur **Suivant**.
11. (Facultatif) Vous pouvez modifier les attributs de colonne.
12. Cliquez sur **Suivant**.

13. Acceptez le nom par défaut ou entrez un autre nom pour le fichier plat.
14. (Facultatif) Entrez une description.
15. Cliquez sur **Terminer**.
Un message de synchronisation vous invite à confirmer l'action.
16. Cliquez sur **Oui** pour synchroniser le fichier plat.
Un message indiquant que la synchronisation est terminée s'affiche. Pour afficher les détails des modifications des métadonnées, cliquez sur **Afficher les détails**.
17. Cliquez sur **OK**.

Synchronisation d'un objet de données relationnel

Vous pouvez synchroniser les modifications apportées à une source de données relationnelle externe avec son objet de données table. Une modification de la source de données externe comprend l'ajout, la modification et la suppression de colonnes sources et de colonnes de règle.

1. Permet d'ouvrir l'espace de travail **Bibliothèque**.
2. Dans la section **Projets**, sélectionnez un objet de données table depuis un projet.
L'outil Analyst affiche l'aperçu des données de la table dans l'onglet **Aperçu des données**.
3. Cliquez sur l'onglet **Propriétés**.
4. Dans le menu Actions, cliquez sur **Synchroniser**.
Un message vous invite à confirmer l'action.
5. Pour terminer la synchronisation, cliquez sur **Oui**.
Un message de statut de synchronisation s'affiche.
6. Un message indiquant que la synchronisation est terminée s'affiche.
Pour afficher les détails des modifications des métadonnées, cliquez sur **Afficher les détails**.
7. Cliquez sur **OK**.

CHAPITRE 7

Résultats de profil de colonne dans Informatica Analyst

Ce chapitre comprend les rubriques suivantes :

- [Présentation des résultats de profil de colonne dans Informatica Analyst, 44](#)
- [Résumé de profil, 45](#)
- [Valeurs de colonne, 46](#)
- [Formes de colonne, 47](#)
- [Types de données de colonne, 48](#)
- [Statistiques de colonne, 49](#)
- [Développement de profil de colonne, 50](#)
- [Conservation dans l'outil Analyst, 51](#)
- [Fichiers d'exportation de profil de colonne dans Informatica Analyst, 52](#)

Présentation des résultats de profil de colonne dans Informatica Analyst

Affichez les résultats de profil pour comprendre la structure des données et analyser leur qualité. Vous pouvez afficher les résultats de profil après avoir exécuté un profil. Vous pouvez afficher un résumé des colonnes et des règles dans le profil et les valeurs, les modèles, les types de données et les statistiques des colonnes et des règles.

Après avoir exécuté un profil, vous pouvez afficher les résultats de profil dans les vues **Profilage de colonnes**, **Propriétés** et **Aperçu des données**. Vous pouvez exporter les fréquences de valeur, les fréquences de modèle ou les données de développement dans un fichier CSV. Vous pouvez exporter la totalité des informations de résumé de profil dans un fichier Microsoft Excel afin d'afficher toutes les données dans un fichier pour les analyser plus précisément.

Dans la vue **Profilage de colonnes**, vous pouvez afficher les informations de résumé des colonnes d'une exécution de profil. Vous pouvez afficher les valeurs, les modèles, les types de données et les statistiques de chaque colonne dans les vues **Valeurs**, **Modèles**, **Types de données** et **Statistiques**.

L'outil Analyst affiche les règles en tant que colonnes dans les résultats de profil. Les résultats de profil d'une règle s'affichent sous la forme d'une colonne profilée. Les résultats de profil qui s'affichent dépendent des options de configuration et d'échantillonnage du profil.

Les résultats de profilage suivants s'affichent dans la vue **Profilage de colonnes** :

- Informations de résumé de l'exécution de profil, y compris le nombre de valeurs uniques et Null, les types de données inférés avec le pourcentage de conformité maximal, ainsi que la date et l'heure de la dernière exécution de profil.
- Les valeurs des colonnes et la fréquence d'affichage de la valeur pour la colonne. La fréquence s'affiche sous la forme d'un nombre, d'un pourcentage et d'un graphique.
- Les modèles de valeur pour les colonnes profilées et la fréquence d'affichage du modèle. La fréquence s'affiche sous la forme d'un nombre et d'un pourcentage.
- Graphique à barres qui représente le pourcentage de lignes avec des valeurs Null, des valeurs uniques et des valeurs non unique dans une colonne.
- Graphiques en secteurs qui affichent les fréquences de valeurs et les fréquences de forme.

Remarque: L'outil Analyst affiche un graphique en secteurs uniquement si le nombre total de lignes des 10 premières valeurs de l'onglet **Valeurs** est supérieur à 25 % du nombre total de lignes dans la source de données.

- Tous les types de données inférés de chaque colonne, avec leurs statuts pour chaque colonne. La fréquence des types de données inférés s'affiche sous la forme d'un nombre et d'un pourcentage.
- Les statistiques sur les valeurs de colonne, telles que la moyenne, la longueur et les valeurs supérieure et inférieure.

Remarque: Vous pouvez sélectionner une valeur ou un modèle et afficher les lignes profilées qui correspondent à la valeur ou au modèle dans le panneau **Détails**

Dans la vue **Propriétés**, vous pouvez afficher les propriétés de profil dans le panneau **Propriétés**. Vous pouvez afficher les propriétés des colonnes et règles dans les panneaux **Colonnes** et **Règles**.

Dans la vue **Aperçu des données**, vous pouvez prévisualiser les données de profil. L'outil Analyst comprend toutes les colonnes dans le profil et affiche les 100 premières lignes des données.

Résumé de profil

Le résumé pour l'exécution de profil inclut le nombre et pourcentage de valeurs uniques et Null, les types de données inférés et la date et l'heure de la dernière exécution. Le résumé inclut également une représentation visuelle du pourcentage de lignes avec des valeurs Null, des valeurs uniques et des valeurs non uniques. Vous pouvez cliquer sur chaque propriété de résumé de profil pour trier les valeurs de la propriété.

Le tableau suivant décrit les propriétés de résumé de profil :

Propriété	Description
Nom	Nom de la colonne dans le profil.
Valeurs uniques	Le nombre de valeurs uniques pour la colonne.
% unique	Pourcentage de valeurs uniques pour la colonne.
Null	Le nombre de valeurs Null pour la colonne.
% Null	Pourcentage de valeurs Null pour la colonne.

Propriété	Description
Type de données	Type de données dérivé des valeurs de la colonne. L'outil Analyst peut dériver les types de données suivants à partir de ceux des valeurs dans les colonnes : <ul style="list-style-type: none"> - String - Varchar - Decimal - Entier - « - » pour les valeurs Null Remarque: L'outil Analyst ne peut pas dériver le type de données des valeurs d'une colonne numérique dont la précision est supérieure à 38. L'outil Analyst ne peut pas dériver le type de données des valeurs d'une colonne de chaîne dont la précision est supérieure à 255. Si vous disposez d'une colonne de date dans laquelle vous créez un profil de colonne avec une valeur d'année antérieure à 1800, il se peut que le type de données inféré s'affiche sous la forme d'une chaîne de longueur fixe. Modifiez la valeur par défaut du paramètre année-minimum dans le fichier InferDateTimeConfig_fr.xml, selon les besoins.
% inférés	Pourcentage de valeurs qui correspondent au type de données inféré par l'outil Analyst.
Type de données consigné	Type de données déclaré pour la colonne de l'objet profilé.
Valeur minimale	Valeur minimale dans la colonne.
Valeur maximale	Valeur maximale dans la colonne.
Dernière exécution de profil	Date et heure de la dernière exécution du profil.
Développement	Si cette option est sélectionnée, permet le développement des données en direct pour la colonne.

Valeurs de colonne

Les valeurs de colonne comprennent les valeurs des colonnes et la fréquence d'affichage de la valeur pour la colonne. Les valeurs de colonne comprennent également un graphique en secteurs qui représente les valeurs et leur pourcentage. L'outil Analyst affiche un graphique en secteurs uniquement si le nombre total de lignes des 10 premières valeurs de l'onglet **Valeurs** est supérieur à 25 % du nombre total de lignes dans la source de données.

Le tableau suivant décrit les propriétés des valeurs de colonne :

Propriété	Description
Valeur	Liste de toutes les valeurs de la colonne dans le profil. Remarque: L'outil Analyst exclut les types de données CLOB, BLOB, brutes et binaires dans les valeurs de colonne d'un profil.
Fréquence	Nombre de fois où une valeur s'affiche pour une colonne, exprimé en nombre, en pourcentage et sous la forme d'un graphique.
Pourcentage	Pourcentage d'affichage d'une valeur pour une colonne.

Propriété	Description
Graphique	Graphique du pourcentage.
Développement	Développe des lignes source spécifiques en fonction d'une valeur de colonne.

Remarque: Pour trier les colonnes **Valeur** et **Fréquence**, sélectionnez les colonnes. Quand vous triez les résultats de la colonne **Fréquence**, l'outil Analyst trie les résultats en fonction du type de données de la colonne.

Formes de colonne

Les modèles de colonne comprennent les modèles pour les valeurs de colonne et la fréquence d'affichage des modèles. Les modèles de colonne comprennent également un graphique en secteurs qui représente les modèles et les fréquences.

L'entrepôt de profilage stocke les 16 000 valeurs de fréquence uniques les plus élevées, y compris les valeurs NULL pour les résultats de profil par défaut. Si les résultats de profil comprennent au moins une valeur NULL, l'outil Analyst peut afficher les valeurs NULL en tant que formes.

Remarque: L'outil Analyst ne peut pas dériver la forme d'une colonne numérique dont la précision est supérieure à 38. L'outil Analyst ne peut pas dériver la forme d'une colonne de chaîne dont la précision est supérieure à 255.

Le tableau suivant décrit les propriétés des formes de colonne :

Propriété	Description
Modèle	Modèle de la colonne dans le profil.
Fréquence	Nombre de fois où un modèle s'affiche pour une colonne, exprimé en nombre.
Pourcentage	Pourcentage d'affichage d'un modèle pour une colonne.
Graphique	Graphique du pourcentage.
Développement	Développe des lignes source spécifiques en fonction d'une forme de colonne.

Le tableau suivant décrit les caractères de forme et ce qu'ils représentent :

Caractère	Description
9	Représente tout caractère numérique. Informatica Analyst affiche jusqu'à trois caractères de manière distincte au format « 9 ». L'outil indique les valeurs de plus de trois caractères à l'aide de parenthèses. Par exemple, le format « 9(8) » représente une valeur numérique composée de huit chiffres.
X	Représente tout caractère alphabétique. Informatica Analyst affiche jusqu'à trois caractères de manière distincte au format « X ». L'outil indique les valeurs de plus de trois caractères à l'aide de parenthèses. Par exemple, le format « X(6) » peut représenter la valeur « Boston ». Remarque: Le X représentant les caractères du modèle n'est pas sensible à la casse et peut représenter des caractères majuscules ou minuscules depuis la source de données.
p	Représente « (», la parenthèse ouvrante.
q	Représente «) », la parenthèse fermante.
b	Représente un espace vide.

Remarque: Les modèles de colonne peuvent également inclure des caractères spéciaux, tels que ~, [], =, -, ?, =, {, *, -, >, < et \$.

Types de données de colonne

Les types de données de colonne comprennent tous les types de données inférés de chaque colonne dans les résultats de profil.

Le tableau suivant décrit les propriétés des types de données de colonne :

Propriété	Description
Type de données	Liste de tous les types de données inférés de la colonne dans le profil.
Fréquence	Nombre de fois où un type de données s'affiche pour une colonne, exprimé en nombre.
Pourcentage	Pourcentage d'affichage d'un type de données pour une colonne.

Propriété	Description
Développement	Développe des lignes source spécifiques en fonction d'un type de données de colonne. Remarque: Vous ne pouvez pas effectuer d'action de développement si vous sélectionnez plusieurs types de données inférés.
Statut	Indique le statut du type de données. Les statuts peuvent être Inféré, Approuvé ou Rejeté. Inféré Indique le type de données de colonne que l'outil Analyst a inféré. Approuvé Indique que le type de données de la colonne est approuvé. Lorsque vous approuvez un type de données, ce dernier est validé pour le référentiel modèle. Rejeté Indique que le type de données de la colonne est rejeté.

Statistiques de colonne

Les statistiques de colonne comprennent les statistiques sur les valeurs de colonne, telles que la moyenne, la longueur et les valeurs supérieure et inférieure. Les statistiques qui s'affichent dépendent du type de colonne.

La table suivante décrit les types de statistiques de colonnes pour chaque type de colonne :

Statistique	Type de colonne	Description
Moyenne	Entier	La moyenne des valeurs pour la colonne.
Écart-type	Entier	L'écart-type ou la variabilité entre les valeurs de colonne, pour toutes les valeurs de la colonne.
Longueur maximale	Entier, chaîne	La longueur de la valeur la plus longue pour la colonne.
longueur minimale	Entier, chaîne	La longueur de la valeur la plus courte pour la colonne.
Somme	Les types de données numériques, comme entier, décimal et double.	La somme de toutes les valeurs dans la colonne.
Inférieur	Entier, chaîne	Les valeurs inférieures pour la colonne.
Supérieur	Entier, chaîne	Les valeurs supérieures pour la colonne.

Développement de profil de colonne

Les options de développement d'un profil de colonne permettent de développer des lignes spécifiques dans la source de données selon une valeur de colonne. Vous pouvez choisir de lire les données actuelles dans une source de données pour le développement ou lire les données de profil stockées temporairement dans l'entrepôt de profils. Lorsque vous développez une ligne spécifique dans des données de profil stockées temporairement, l'outil Analyst crée un filtre de développement pour la valeur de colonne correspondante. Une fois le développement terminé, vous pouvez éditer, rappeler, réinitialiser et enregistrer le filtre de développement.

Vous pouvez sélectionner des colonnes pour le développement même si vous n'avez pas choisi ces colonnes pour le profilage. Vous pouvez choisir de lire les données actuelles dans une source de données pour le développement ou lire les données de profil stockées temporairement dans l'entrepôt de profilage. Après avoir effectué un développement dans une valeur de colonne, vous pouvez exporter les données de développement pour les valeurs sélectionnées ou les formes vers un fichier CSV à l'emplacement choisi. Bien que Informatica Analyst affiche les 200 premières valeurs pour les données de développement, l'outil exporte toutes les valeurs vers le fichier CSV.

Développement des données de ligne

Après avoir exécuté un profil, vous pouvez développer des lignes spécifiques qui correspondent à la valeur ou la forme de la colonne.

1. Exécutez un profil.
Le profil s'affiche dans un onglet.
2. Dans la vue **Résumé**, sélectionnez un nom de colonne pour afficher les résultats de profil pour la colonne.
3. Sélectionnez une valeur de colonne dans l'onglet **Valeurs** ou sélectionnez une forme de colonne dans l'onglet **Formes**.
4. Cliquez sur **Actions > Développement** pour afficher les lignes de données.

Le panneau **Développement** affiche les lignes qui contiennent les valeurs ou les formes. La valeur ou la forme de colonne s'affiche en haut du panneau.

Remarque: Vous pouvez choisir de développer des données actives ou stockées.

Application de filtres aux données de développement

Vous pouvez filtrer les données de développement de manière itérative afin de pouvoir analyser les irrégularités de données dans les sous-ensembles de résultats du profil.

1. Développez les données de la ligne dans les résultats de profil.
2. Sélectionnez une valeur de colonne sur l'onglet **Valeurs**.
3. Cliquez avec le bouton droit de la souris et sélectionnez **Filtre de développement > Modifier** pour ouvrir la boîte de dialogue **Filtre de développement**.
4. Pour ajouter une condition de filtre basée sur les résultats de développement de colonne, cliquez avec le bouton droit sur une valeur de colonne dans les résultats de développement et sélectionnez **Ajouter au filtre**.

La boîte de dialogue **Filtre de développement** s'affiche avec la condition de filtre.

5. Ajoutez les conditions de filtre requises et cliquez sur **Exécuter**.

Vous ne pouvez pas appliquer de filtre de développement à des types de données inférés.

6. Afin de gérer les filtres de développement actuels, vous pouvez les enregistrer, les rappeler ou les réinitialiser.
 - Pour enregistrer un filtre, sélectionnez **Filtre de développement > Enregistrer**.
 - Pour retourner aux derniers résultats de filtres de développement enregistrés, sélectionnez **Filtre de développement > Rappeler**.
 - Pour réinitialiser les résultats des filtres de développement, sélectionnez **Filtre de développement > Réinitialiser**.

Conservation dans l'outil Analyst

La conservation est le processus de validation et de gestion des métadonnées d'une source de données découvertes, de sorte à les préparer afin d'être utilisées et de figurer dans des rapports. Lorsque vous conservez des métadonnées dans l'outil Analyst, vous pouvez approuver, rejeter et rétablir les types de données ou domaines de données inférés dans les résultats de profil.

Vous pouvez approuver un type de données par colonne et masquer les types de données ou domaines de données rejetés dans cette dernière. Après avoir approuvé ou rejeté un type de données ou domaine de données inféré, vous pouvez le rétablir afin de restaurer le statut inféré.

Approbation des types de données

Les résultats de profil comprennent les types de données inférés de chaque colonne dans la source de données. Vous ne pouvez choisir et approuver qu'un seul type de données pour chaque colonne dans l'outil Analyst.

1. Dans le navigateur, sélectionnez le projet ou le dossier qui contient le profil.
2. Cliquez sur le profil pour l'ouvrir.
Le profil s'ouvre dans un onglet.
3. Dans la vue **Profilage de colonnes**, sélectionnez une colonne pour en afficher les fréquences de valeur, les modèles, les types de données et les statistiques dans le panneau de droite.
4. Dans le panneau de droite, cliquez sur **Types de données**.
Les types de données inférés de la colonne s'affichent.
5. Sélectionnez la colonne à approuver, puis cliquez sur **Actions > Approuver**.
Le statut du type de données devient Approuvé.
6. Pour restaurer le statut inféré du type de données, sélectionnez le type de données, puis cliquez sur **Actions > Réinitialiser**.

Rejet des types de données

Lorsque vous ouvrez les résultats de profil, l'outil Analyst affiche les types de données approuvés par défaut. Vous pouvez afficher ou masquer les types de données rejetés.

1. Dans le navigateur, sélectionnez le projet ou le dossier qui contient le profil.
2. Cliquez sur le profil pour l'ouvrir.
Le profil s'ouvre dans un onglet.

3. Sélectionnez une ligne dans la vue **Profilage de colonnes**.
4. Sélectionnez la vue **Types de données** dans le panneau de droite pour rejeter les types de données de colonnes inférés. Sélectionnez le type de données inféré que vous souhaitez rejeter, puis cliquez sur **Actions > Rejeter**.
L'outil Analyst retire le type de données rejeté de la liste des types de données.
5. Pour afficher les types de données rejetés, cliquez sur **Actions > Afficher les éléments rejetés**.
6. Pour masquer les types de données rejetés, cliquez sur **Actions > Masquer les éléments rejetés**.

Fichiers d'exportation de profil de colonne dans Informatica Analyst

Vous pouvez exporter les résultats de profil de colonne dans un fichier CSV ou un fichier Microsoft Excel selon que vous choisissez une partie des résultats de profil ou le résumé des résultats complet.

Vous pouvez exporter les fréquences de valeur, les fréquences de modèle, les types de données ou les données de développement des valeurs et des modèles sélectionnés dans un fichier CSV. Vous pouvez exporter le résumé des résultats de profilage pour toutes les colonnes dans un fichier Microsoft Excel. Servez-vous des **résultats de développement et d'exportation** du privilège de Data Integration Service pour déterminer, par utilisateur ou par groupe, qui exporte les résultats de profil.

Résultats de l'exportation de profil dans un fichier CSV

Vous pouvez exporter des fréquences de valeur, des fréquences de modèle, des types de données ou des données de développement pour afficher les données d'un fichier. L'outil Analyst enregistre les informations dans un fichier CSV.

Lorsque vous exportez les formes de colonne inférées, l'outil Analyst exporte un format différent de la forme de colonne. Par exemple, quand vous exportez la forme de colonne inférée X(5), l'outil Analyst affiche le format de forme de colonne suivant dans le fichier CSV : XXXXX.

Résultats de l'exportation de profil dans Microsoft Excel

Lorsque vous exportez le résumé des résultats de profil complet, l'outil Analyst enregistre les informations dans plusieurs feuilles de calcul dans un fichier Microsoft Excel. L'outil Analyst enregistre le fichier au format « xlsx ».

Le tableau suivant décrit les informations qui apparaissent sur chaque feuille de calcul du fichier d'exportation :

Tab	Description
Profil de colonne	Résumé des informations exportées à partir de la vue de profilage des colonnes après les exécutions du profil. Ces informations sont par exemple : les noms des colonnes, les noms des règles, le nombre de valeurs uniques, le nombre de valeurs null, les types de données inférés, ainsi que la date et l'heure de la dernière exécution de profil.
Valeurs	Valeurs des colonnes et des règles et la fréquence à laquelle les valeurs apparaissent pour chaque colonne.

Tab	Description
Modèles	Formes des valeurs des colonnes et des règles sur lesquelles vous avez exécuté le profil et la fréquence à laquelle les formes apparaissent.
Types de données	Tous les types de données de la colonne, la fréquence de chaque type de données, la valeur de pourcentage et le statut du type de données, par exemple, Inféré, Approuvé ou Rejeté.
Statistiques	Les statistiques relatives à chaque colonne et à chaque règle. Par exemple, la moyenne, la longueur, les valeurs supérieures, les valeurs inférieures et l'écart-type.
Propriétés	Informations de la vue des propriétés, notamment le nom du profil, son type, la stratégie d'échantillonnage et le nombre de lignes.

Exportation des résultats de profil depuis Informatica Analyst

Vous pouvez exporter les résultats d'un profil dans un fichier « .csv » ou « .xlsx » pour afficher les données dans un fichier.

1. Dans le navigateur, sélectionnez le projet ou le dossier qui contient le profil.
2. Cliquez sur le profil pour l'ouvrir.
Le profil s'ouvre dans un onglet.
3. Dans la vue **Profilage de colonne**, sélectionnez la colonne à exporter.
4. Cliquez sur **Actions > Exporter les données**.
La fenêtre **Exporter les données dans un fichier** s'affiche.
5. Entrez le nom de fichier. Utilisez éventuellement le nom de fichier par défaut.
6. Sélectionnez le type de données à exporter.
 - Tout (résumé, valeurs, formes, statistiques, propriétés)
 - Fréquences de valeur pour la colonne sélectionnée.
 - Fréquences de forme pour la colonne sélectionnée.
 - Développez les données pour les valeurs ou formes sélectionnées.
7. Entrez un format de fichier. Le format est **Excel** pour l'option **Tout** et **CSV** pour le reste des options.
8. Sélectionnez la page de codes du fichier.
9. Cliquez sur **OK**.

CHAPITRE 8

Règles dans Informatica Analyst

Ce chapitre comprend les rubriques suivantes :

- [Présentation des règles dans Informatica Analyst, 54](#)
- [Règles prédéfinies, 55](#)
- [Règles d'expression, 56](#)

Présentation des règles dans Informatica Analyst

Une règle est une logique métier qui définit les conditions appliquées aux données source lorsque vous exécutez un profil. Vous pouvez ajouter une règle au profil pour nettoyer, modifier ou valider les données.

Vous souhaitez peut-être utiliser une règle dans des circonstances différentes. Vous pouvez ajouter une règle pour nettoyer une ou plusieurs colonnes de données. Vous pouvez ajouter une règle de recherche qui contient des informations que les données source ne fournissent pas. Vous pouvez ajouter une règle pour valider une règle de nettoyage pour un projet d'intégration ou de qualité des données.

Vous pouvez ajouter une règle avant ou après avoir exécuté un profil. Lors de l'ajout d'une règle à un profil, vous pouvez créer ou appliquer une règle. Vous pouvez créer ou appliquer les types de règles suivants pour un profil :

- **Règles d'expression** Utilisez les fonctions d'expression et les colonnes pour définir la logique de règle. Créez des règles d'expression dans l'outil Analyst. Un analyste peut créer une règle d'expression et en faire une règle réutilisable que d'autres analystes peuvent utiliser dans plusieurs profils.
- **Règles prédéfinies** Comprend les règles réutilisables qu'un développeur crée dans l'outil Developer. Les règles qu'un développeur crée dans l'outil Developer comme maplets peuvent s'afficher dans l'outil Analyst en tant que règles réutilisables.

Après avoir ajouté une règle à un profil, vous pouvez exécuter le profil à nouveau pour la colonne de règle. L'outil Analyst affiche les résultats de profil pour la colonne de règle. Vous pouvez modifier la règle et exécuter le profil à nouveau pour afficher les modifications dans les résultats de profil. La sortie d'une règle peut être une ou plusieurs colonnes virtuelles. Les colonnes virtuelles existent dans les résultats de profils. L'outil Analyst profile les colonnes virtuelles. Par exemple, vous devez utiliser une règle prédéfinie qui divise une colonne qui contient le prénom et le nom en colonnes FIRST_NAME et LAST_NAME virtuelles. L'outil Analyst profile les colonnes FIRST_NAME et LAST_NAME.

Remarque: Si vous supprimez un objet de règle que d'autres types d'objets référencent, l'outil Analyst affiche un message qui indique ces types d'objets. Déterminez l'impact de la suppression de la règle avant de la supprimer.

Règles prédéfinies

Les règles prédéfinies sont créées dans Developer Tool ou fournies avec Developer Tool et Analyst Tool. Appliquez les règles prédéfinies aux profils d'Analyst Tool pour modifier ou valider les données source.

Les règles prédéfinies utilisent des transformations pour définir la logique de règle. Vous pouvez utiliser des règles prédéfinies avec des profils multiples. Dans le référentiel Modèle, une règle prédéfinie est une mapplet avec un groupe d'entrée, un groupe de sortie et des transformations qui définissent la logique de règle.

Processus de règles prédéfinies

Utilisez l'**Assistant Nouvelle règle** pour appliquer une règle prédéfinie à un profil.

Vous pouvez effectuer les étapes suivantes pour appliquer une règle prédéfinie :

1. Ouvrir un profil.
2. Sélectionner une règle prédéfinie.
3. Examiner les paramètres de règles.
4. Sélectionner la colonne d'entrée. Vous pouvez sélectionner plusieurs colonnes si vous souhaitez appliquer la règle à plusieurs colonnes.
5. Configurer les options de profilage.

Application d'une règle prédéfinie

Utilisez l'assistant **Éditer le profil** pour appliquer une règle prédéfinie à un profil. Lorsque vous appliquez une règle prédéfinie, sélectionnez la règle et configurez les colonnes d'entrée et de sortie de la règle. Appliquez une règle prédéfinie pour utiliser une règle promue comme règle réutilisable ou utiliser une règle créée par un développeur.

1. Dans le navigateur, sélectionnez le projet ou le dossier qui contient le profil auquel ajouter la règle.
2. Cliquez sur le profil pour l'ouvrir.
Le profil s'affiche dans un onglet.
3. Cliquez sur **Actions > Éditer > Règles de profilage de colonnes**.
La boîte de dialogue **Éditer le profil** s'affiche.
4. Pour ouvrir la boîte de dialogue **Nouvelle règle**, cliquez sur **+**.
5. Sélectionnez **Appliquer une règle**.
6. Cliquez sur **Suivant**.
7. Dans le panneau **Règles**, sélectionnez la règle à appliquer.
Les colonnes de nom, type de données, description et précision s'affichent pour les colonnes **Entrées** et **Sorties** dans le panneau **Paramètres de règles**.
8. Cliquez sur **Suivant**.
9. Dans le panneau **Entrées et sorties de la règle**, cliquez sur l'icône **Ajouter une entrée**.
Une boîte de dialogue s'affiche.
10. Dans la section **Entrées**, sélectionnez les colonnes d'entrée, puis cliquez sur **OK**. La colonne d'entrée est un nom de colonne dans le profil.
L'outil Analyst génère des colonnes de sortie à partir des colonnes d'entrée sélectionnées.

11. Dans la section **Sorties**, configurez éventuellement le libellé des colonnes de sortie. Par défaut, l'outil Analyst ajoute le préfixe « output » aux libellés des colonnes de sortie.
12. Cliquez sur **Suivant**.
13. Dans le panneau **Colonnes**, sélectionnez les colonnes dans lesquelles vous souhaitez exécuter un profil. Les colonnes comprennent les règles appliquées au profil. Sélectionnez éventuellement **Nom** pour inclure toutes les colonnes.
L'outil Analyst indique le nom, le type de données, la précision et l'échelle pour chaque colonne.
14. Dans le panneau **Options d'échantillonnage**, configurez les options correspondantes.
15. Dans le panneau **Options de développement**, configurez les options correspondantes.
16. Cliquez sur **Enregistrer** pour appliquer la règle ou sur **Enregistrer et exécuter** pour appliquer la règle puis exécuter le profil.

Règles d'expression

Les règles d'expression utilisent des fonctions d'expression et des colonnes pour définir la logique de règle. Créez des règles d'expression et ajoutez-les à un profil dans Analyst Tool.

Utilisez les règles d'expression pour changer ou valider des valeurs de colonnes dans un profil. Vous pouvez créer une ou plusieurs règles d'expression à utiliser dans un profil. Les fonctions d'expression sont des fonctions semblables à SQL utilisées pour transformer les données source. Vous pouvez créer la logique de règle d'expression avec les types suivants de fonctions :

- Caractère
- Conversion
- Nettoyage des données
- Date
- Codage
- Financier
- Numérique
- Scientifique
- Spécial
- Test

Processus de règles d'expression

Utilisez l'assistant **Éditer le profil** pour créer une règle d'expression et l'ajouter à un profil.

L'assistant **Éditer le profil** inclut un éditeur d'expression. Utilisez-le pour ajouter des fonctions d'expression, configurer les colonnes comme entrée aux fonctions, valider l'expression et configurer le type de renvoi, la précision et l'échelle.

La sortie d'une règle d'expression est une colonne virtuelle qui utilise le nom de la règle comme nom de colonne. L'outil Analyst exécute un profil de colonne dans la colonne virtuelle. Par exemple, vous devez utiliser une règle d'expression pour valider un code postal. La règle renvoie 1 si le code postal est valide et 0 s'il n'est pas valide. Informatica Analyst exécute un profil de colonne sur les valeurs de sorties 1 et 0 de la règle.

Vous pouvez effectuer les étapes suivantes pour créer une règle d'expression :

1. Ouvrir un profil.
2. Configurer la logique de règle à l'aide des fonctions d'expression et des colonnes en tant que paramètres.
3. Configurer les options de profilage.

Création d'une règle d'expression

Utilisez l'assistant **Éditer le profil** pour créer une règle d'expression et l'ajouter à un profil. Créez une règle d'expression pour valider des valeurs pour les colonnes d'un profil.

1. Dans le navigateur, sélectionnez le projet ou le dossier qui contient le profil auquel ajouter la règle.
2. Dans le panneau **Contenu**, cliquez sur le profil pour l'ouvrir.
Le profil s'affiche dans un onglet.
3. Cliquez sur **Actions > Éditer > Règles de profilage de colonnes**.
La boîte de dialogue **Éditer le profil** s'affiche.
4. Cliquez sur **Nouveau**.
5. Sélectionnez **Créer une règle**.
6. Cliquez sur **Suivant**.
7. Entrez un nom et une description facultative pour la règle.
8. Transformez éventuellement la règle en règle réutilisable et configurez l'emplacement du projet et du dossier.
Si vous transformez une règle en règle réutilisable, vous ou d'autres utilisateurs peuvent utiliser la règle dans un autre profil comme règle prédéfinie.
9. Dans l'onglet **Fonctions**, sélectionnez une fonction et cliquez sur la flèche droite pour entrer les paramètres de la fonction.
10. Dans l'onglet **Colonnes**, sélectionnez une colonne d'entrée et cliquez sur la flèche droite pour ajouter l'expression dans l'éditeur **Expression**. Vous pouvez également ajouter des opérateurs logiques à l'expression.
11. Cliquez sur **Valider**. Vous pouvez passer à l'étape suivante si l'expression est valide.
12. Cliquez éventuellement sur **Éditer** pour configurer le type de renvoi, la précision et l'échelle.
13. Cliquez sur **Suivant**.
14. Dans le panneau **Profilage de colonnes**, sélectionnez les colonnes dans lesquelles vous souhaitez exécuter un profil. Les colonnes comprennent toutes les règles appliquées au profil. Sélectionnez éventuellement **Nom** pour sélectionner toutes les colonnes.
L'outil Analyst indique le nom, le type de données, la précision et l'échelle pour chaque colonne.
15. Dans le panneau **Options d'échantillonnage**, configurez les options correspondantes.
16. Dans le panneau **Options de développement**, configurez les options correspondantes.
17. Cliquez sur **Enregistrer** pour créer la règle ou sur **Enregistrer et exécuter** pour créer la règle puis exécuter le profil.

CHAPITRE 9

Fiches d'évaluation dans Informatica Analyst

Ce chapitre comprend les rubriques suivantes :

- [Présentation des fiches d'évaluation dans Informatica Analyst, 58](#)
- [Processus de fiche d'évaluation de Informatica Analyst, 59](#)
- [Mesures, 60](#)
- [Groupes de mesures, 63](#)
- [Développement dans les colonnes, 65](#)
- [Graphes de tendance, 66](#)
- [Fichier d'exportation de fiche d'évaluation dans Informatica Analyst, 69](#)
- [Notifications de la fiche d'évaluation, 70](#)
- [Intégration des fiches d'évaluation à des applications externes, 72](#)
- [Lignage des fiches d'évaluation, 74](#)

Présentation des fiches d'évaluation dans Informatica Analyst

Une fiche d'évaluation est la représentation graphique de valeurs valides pour une colonne dans un profil. Vous pouvez créer des fiches d'évaluation pour développer des données en direct ou des données stockées temporairement.

Utilisez des fiches d'évaluation pour mesurer la progression de la qualité des données. Par exemple, vous pouvez créer une fiche d'évaluation pour mesurer la qualité des données avant d'appliquer les règles de qualité des données. Après avoir appliqué des règles de qualité des données, vous pouvez créer une autre fiche d'évaluation pour comparer les effets des règles sur la qualité des données.

Les fiches d'évaluation affichent la fréquence des valeurs pour les colonnes sous forme de scores. Les scores reflètent le pourcentage de valeurs valides dans les colonnes. Une fois le profil exécuté, vous pouvez ajouter des colonnes de profil comme mesures à une fiche d'évaluation. Vous pouvez créer des groupes de mesures afin de regrouper les mesures associées dans une seule entité. Vous pouvez définir des seuils indiquant dans un enregistrement la plage de données incorrectes acceptables pour les colonnes et assigner des poids métriques à chaque mesure. Lorsque vous exécutez une fiche d'évaluation, l'outil Analyst génère les valeurs moyennes pondérées de chaque groupe de mesures. Pour évaluer plus précisément la qualité des données, vous pouvez également affecter un coût fixe ou variable à chaque mesure. Lorsque vous exécutez

la fiche d'évaluation, l'outil Analyst calcule la somme du coût des données incorrectes pour chaque mesure et affiche le coût total.

Pour identifier des enregistrements de données valides et des enregistrements qui ne sont pas valides, vous pouvez développer chaque mesure. Vous pouvez utiliser des graphes de tendance pour suivre la manière dont évoluent les scores de mesures et les coût des données incorrectes dans les mesures sur une période.

Processus de fiche d'évaluation de Informatica Analyst

Vous pouvez exécuter et éditer la fiche d'évaluation dans l'outil Analyst. Vous pouvez créer et afficher une fiche d'évaluation dans l'outil Developer. Vous pouvez exécuter la fiche d'évaluation sur les données actuelles dans l'objet de données ou sur les données stockées dans l'entrepôt de profilage.

Vous pouvez afficher une fiche d'évaluation dans l'espace de travail **Fiches d'évaluation**. Une fois la fiche d'évaluation exécutée, vous pouvez afficher les scores dans le panneau **Fiche d'évaluation**. Vous pouvez sélectionner l'objet de données et y accéder depuis un score de la fiche d'évaluation. L'outil Analyst ouvre l'objet de données dans un autre onglet.

Vous pouvez effectuer les tâches suivantes lors de l'utilisation des fiches d'évaluation :

1. Créer une fiche d'évaluation dans l'outil Developer et ajouter des colonnes depuis un profil.
2. Se connecter éventuellement à l'outil Analyst et ouvrir la fiche d'évaluation dans l'outil Analyst.
3. Après avoir exécuté un profil, ajoutez des colonnes de profil comme mesures à la fiche d'évaluation.
4. Vous pouvez également configurer le coût des données non valides pour chaque mesure.
5. Exécuter la fiche d'évaluation pour générer les scores pour les colonnes.
6. Afficher une fiche d'évaluation pour consulter les scores de chaque colonne dans un enregistrement.
7. Développer un score dans les colonnes.
8. Éditer une fiche d'évaluation.
9. Définissez les seuils de chaque mesure dans une fiche d'évaluation.
10. Créez un groupe pour ajouter ou déplacer les mesures associées de la fiche d'évaluation.
11. Éditer ou supprimer un groupe, selon les besoins.
12. Afficher le graphe de tendance de score pour chaque score afin de surveiller la manière dont le score change au cours du temps.
13. Vous pouvez également afficher le graphe de tendance de coût pour chaque mesure pour surveiller la valeur de la qualité des données.
14. Affichez le lignage des fiches d'évaluation pour chaque métrologie ou groupe de métrologies.

Mesures

Une mesure est la colonne d'une source de données ou de sortie d'une règle faisant partie d'une fiche d'évaluation. Lorsque vous créez une fiche d'évaluation, vous pouvez affecter un poids à chaque mesure. Créez un groupe de mesures pour classer les mesures associées d'une fiche d'évaluation dans un ensemble.

Poids métriques

Lorsque vous créez une fiche d'évaluation, vous pouvez affecter un poids à chaque mesure. La valeur par défaut d'un poids est de 0.

Lorsque vous exécutez une fiche d'évaluation, l'outil Analyst calcule la moyenne pondérée de chaque groupe de mesures en fonction du score de mesure et du poids que vous avez affecté à chaque mesure.

Par exemple, vous affectez un poids de W1 à la mesure M1 et vous assignez un poids de W2 à la mesure M2. L'outil Analyst utilise la formule suivante pour calculer la moyenne pondérée :

$$(M1 \times W1 + M2 \times W2) / (W1 + W2)$$

Valeur de la qualité des données

Une mesure de la qualité des données dans les données source est une information critiques pour la gestion des ressources des données dans l'entreprise. Le coût de données non valides dans les mesures représentées dans une fiche d'évaluation permet aux entreprises d'avoir une idée de la valeur de la surveillance de la qualité des données de la source de données. En tant qu'analyste de données, vous pouvez souhaiter associer une valeur, telle qu'une unité de devise ou une unité personnalisée, à des mesures et des groupes de mesures. Vous pouvez ensuite exécuter la fiche d'évaluation pour afficher le coût total des données non valides dans les données source.

Vous pouvez définir l'unité de coût pour une mesure en fonction des besoins professionnels. Vous pouvez également configurer un coût variable ou fixe pour chaque mesure lorsque vous créez ou modifiez une fiche d'évaluation.

Coût fixe

Le coût fixe est une valeur constante que vous pouvez affecter à une mesure dans une fiche d'évaluation. Vous pouvez choisir une unité de coût prédéfinie ou créer une unité de coût personnalisée qui répond aux besoins professionnels.

Coût variable

Le coût variable est une valeur que vous attribuez à une mesure en fonction des valeurs d'une colonne numérique d'une source de données. Le service d'intégration de données calcule le coût variable de la mesure en fonction de la colonne ou de la colonne virtuelle que vous affectez au coût.

Exemple

En tant que courtier hypothécaire, vous devez fournir à vos clients des livrets de paiement afin que les clients puissent envoyer des paiement hypothécaires. Vous pouvez utiliser une fiche d'évaluation pour mesurer la précision de vos adresses client afin de garantir l'envoi des livrets de paiement. Vous pouvez souhaiter définir la variable coût sur la colonne Montant du paiement mensuel pour la mesure d'exactitude des adresses. Exécutez la fiche d'évaluation pour calculer le total coût que l'entreprise d'hypothèques perd si les clients ne paient pas le montant mensuel à temps.

Ajout de colonnes à une fiche d'évaluation

Après avoir exécuté un profil, vous pouvez ajouter les colonnes d'un profil à une fiche d'évaluation, configurer les valeurs valides pour les colonnes et ajouter le coût des données non valides pour chaque mesure. Si vous ajoutez une colonne à une fiche d'évaluation à partir d'un profil possédant un filtre ou une option d'échantillonnage autre que **Toutes les lignes**, les résultats du profil peuvent ne pas refléter pas les résultats de la fiche d'évaluation.

1. Cliquez sur un profil pour l'ouvrir.

Le profil s'affiche dans l'espace de travail **Découverte**.

2. Cliquez sur **Actions > Exécuter le profil** pour exécuter le profil.
3. Cliquez sur **Actions > Ajouter à > Fiche d'évaluation**.

L'assistant **Ajouter à la fiche d'évaluation** s'affiche.

Remarque: Utiliser les règles et les directives suivantes avant d'ajouter des colonnes à une fiche d'évaluation :

- Vous ne pouvez pas ajouter une colonne à une fiche d'évaluation si le nom de la colonne et le nom de la fiche d'évaluation sont identiques.
- Vous ne pouvez pas ajouter deux fois la même colonne à une fiche d'évaluation même si vous changez le nom de la colonne.

4. Sélectionnez **Fiche d'évaluation existante** pour ajouter les colonnes à une fiche d'évaluation prédéfinie. L'option **Nouvelle fiche d'évaluation** est sélectionnée par défaut.

5. Cliquez sur **Suivant**.

6. Sélectionnez la fiche d'évaluation à laquelle vous voulez ajouter les colonnes et cliquez sur **Suivant**.

7. Sélectionnez les colonnes et les règles que vous voulez ajouter comme scores à la fiche d'évaluation. Cochez éventuellement la case dans l'en-tête de colonne gauche pour sélectionner toutes les colonnes. Sélectionnez éventuellement **Nom de colonne** pour trier les noms de colonne.

8. Cliquez sur **Suivant**.

9. Sélectionnez chaque score dans le volet **Scores** et configurez les valeurs valides dans la liste de toutes les valeurs du volet **Score avec : des valeurs**.

Vous pouvez sélectionner plusieurs valeurs dans le volet **Valeurs disponibles** et cliquer sur la flèche de droite pour les déplacer vers le volet **Valeurs valides**. Le nombre total de valeurs valides pour une mesure s'affiche en haut du volet **Valeurs disponibles**.

10. Sélectionnez chaque mesure et configurez le coût des données non valides. Pour affecter une valeur constante au coût de la mesure, sélectionnez **Coût fixe**.

Vous pouvez également cliquer sur **Changer l'unité de coût** pour modifier l'unité de coût.

11. Pour joindre une colonne numérique, telle qu'une variable de coût, à la mesure, sélectionnez **Coût variable** et cliquez sur **Sélectionner une colonne** pour sélectionner une colonne numérique.

12. Dans le volet **Mesures**, sélectionnez chaque mesure.

13. Dans le volet **Seuils des mesures**, configurez les seuils des mesures.

Vous pouvez définir les seuils pour les scores **Satisfaisant**, **Acceptable** et **Inacceptable**.

14. Cliquez sur **Suivant**.

15. Définissez le poids de chaque score. Vous pouvez double-cliquer sur le poids de score par défaut, qui est de 0, pour le modifier.

16. Dans le volet **Seuils des groupes de scores**, configurez les seuils des groupes de scores.

17. Cliquez sur **Enregistrer** pour enregistrer la fiche d'évaluation ou sur **Enregistrer et exécuter** pour enregistrer et exécuter la fiche d'évaluation.

Exécution d'une fiche d'évaluation

Exécutez une fiche d'évaluation pour générer des scores pour les colonnes.

1. Dans le panneau **Ressources**, choisissez la fiche d'évaluation que vous voulez exécuter.
2. Cliquez sur la fiche d'évaluation pour l'ouvrir.
La fiche d'évaluation s'affiche dans l'espace de travail **Fiches d'évaluation**.
3. Cliquez sur **Actions > Exécuter la fiche d'évaluation**.
4. Sélectionnez un score dans le panneau **Mesures**, puis sélectionnez les colonnes dans le panneau **Colonnes** pour effectuer le développement.
5. Dans l'option **Développement**, choisissez de développer les données en direct ou stockées.
Pour des performances optimales, développez les données actives.
6. Cliquez sur **Exécuter**.

Affichage d'une fiche d'évaluation

Exécutez une fiche d'évaluation pour voir les scores de chaque mesure. Une fiche d'évaluation affiche le score en pourcentage et sous la forme d'un graphique à barres. Afficher les données valides ou non valides. Vous pouvez également afficher les informations de la fiche d'évaluation, comme le poids métrique, le score du groupe de mesures, la tendance du score ainsi que le nom de l'objet de données.

1. Exécutez une fiche d'évaluation pour afficher les scores.
2. Sélectionnez une mesure qui contient le score à afficher.
3. Cliquez sur **Actions > Développement** pour afficher les lignes de données valides ou les lignes de données non valides de la colonne.
L'outil Analyst affiche les lignes de données non valides par défaut dans la section **Développement**.

Édition d'une fiche d'évaluation

Modifiez les valeurs valides des mesures dans une fiche d'évaluation. Vous devez exécuter une fiche d'évaluation avant de pouvoir l'éditer.

1. Dans le panneau **Ressources**, choisissez une fiche d'évaluation.
2. Cliquez sur la fiche d'évaluation pour l'ouvrir.
La fiche d'évaluation s'affiche dans l'espace de travail **Fiches d'évaluation**.
3. Pour modifier les propriétés générales de la fiche d'évaluation, cliquez sur **Actions > Modifier > Général**.
La boîte de dialogue **Éditer la fiche d'évaluation** s'ouvre.
4. Modifiez le nom et description de la fiche d'évaluation si nécessaire.
5. Cliquez sur l'onglet **Mesures**.
6. Sélectionnez chaque mesure dans le volet **Mesures** et configurez les valeurs valides dans la liste de toutes les valeurs du volet **Score avec : des valeurs**.
7. Dans le volet **Seuils des mesures**, apportez des modifications aux seuils de scores, si nécessaire.

8. Vérifiez le coût des données non valides pour chaque mesure et effectuez des modifications, si nécessaire.
9. Cliquez sur l'onglet **Groupes de mesures**.
10. Créez, modifiez ou supprimez des groupes de scores.
Vous pouvez également modifier le poids métrique et les seuils de mesures sur l'onglet **Groupes de mesures**.
11. Cliquez sur l'onglet **Notifications**.
12. Modifiez les paramètres de notification de la fiche d'évaluation si nécessaire.
Vous pouvez configurer des paramètres généraux et personnalisés pour les mesures et les groupes de mesures.
13. Cliquez sur **Enregistrer** pour enregistrer les modifications apportées à la fiche d'évaluation ou cliquez sur **Enregistrer et exécuter** pour enregistrer les modifications et exécuter la fiche d'évaluation.

Définition des seuils

Vous pouvez définir des seuils pour chaque score dans une fiche d'évaluation. Un seuil détermine la plage en pourcentage de données incorrectes acceptables pour les colonnes dans un enregistrement. Vous pouvez définir des seuils pour des plages de données satisfaisantes, acceptables ou inacceptables. Vous pouvez définir des seuils pour chaque colonne quand vous ajoutez des colonnes à une fiche d'évaluation, ou quand vous éditez une fiche d'évaluation.

Effectuez l'une des tâches préalables suivantes avant de définir des seuils pour les colonnes dans une fiche d'évaluation :

- Ouvrez un profil et ajoutez des colonnes du profil à la fiche d'évaluation dans la boîte de dialogue **Ajouter à la fiche d'évaluation**.
 - Vous pouvez cliquer sur une fiche d'évaluation dans l'espace de travail **Bibliothèque** et sélectionner **Actions > Modifier** pour modifier la fiche d'évaluation dans la boîte de dialogue **Modifier la fiche d'évaluation**.
1. Dans la boîte de dialogue **Ajouter à la fiche d'évaluation** ou **Modifier la fiche d'évaluation**, sélectionnez chaque score dans le volet **Scores**.
 2. Dans le volet **Seuils des scores**, entrez les seuils qui représentent la limite supérieure de la plage inacceptable et la limite inférieure de la plage satisfaisante.
 3. Cliquez sur **Suivant** ou **Enregistrer**.

Groupes de mesures

Créez un groupe de mesures pour classer les mesures associés d'une fiche d'évaluation dans un ensemble. Par défaut, l'outil Analyst classe tous les scores dans un groupe de mesures par défaut.

Après avoir créé un groupe de mesures, vous pouvez déplacer les scores du groupe de mesures par défaut vers un autre groupe de mesures. Vous pouvez modifier le nom et la description d'un groupe de mesures, y compris pour le groupe de mesures par défaut. Vous pouvez supprimer les groupes de mesures que vous n'utilisez plus. Vous ne pouvez pas supprimer le groupe de mesures par défaut.

Création d'un groupe de mesures

Créez un groupe de mesures pour ajouter les scores associés de la fiche d'évaluation au groupe.

1. Dans le navigateur, sélectionnez le projet ou le dossier qui contient la fiche d'évaluation.
2. Cliquez sur la fiche d'évaluation pour l'ouvrir.
La fiche d'évaluation s'affiche dans un onglet.
3. Cliquez sur **Actions > Éditer**.
La fenêtre **Éditer la fiche d'évaluation** s'affiche.
4. Cliquez sur l'onglet **Groupes de mesures**.
Le groupe par défaut s'affiche dans le panneau **Groupes de mesures** et les scores du groupe par défaut s'affichent dans le panneau **Mesures**.
5. Cliquez sur l'icône **Nouveau groupe** pour créer un groupe de mesures.
La boîte de dialogue **Groupes de mesures** s'ouvre.
6. Entrez un nom et une description facultative.
7. Cliquez sur **OK**.
8. Cliquez sur **Enregistrer** pour enregistrer les modifications dans la fiche d'évaluation.

Déplacement des scores vers un groupe de mesures

Une fois le groupe de mesures créé, vous pouvez déplacer les scores associés vers le groupe de mesures.

1. Dans le navigateur, sélectionnez le projet ou le dossier qui contient la fiche d'évaluation.
2. Cliquez sur la fiche d'évaluation pour l'ouvrir.
La fiche d'évaluation s'affiche dans un onglet.
3. Cliquez sur **Actions > Éditer**.
La fenêtre **Éditer la fiche d'évaluation** s'affiche.
4. Cliquez sur l'onglet **Groupes de mesures**.
Le groupe par défaut s'affiche dans le panneau **Groupes de mesures** et les scores du groupe par défaut s'affichent dans le panneau **Mesures**.
5. Sélectionnez une mesure depuis le panneau **Mesures** et cliquez sur l'icône **Déplacer les mesures**.
La boîte de dialogue **Déplacer les mesures** s'ouvre.
Remarque: Pour sélectionner plusieurs scores, maintenez enfoncée la touche Maj.
6. Sélectionnez le groupe de mesures vers lequel déplacer les scores.
7. Cliquez sur **OK**.

Modification d'un groupe de mesures

Modifier le nom et la description d'un groupe de mesures. Vous pouvez changer le nom du groupe de mesures par défaut.

1. Dans le navigateur, sélectionnez le projet ou le dossier qui contient la fiche d'évaluation.
2. Cliquez sur la fiche d'évaluation pour l'ouvrir.
La fiche d'évaluation s'ouvre dans un onglet.

3. Cliquez sur **Actions > Éditer**.
La fenêtre **Éditer la fiche d'évaluation** s'affiche.
4. Cliquez sur l'onglet **Groupes de mesures**.
Le groupe de mesures par défaut s'affiche dans le panneau **Groupes de mesures** et les mesures du groupe de mesures par défaut s'affichent dans le panneau **Mesures**.
5. Sur le panneau **Groupes de mesures**, cliquez sur l'icône **Éditer le groupe**.
La boîte de dialogue **Éditer** s'ouvre.
6. Entrez un nom et une description facultative.
7. Cliquez sur **OK**.

Suppression d'un groupe de mesures

Vous pouvez supprimer un groupe de mesures qui n'est plus valide. Quand vous supprimez un groupe de mesures, vous pouvez choisir de déplacer les scores du groupe de mesures vers le groupe de mesures par défaut. Vous ne pouvez pas supprimer le groupe de mesures par défaut.

1. Dans le navigateur, sélectionnez le projet ou le dossier qui contient la fiche d'évaluation.
2. Cliquez sur la fiche d'évaluation pour l'ouvrir.
La fiche d'évaluation s'ouvre dans un onglet.
3. Cliquez sur **Actions > Éditer**.
La fenêtre **Éditer la fiche d'évaluation** s'affiche.
4. Cliquez sur l'onglet **Groupes de mesures**.
Le groupe de mesures par défaut s'affiche dans le panneau **Groupes de mesures** et les mesures du groupe de mesures par défaut s'affichent dans le panneau **Mesures**.
5. Sélectionnez un groupe de mesures dans le panneau **Groupes de mesures** et cliquez sur l'icône **Supprimer le groupe**.
La boîte de dialogue **Supprimer les groupes** s'ouvre.
6. Choisissez l'option qui consiste à supprimer les mesures du groupe de mesures ou l'option pour déplacer les mesures vers le groupe de mesures par défaut avant de supprimer le groupe de mesures.
7. Cliquez sur **OK**.

Développement dans les colonnes

Développez les colonnes d'un score pour sélectionner des colonnes qui apparaissent quand vous affichez les lignes de données valides ou non valides. Les colonnes choisies pour effectuer le développement s'affichent dans le panneau **Développement**.

1. Exécutez une fiche d'évaluation pour afficher les scores.
2. Sélectionnez une colonne qui contient le score à afficher.
3. Cliquez sur **Actions > Développer** pour afficher les lignes des données valides ou non valides de la colonne.

4. Cliquez sur **Actions > Colonnes de développement**.

Les colonnes s'affichent dans le panneau **Développement** pour le score sélectionné. L'outil Analyst affiche les lignes des données valides pour les colonnes par défaut. Cliquez éventuellement sur **Non valide** pour afficher les lignes de données non valides.

Graphes de tendance

Utilisez les graphes de tendances pour surveiller la manière dont évoluent les scores de mesures et le coût des données non valides sur une période.

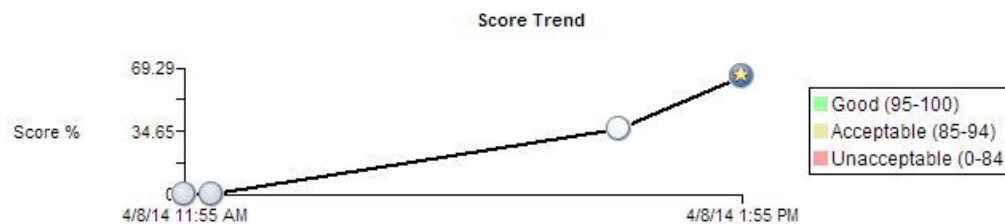
Les graphes de tendance contiennent les graphes de score et de coût qui tracent les valeurs de score ou de coût selon l'axe vertical en fonction de toutes les exécutions de fiche d'évaluation selon l'axe horizontal. Par défaut, le graphe de tendance affiche les données à partir des 10 dernières exécutions de la fiche d'évaluation. Vous pouvez afficher le nombre de lignes totales et de lignes non valides pour la mesure dans le graphe de tendance. Le graphe de tendance affiche également si les tendances de score et de coût demeurent constantes ou augmentent ou diminuent en fonction de la dernière exécution de la fiche d'évaluation.

L'outil Analyst utilise les données historiques d'exécution de la fiche d'évaluation pour chaque date et les dernières valeurs valides de score pour calculer le score. L'outil Analyst utilise les derniers paramètres de seuil dans le graphe pour décrire la couleur des points de score. Vous pouvez voir les seuils Bon, Acceptable, et Inacceptable pour le score. Les seuils changent à chaque fois que vous exécutez la fiche d'évaluation après avoir modifié les valeurs des scores dans la fiche d'évaluation. Lorsque vous exportez une fiche d'évaluation, l'outil Analyst inclut les informations relatives au graphe de tendance, notamment les informations concernant les scores et les coûts dans le fichier exporté.

Graphe de tendance de score

Un graphe de tendance de score est une représentation graphique de la manière dont évoluent les scores de mesure sur plusieurs exécutions de profils. Le graphe de tendance de score trace les valeurs de score de mesure selon l'axe vertical en fonction de toutes les exécutions de fiche d'évaluation selon l'axe horizontal.

L'image suivante affiche un exemple de graphe de tendance de score :



Exemple

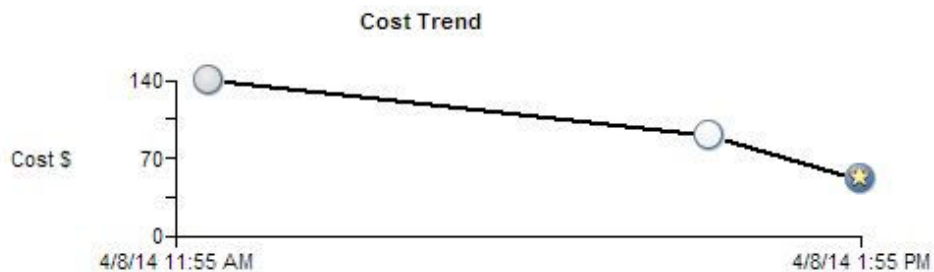
En tant qu'analyste de données, vous pouvez surveiller la qualité des données afin de déterminer si les mappages et d'autres processus modifient le résultat en améliorant le score de la qualité des données. Après avoir mesuré la modification de qualité des données, vous pouvez signaler la modification de la qualité des données afin que l'entreprise les analyse et les utilise. Par exemple, à la fin de plusieurs exécutions de la fiche d'évaluation, le pourcentage de valeurs valides dans une colonne de numéro de sécurité sociale peut être passé de 84 à 90. Vous pouvez signaler cette modification de qualité des données sous la forme d'un graphe visuel permettant une analyse rapide.

Graphes de tendance de coût

Un graphe de tendance de coût est une représentation graphique de la manière dont varie le coût des données non valides dans les mesures sur plusieurs exécutions de profils. Le graphe de tendance de coût peut mesurer l'incidence de la qualité des données dans une entreprise. Le graphe de tendance de coût trace les valeurs de coût selon l'axe vertical en fonction de toutes les exécutions de fiche d'évaluation selon l'axe horizontal. Vous pouvez également afficher le coût total des données non valides et des valeurs valides pour la mesure dans une grille sous le graphe de tendance de coût.

Un graphe de tendance de coût vous permet de suivre l'incidence des données non valides dans les enregistrements de valeur élevée. Parfois, lorsque vous utilisez un coût fixe pour calculer des données non valides, vous pouvez manquer l'incidence des données non valides dans les enregistrements de valeur élevée. Ce problème se produit, car les graphes de tendance peuvent afficher une amélioration du score et une réduction du coût global sur plusieurs exécutions de la fiche d'évaluation. Cependant, des problèmes de qualité des données moins importants représentés dans la fiche d'évaluation peuvent exister des enregistrements de valeur élevée.

L'image suivante affiche un exemple de graphe de tendance de coût :



Exemple

Dans une institution financière, vous avez plusieurs clients disposant d'un solde élevé avec des dépôts et des investissements importants, par exemple 10 millions de dollars, à la banque. Vous avez également un grand nombre de clients ayant un solde peu élevé. Le graphe de tendance de score peut afficher une amélioration de scores sur une période. Toutefois, une adresse incorrecte ou erreur sur le sexe concernant quelques comptes de quelques clients disposant d'un solde élevé peuvent avoir une incidence sur les relations avec les clients les plus précieux de l'organisation. Vous pouvez définir la colonne Solde de compte comme colonne de coût variable pour le calcul des données non valides. Si le coût des données non valides lié à la colonne est élevé, vous pouvez prendre en compte la valeur totale du risque et prendre une action corrective immédiate.

Affichage des graphes de tendances

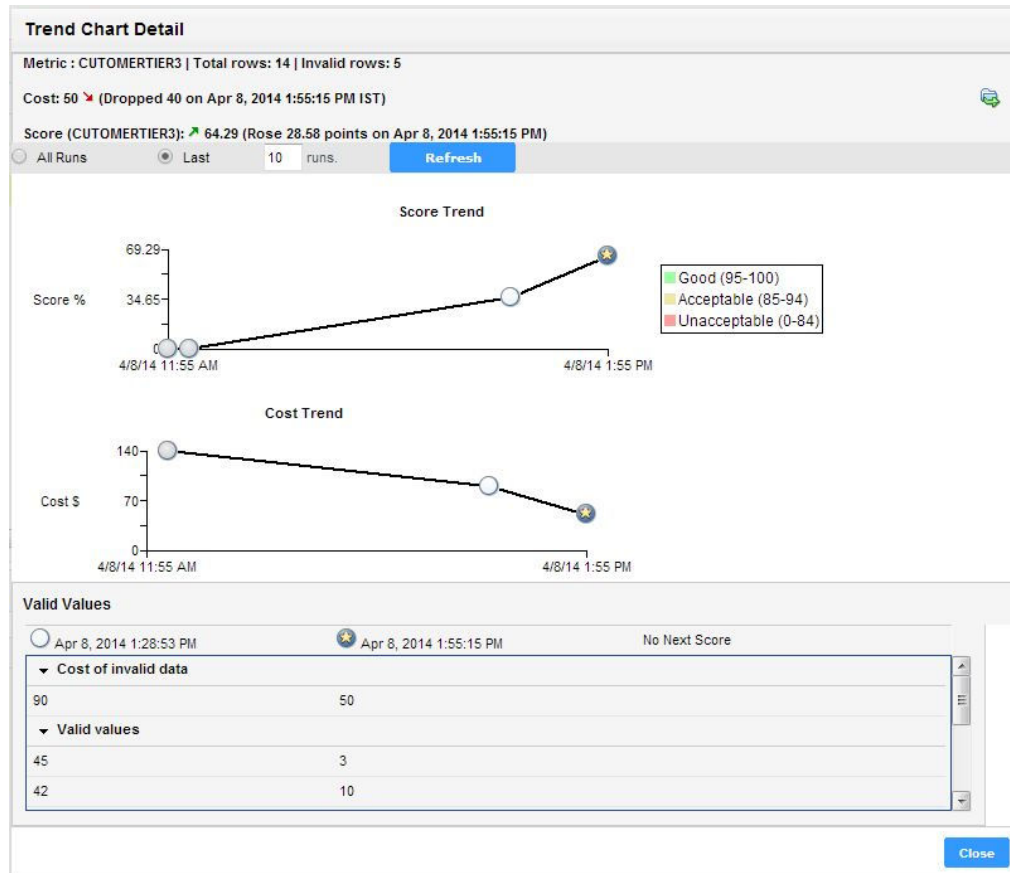
Vous pouvez afficher des graphes de tendance pour chaque mesure afin de surveiller la manière selon laquelle évolue le score ou le coût des données non valides dans le temps.

1. Dans le navigateur de bibliothèque, sélectionnez le projet ou le dossier qui contient la fiche d'évaluation.
2. Cliquez sur la fiche d'évaluation pour l'ouvrir.
La fiche d'évaluation s'affiche dans l'espace de travail **Fiches d'évaluation**.
3. Dans la vue **Fiche d'évaluation**, sélectionnez une mesure.

4. Cliquez sur **Actions > Afficher le graphe de tendance**.

La boîte de dialogue **Détail de graphe de tendance** s'affiche.

L'image suivante affiche la boîte de dialogue **Détail de graphe de tendance** :



Vous pouvez afficher les valeurs de score et de coût qui ont changé au cours du temps. En haut de la boîte de dialogue, vous pouvez afficher le nombre total de lignes et le nombre de lignes non valides. L'outil Analyst utilise les données historiques d'exécution de fiche d'évaluation pour chaque date et les dernières valeurs valides de score pour calculer le score. Sous les graphes de tendance de score et de coût, vous pouvez afficher les valeurs valides pour les mesures et le coût des données non valides.

Exportation des graphes de tendance

Vous pouvez exporter les graphes de score et de tendance vers un fichier « .xlsx » pour afficher les données dans un fichier.

1. Ouvrez une fiche d'évaluation.
2. Sélectionnez une mesure et cliquez sur **Actions > Afficher le graphe de tendance**.
La boîte de dialogue **Détail de graphe de tendance** s'affiche.
3. Cliquez sur l'icône **Exporter les données**.
La boîte de dialogue **Exporter les données vers un fichier** s'ouvre.
4. Entrez un nom de fichier. Utilisez éventuellement le nom de fichier par défaut.
Le format de fichier par défaut est Microsoft Excel.
5. Sélectionnez la page de code du fichier.

6. Cliquez sur **OK**.

Fichier d'exportation de fiche d'évaluation dans Informatica Analyst

Vous pouvez exporter les résultats d'une fiche d'évaluation vers un fichier Microsoft Excel. L'outil Analyst enregistre le fichier au format « .xlsx ».

Le fichier exporté contient le résumé de la fiche d'évaluation, des graphes de tendance, des lignes non valides et les propriétés de la fiche d'évaluation.

Résultats de l'exportation de fiche d'évaluation dans Microsoft Excel

Lorsque vous exportez les résultats d'une fiche d'évaluation, l'outil Analyst enregistre les informations dans plusieurs feuilles de calcul dans un fichier Microsoft Excel. L'outil Analyst enregistre le fichier au format « .xlsx ».

Le tableau suivant décrit les informations qui apparaissent sur chaque feuille de calcul du fichier d'exportation :

Tab	Description
Résumé de la fiche d'évaluation	Résumé des résultats de la fiche d'évaluation exportés. Les informations comprennent le nom de la fiche d'évaluation, le nombre total de lignes pour chaque colonne, le nombre de lignes qui ne sont pas valides, le score et le poids métrique.
Graphe de tendance	Graphes de tendance pour les scores.
Lignes non valides	Détails des lignes qui ne sont pas valides pour chaque colonne.
Propriétés	Propriétés de la fiche d'évaluation, telles que le nom, le type, la description et l'emplacement.

Exportation des résultats de fiche d'évaluation à partir d'Informatica Analyst

Vous pouvez exporter les résultats d'une fiche d'évaluation vers un fichier « .xlsx » pour afficher les données dans un fichier.

1. Ouvrez une fiche d'évaluation.
2. Cliquez sur **Actions > Exporter les données**.
La boîte de dialogue **Exporter les données vers un fichier** s'ouvre.
3. Entrez un nom de fichier. Utilisez éventuellement le nom de fichier par défaut.
Le format de fichier par défaut est Microsoft Excel.
4. Sélectionnez la page de code du fichier.
5. Cliquez sur **OK**.

Notifications de la fiche d'évaluation

Configurez les paramètres de notification de la fiche d'évaluation pour que l'outil Analyst envoie des e-mails lorsque les scores de mesure spécifiques, les scores de groupe de mesures ou les coûts de mesures dépassent les seuils. Les scores de mesures ou les scores des groupes de mesures peuvent dépasser les seuils ou rester dans des plages de scores spécifiques, par exemple Inacceptable, Acceptable et Satisfaisant. Les valeurs de coût des mesures peuvent dépasser les seuils de coût maximal et minimal que vous avez définis.

Vous pouvez configurer des notifications par e-mail pour des scores de mesures individuelles, des groupes de mesures et des coûts de mesures. Si vous utilisez les paramètres généraux pour les scores, l'outil Analyst envoie des notifications par e-mail lorsque des scores de mesures spécifiques dépassent le seuil à partir duquel le score passe de Satisfaisant à Acceptable et d'Acceptable à Insatisfaisant. Vous obtenez également des notifications par e-mail pour chaque exécution de la fiche d'évaluation si le score demeure dans la plage de scores Inacceptable après plusieurs exécutions consécutives de la fiche d'évaluation. Si vous utilisez les paramètres généraux pour les coûts de mesures, l'outil Analyst envoie des notifications par e-mail lorsque le coût des données non valides dans les mesures sélectionnées dépasse les seuils maximal et minimal.

Vous pouvez personnaliser les paramètres de notification pour que les utilisateurs de la fiche d'évaluation reçoivent des notifications par e-mail lorsque les scores passent de la plage de scores Inacceptable à Acceptable et Acceptable à Satisfaisant. Vous pouvez choisir d'envoyer des notifications par e-mail si un score de mesure ou un coût de mesure demeure dans les plages spécifiques pour chaque exécution de la fiche d'évaluation. Vous pouvez afficher le coût actuel des données non valides pour chaque mesure dans les paramètres de notification en fonction desquels vous pouvez définir les seuils de coût.

Modèle de notification de message électronique

Vous pouvez configurer le message et la structure du message électronique que l'outil Analyst envoie aux destinataires comme partie des notifications de la fiche d'évaluation. Le modèle dispose de sections facultatives pour le texte d'introduction et le texte de clôture ainsi que d'une section relative au corps du message en lecture seule.

Le tableau suivant décrit les balises dans le modèle d'e-mail :

Balise	Description
ScorecardName	Nom de la fiche d'évaluation.
ObjectURL	Lien hypertexte vers la fiche d'évaluation. Vous devez fournir le nom d'utilisateur et le mot de passe.
MetricGroupName	Nom du groupe de mesures auquel la mesure appartient.
CurrentWeightedAverage	Valeur moyenne pondérée du groupe de mesures dans l'exécution actuelle de la fiche d'évaluation.
CurrentRange	Score de plage, par exemple Inacceptable, Acceptable et Satisfaisant pour le groupe de mesures de l'exécution actuelle de la fiche d'évaluation.
PreviousWeightedAverage	Valeur moyenne pondérée du groupe de mesures de l'exécution précédente de la fiche d'évaluation.
PreviousRange	Score de plage, par exemple Inacceptable, Acceptable et Satisfaisant pour le groupe de mesures de l'exécution précédente de la fiche d'évaluation.

Balise	Description
MetricName	Nom de la mesure.
MetricGroupName	Nom du groupe de mesures.
CurrentScore	Score basé sur la dernière exécution de la fiche d'évaluation.
CurrentRange	Plage de score dans laquelle le score actuel demeure basé sur la dernière exécution de la fiche d'évaluation.
PreviousScore	Score basé sur l'exécution précédente de la fiche d'évaluation.
PreviousRange	Plage de score basée sur l'exécution précédente de la fiche d'évaluation.
CurrentCost	Coût des données non valides dans la mesure basée sur la dernière exécution de la fiche d'évaluation.
PreviousCost	Coût des données non valides dans la mesure basée sur l'exécution précédente de la fiche d'évaluation.
ColumnName	Nom de la colonne source à laquelle la mesure est affectée.
ColumnType	Type de la colonne source.
RuleName	Nom de la règle.
RuleType	Type de la règle.
DataObjectName	Nom de l'objet de données source.

Configuration des notifications de la fiche d'évaluation

Vous pouvez configurer les notifications de la fiche d'évaluation au niveau des mesures et au niveau du groupe de mesures. Les paramètres généraux de notification s'appliquent aux mesures et aux groupes de mesures qui n'ont pas de paramètres de notification individuels.

1. Exécuter une fiche d'évaluation dans l'outil Analyst.
2. Cliquez sur **Actions > Éditer**.
3. Cliquez sur l'onglet **Notifications**.
4. Sélectionnez **Activer les notifications** pour démarrer la configuration des notifications de la fiche d'évaluation.
5. Sélectionnez une mesure ou un groupe de mesures.
6. Cochez la case **Notifications** pour activer les paramètres généraux d'une mesure ou d'un groupe de mesures.
7. Sélectionnez **Utiliser les paramètres personnalisés** pour modifier les paramètres de la mesure ou du groupe de mesures.

Vous pouvez choisir d'envoyer une notification par e-mail lorsque le score se trouve dans une plage **Inacceptable**, **Acceptable** et **Satisfaisant** et passe d'un seuil à un autre. Vous pouvez également envoyer un e-mail de notification lorsque le coût de la mesure dépasse les seuils maximal ou minimal.

8. Pour éditer les paramètres généraux de notification de la fiche d'évaluation, cliquez sur l'icône **Éditer les paramètres généraux**.
La boîte de dialogue **Éditer les paramètres généraux** s'ouvre et vous pouvez y modifier les paramètres, y compris le modèle d'e-mail.

Configuration des paramètres généraux des notifications de la fiche d'évaluation

Si vous sélectionnez les paramètres généraux des notifications de la fiche d'évaluation, l'outil Analyst envoie des e-mails aux utilisateurs cibles lorsque le score se trouve dans la plage **Inacceptable**. Vous pouvez également configurer les paramètres de notification pour envoyer des e-mails lorsque les scores des mesures ou les coûts des mesures déplacent les seuils. Vous pouvez configurer le modèle d'e-mail, notamment les adresses de courriel et le texte du message d'une fiche d'évaluation.

1. Exécuter une fiche d'évaluation dans l'outil Analyst.
2. Cliquez sur **Actions > Éditer > Notifications** pour ouvrir la boîte de dialogue **Modifier la fiche d'évaluation**.
3. Sélectionnez **Activer les notifications** pour démarrer la configuration des notifications de la fiche d'évaluation.
4. Cliquez sur l'icône **Éditer les paramètres généraux**.
La boîte de dialogue **Éditer les paramètres généraux** s'ouvre et vous pouvez y modifier les paramètres, y compris le modèle d'e-mail.
5. Choisissez le moment d'envoi des notifications par e-mail pour les scores des mesures à l'aide des cases à cocher **Score dans** et **Déplacements des scores**.
6. Choisissez à quel moment envoyer des notifications par e-mail pour les coûts des mesures à l'aide des cases à cocher **Le coût va**.
7. Dans le champ **Destinataire**, entrez l'ID d'e-mail du destinataire. Utilisez un point-virgule pour séparer plusieurs identifiants d'e-mail.
Par défaut, l'outil Analyst utilise la propriété **Adresse de messagerie de l'expéditeur** du service d'intégration de données comme ID d'e-mail de l'expéditeur.
8. Entrez le texte de l'objet de l'e-mail.
9. Dans le champ **Corps**, ajoutez le texte d'introduction et de clôture du message électronique.
10. Pour appliquer les paramètres généraux, sélectionnez **Appliquer les paramètres à toutes les mesures et tous les groupes de mesures**.
11. Cliquez sur **OK**.

Intégration des fiches d'évaluation à des applications externes

Vous pouvez créer une fiche d'évaluation dans l'outil Analyst et en consulter les résultats dans des applications externes ou dans des portails Web. Précisez l'URL des résultats de fiche d'évaluation dans un

format incluant le nom d'hôte, le numéro de port, l'identifiant de projet et l'identifiant de fiche d'évaluation afin de consulter les résultats dans des applications externes.

Ouvrez une fiche d'évaluation après l'avoir exécutée et copié son URL à partir de votre navigateur. L'URL doit être au format suivant :

```
http://{HOST_NAME}:{PORT}/AnalystTool/com.informatica.at.AnalystTool/index.jsp?
mode=scorecard&project={MRS_PROJECT_ID}&id={SCORECARD_ID}&parentpath={MRS_PARENT_PATH}&view={
VIEW_MODE}&pcsfcred={CREDENTIAL}
```

Le tableau suivant décrit les attributs de l'URL de fiche d'évaluation :

Attribut	Description
HOST_NAME	Nom de l'hôte du service Analyst.
PORT	Numéro de port pour le service Analyst.
MRS_PROJECT_ID	Identifiant de projet dans le référentiel modèle.
SCORECARD_ID	Identifiant de la fiche d'évaluation.
MRS_PARENT_PATH	Emplacement de la fiche d'évaluation dans l'outil Analyst. Par exemple, /project1/folder1/sub_folder1.
VIEW_MODE	Détermine si un affichage de la fiche d'évaluation en lecture seule ou modifiable est intégré avec l'application externe.
CREDENTIAL	Dernière partie de l'URL générée par la fonctionnalité d'authentification unique qui représente le type d'objet, fiche d'évaluation par exemple.

L'attribut VIEW_MODE apparaissant dans l'URL de la fiche d'évaluation détermine si vous pouvez intégrer un affichage en lecture seule ou modifiable de la fiche d'évaluation avec l'application externe :

`view=objectonly`

Offre un affichage en lecture seule des résultats de la fiche d'évaluation.

`view=objectrunonly`

Affiche les résultats de fiche d'évaluation dans lesquels vous pouvez exécuter les fiches d'évaluation et le développement des résultats.

`view=full`

Ouvre les résultats de la fiche d'évaluation dans l'outil Analyst avec un accès complet.

Consultation d'une fiche d'évaluation dans des applications externes

Vous pouvez consulter une fiche d'évaluation en utilisant l'URL de la fiche d'évaluation dans des applications externes ou dans des portails Web. Copiez l'URL de la fiche d'évaluation à partir de l'outil Analyst et ajoutez-la au code source des applications externes ou des portails Web.

1. Exécuter une fiche d'évaluation dans l'outil Analyst.
2. Copiez l'URL de la fiche d'évaluation à partir de votre navigateur.
3. Vérifiez que l'URL correspond au format `http://{HOST_NAME}:{PORT}/AnalystTool/com.informatica.at.AnalystTool/index.jsp?`

```
mode=scorecard&project={MRS_PROJECT_ID}&id={SCORECARD_ID}&parentpath={MRS_PARENT_PATH}&view={VIEW_MODE}&pcsfcrcd={CREDENTIAL}.
```

4. Ajoutez l'URL au code source de l'application externe ou du portail Web.

Lignage des fiches d'évaluation

Le lignage des fiches d'évaluation indique l'origine des données, décrit le chemin et indique comment les données circulent pour une métrologie ou un groupe de métrologies. Vous pouvez utiliser le lignage des fiches d'évaluation pour analyser la cause racine d'un écart de score inacceptable dans des métrologies ou des groupes de métrologies. Vous pouvez consulter le lignage des fiches d'évaluation dans l'outil Analyst.

Effectuez les tâches suivantes pour afficher le lignage des fiches d'évaluation :

1. Dans Informatica Administrator, associez un service Metadata Manager au service Analyst.
2. Sélectionnez un projet et exportez les objets de fiche d'évaluation qu'il contient dans un fichier XML à l'aide de l'option Exporter un fichier de ressource pour Metadata Manager dans l'outil Developer ou de la commande `infacmd oie exportResources`.
3. Dans Metadata Manager, utilisez le fichier XML exporté pour créer une ressource et la charger.
Remarque: Le nom du fichier de ressource créé et chargé dans Metadata Manager doit respecter la convention de dénomination suivante : `<nom MRS>_<nom de projet>`. Pour plus d'informations sur la création et le chargement d'un fichier de ressource, voir *Guide de l'utilisateur Informatica PowerCenter Metadata Manager*.
4. Dans l'outil Analyst, ouvrez la fiche d'évaluation et sélectionnez une métrologie ou un groupe de métrologies.
5. Consultez le lignage des fiches d'évaluation.

Affichage du lignage des fiches d'évaluation dans Informatica Analyst

Vous pouvez afficher un diagramme du lignage des fiches d'évaluation pour une métrologie ou un groupe de métrologies. Avant de pouvoir afficher un diagramme de lignage des fiches d'évaluation dans l'outil Analyst, vous devez charger le lignage des fiches d'évaluation et les métadonnées dans Metadata Manager.

1. Dans le navigateur, sélectionnez le projet ou le dossier qui contient la fiche d'évaluation.
2. Cliquez sur la fiche d'évaluation pour l'ouvrir.
La fiche d'évaluation s'affiche dans un onglet.
3. Dans la vue **Fiche d'évaluation**, sélectionnez une métrologie ou un groupe de métrologies.
4. Faites un clic droit et sélectionnez **Afficher le lignage**.

Le diagramme de lignage des fiches d'évaluation s'affiche dans une nouvelle fenêtre.

Important: si vous ne créez pas et ne chargez pas une ressource dans Metadata Manager avec un fichier XML exporté des objets de la fiche d'évaluation, un message d'erreur indiquant que la ressource n'est pas disponible dans le catalogue peut s'afficher. Pour plus d'informations sur l'exportation d'un fichier XML pour un lignage des fiches d'évaluation, consultez ["Exportation d'un fichier de ressource pour le lignage des fiches d'évaluation" à la page 130](#).

CHAPITRE 10

Découverte des domaines de données dans Informatica Analyst

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la découverte des domaines de données dans Informatica Analyst, 75](#)
- [Glossaire de domaine de données dans Informatica Analyst, 75](#)
- [Options de découverte des domaines de données dans Informatica Analyst, 78](#)
- [Création d'un profil pour découvrir des domaines de données dans Informatica Analyst, 80](#)
- [Modification d'un profil dans Informatica Analyst, 81](#)
- [Exécution d'un profil pour effectuer la découverte des domaines de données, 81](#)
- [Résultats de la découverte de domaines de données dans Informatica Analyst, 81](#)
- [Fichiers d'exportation de la découverte des domaines de données dans Informatica Analyst, 84](#)

Présentation de la découverte des domaines de données dans Informatica Analyst

Lorsque vous créez un profil pour effectuer la découverte des domaines de données, sélectionnez les colonnes source, les domaines de données avec lesquels vous voulez comparer les données et le nom des colonnes ainsi que les options d'échantillonnage. Vous pouvez également choisir le nombre maximal de lignes sur lesquelles vous voulez exécuter la découverte des domaines de données et les critères de pourcentage de conformité minimaux.

Glossaire de domaine de données dans Informatica Analyst

Le glossaire de domaine de données répertorie les domaines de données et les groupes de domaines de données. Vous pouvez trier la liste par domaines de données ou groupes de domaines de données. Utilisez le glossaire de domaine de données pour ajouter, modifier et supprimer les domaines de données et groupes de domaines de données. Vous pouvez afficher et modifier la logique de règle associée aux domaines de données dans l'outil Developer.

Création d'un groupe de domaines de données dans Informatica Analyst

Un groupe de domaines de données organise les domaines de données en groupes spécifiques, comme par exemple Informations médicales personnelles (PHI), Informations personnelles identifiables (PII) ou tout autre groupe conceptuel pertinent pour le projet.

1. Cliquez sur **Gérer > Glossaire de domaine de données**.

Le glossaire de domaine de données s'ouvre dans un onglet qui répertorie les domaines de données actuels et les groupes de domaines de données.

2. Dans le navigateur, cliquez sur **Actions > Nouveau > Groupe de domaine de données**.

La boîte de dialogue **Créer un groupe de domaines de données** s'affiche.

3. Entrez un nom et une description pour le groupe de domaines de données.

4. Cliquez sur **Suivant**.

5. Dans le volet **Domaines de données disponibles**, sélectionnez les domaines de données que vous voulez ajouter au groupe de domaines de données et cliquez sur **Ajouter**.

L'outil Analyst déplace les domaines de données sélectionnés dans le volet **Domaines de données sélectionnés**.

6. Cliquez sur **Terminer**.

L'outil Analyst ajoute le groupe de domaines de données au glossaire de domaine de données.

Création d'un domaine de données dans Informatica Analyst

Vous pouvez créer des domaines de données, les ajouter au glossaire de domaine de données et regrouper les domaines de données dans un ou plusieurs groupes de domaines de données. Pour créer un domaine de données, vous pouvez utiliser les règles de données prédéfinies et les règles de noms de colonnes ou générer des domaines de données à partir des valeurs et des modèles des résultats de profil de colonne.

Lorsque vous créez un domaine de données, l'outil Analyst copie les règles et autres objets dépendants associés au domaine de données dans le glossaire de domaine de données. Pour modifier une règle associée à un domaine de données, vous devez accéder à la règle d'origine et la modifier. Vous pouvez ensuite associer à nouveau la règle modifiée au domaine de données.

1. Cliquez sur **Gérer > Glossaire de domaine de données**.

Le glossaire de domaine de données s'ouvre dans un onglet qui répertorie les domaines de données actuels et les groupes de domaines de données.

2. Dans le navigateur, cliquez sur **Actions > Nouveau > Domaine de données**.

La boîte de dialogue **Créer un domaine de données** s'affiche.

3. Entrez un nom et une description pour le domaine de données.

4. Cochez la case **Règle de données** pour rechercher les domaines de données en fonction des données de colonne. Vous pouvez également cocher la case **Règle de nom de colonne** pour rechercher les domaines de données en fonction des titres de colonne dans la source de données.

Le bouton **Sélectionner** est activé.

5. Cliquez sur **Sélectionner** pour ouvrir la boîte de dialogue **Sélectionner une règle**.

6. Sélectionnez une règle appropriée et cliquez sur **OK**.

Les règles que vous avez sélectionnées s'affichent dans les champs **Règle de données** et **Règle de nom de colonne**.

7. Cliquez sur **Suivant**.
8. Dans le volet **Groupes de domaines de données disponibles**, sélectionnez les groupes de domaines de données que vous voulez inclure au domaine de données et cliquez sur **Ajouter**.
L'outil Analyst déplace les groupes de domaines de données sélectionnés dans le volet **Groupes de domaines de données sélectionnés**.
9. Cliquez sur **Terminer**.
L'outil Analyst ajoute le domaine de données au glossaire de domaine de données.

Création d'un domaine de données depuis les résultats de profil dans Informatica Analyst

Exécutez un profil de colonne pour afficher les valeurs et les modèles des données source. Vous pouvez ensuite vérifier les résultats de profil et les utiliser pour créer un domaine de données.

1. Exécutez un profil de colonne pour afficher ses résultats.
Les résultats de profil s'ouvrent dans un nouvel onglet.
2. Dans le volet de gauche, sélectionnez un nom de colonne.
Les valeurs, les modèles et les statistiques de la colonne s'affichent dans le volet de droite.
3. Sélectionnez les valeurs ou les modèles en fonction de ce que vous voulez utiliser pour créer un domaine de données.
4. Effectuez un clic droit sur les valeurs ou les modèles, puis sélectionnez **Nouveau domaine de données**.
La boîte de dialogue **Nouveau domaine de données** s'ouvre.
5. Entrez le nom du domaine de données et éventuellement une description.
L'emplacement est défini sur le glossaire de domaine de données par défaut.
6. Cliquez sur **OK**.
Le domaine de données est ajouté au glossaire de domaine de données.

Rechercher des domaines de données et des groupes de domaines de données dans Informatica Analyst

Vous pouvez rechercher des domaines de données et des groupes de domaines de données spécifiques dans le glossaire de domaine de données. Vous pouvez choisir entre la vue **Domaine de données** et **Groupe de domaines de données** pour afficher la liste des domaines de données dans le glossaire de domaine de données.

Par exemple, vous pouvez avoir un domaine de données **Code postal** que vous avez ajouté au groupe de domaines de données **PII**. Vous pouvez trouver de plus amples informations sur Code postal et son groupe de domaines de données PII des manières suivantes :

Recherchez les domaines de données.

Saisissez une partie du nom de domaine de données comme par exemple **code** ou **postal** dans le champ texte en haut du navigateur. Si vous êtes dans la **Vue Groupes de domaines de données**, l'outil Analyst répertorie **PII**, qui est le groupe de domaines de données contenant **Code postal**. Si vous êtes dans la **Vue de domaine de données**, l'outil Analyst répertorie tous les domaines de données contenant la chaîne de recherche, **Code** ou **postal**.

Remarque: La recherche n'est pas sensible à la casse.

Affichez tous les groupes de domaines de données et les domaines de données qu'ils contiennent.

Dans le navigateur, cliquez sur **Afficher la vue Groupe de domaines de données**.

Affichez tous les domaines de données.

Dans le navigateur, cliquez sur **Afficher la vue Domaine de données**.

Affichez les propriétés d'un domaine de données.

Vérifiez que vous êtes dans **Afficher la vue Domaine de données**. Dans le navigateur, cliquez sur **Code postal** pour afficher ses propriétés dans le volet de droite. Vous pouvez afficher le nom, le type, la description, les règles associées et les groupes de domaine auquel il appartient, dans ce cas, **PII**.

Affichez les propriétés d'un groupe de domaines de données.

Vérifiez que vous êtes dans **Afficher la vue Groupe de domaines de données**. Dans le navigateur, cliquez sur **PII** pour afficher ses propriétés dans le volet de droite. Vous pouvez afficher le nom, le type, la description et la liste des domaines de données, y compris **Code postal**, dans **PII**.

Actualisez le glossaire de domaine de données.

Dans le navigateur, cliquez sur **Actions > Actualiser**. Le glossaire de domaine de données affiche la liste actuelle des domaines de données ou groupes de domaines de données en fonction de la vue dans laquelle vous êtes.

Options de découverte des domaines de données dans Informatica Analyst

Utilisez les options de découverte des domaines de données pour sélectionner les colonnes, les domaines de données et les options d'inférence pour la découverte des domaines de données. Les options d'inférence incluent de choisir si voulez exécuter la découverte des domaines de données en fonction d'une règle sur les données de colonne, le nom des colonnes ou les deux.

Sélection de colonne de domaine de données dans Informatica Analyst

Les options **Colonnes** répertorient toutes les colonnes dans la source de données. Vous pouvez sélectionner les colonnes que vous voulez exécuter dans le cadre d'une découverte de domaines de données.

Le tableau suivant décrit les options **Colonnes** pour la découverte de domaines de données :

Option	Description
Nom	Nom de la colonne.
Type de données	Type de données de la colonne.
Précision	Précision maximum de la colonne.
Échelle	Échelle de la colonne.

Option	Description
Valeur Null autorisée	Indique une colonne pouvant posséder des valeurs null.
Description	Description de la colonne.

Sélection des domaines de données dans Informatica Analyst

Les options **Domaines de données** répertorient tous les domaines de données du glossaire de domaine de données. Vous pouvez choisir les domaines de données que vous voulez exécuter dans le cadre de la découverte de domaines de données.

Le tableau suivant décrit les options **Domaines de données** pour la découverte de domaines de données :

Option	Description
Afficher le groupe de domaines de données dans la hiérarchie	Répertorie tous les groupes de domaine de données, avec les domaines de données regroupés sous chaque groupe de domaines de données.
Nom	Nom du domaine de données.
Groupe de domaines de données	Nom du groupe de domaines de données auquel le domaine de données appartient.
Description	Description du domaine de données.

Options d'inférence des domaines de données dans Informatica Analyst

Les options d'inférence déterminent si la découverte des domaines de données doit s'exécuter sur les données de colonne, les noms de colonne ou les deux. Vous pouvez également spécifier le nombre maximum de lignes que le profil peut analyser et le pourcentage minimum de conformité pour la correspondance du domaine de données.

Le tableau suivant décrit les options d'inférence pour la découverte de domaines de données :

Option	Description
Données	Exécute le profil sur les données de colonne.
Nom de colonne	Exécute le profil sur les titres de colonne.
Nom des colonnes et des données	Exécute le profil sur les données de colonne et les titres de colonne.
Toutes les lignes	Exécute le profil sur toutes les lignes de la source.
Nombre maximum de lignes sur lequel s'exécute un profil	Nombre de lignes maximum sur lequel le profil peut s'exécuter. L'outil Analyst choisit les lignes à partir de la première ligne dans la source.

Option	Description
Pourcentage minimum de conformité	Pourcentage de conformité minimum des données de colonne éligibles pour la correspondance de domaine de données. Le pourcentage de conformité est le rapport entre le nombre de lignes correspondantes et le nombre total de lignes. Remarque: L'outil Analyst considère les valeurs Null comme des lignes qui ne correspondent pas. Les colonnes contenant un grand nombre de valeurs Null ne peuvent pas entraîner d'inférence du domaine de données, sauf si vous spécifiez une valeur faible pour le pourcentage minimum de conformité.
Exclure les colonnes avec un domaine de données approuvé	Exclut les colonnes qui comprennent un domaine de données approuvé issu de l'inférence du domaine de données de l'exécution du profil.

Création d'un profil pour découvrir des domaines de données dans Informatica Analyst

Vous devez créer au moins un domaine de données avant de créer un profil pour effectuer une découverte de domaines de données dans l'outil Analyst. Le profil peut découvrir à la fois les noms de colonne et les données de colonne qui correspondent aux domaines de données prédéfinis.

1. Dans le navigateur, sélectionnez le projet qui contient l'objet de données pour lequel créer un profil.
2. Dans le panneau **Contenu**, cliquez avec le bouton droit sur l'objet de données, puis sélectionnez **Nouveau > Profil**.
L'assistant **Nouveau profil** s'affiche.
3. Sélectionnez **Découverte de domaines de données** ou **Profilage de colonnes et découverte de domaines de données** selon si vous voulez exécuter ou pas un profil de colonne, en plus de la découverte de domaines de données.
4. Cliquez sur **Suivant**.
L'outil Analyst affiche l'objet de données que vous avez sélectionné, ses propriétés, et l'aperçu des données.
5. Cliquez sur **Suivant**.
6. Modifiez le nom par défaut et la description du profil, si nécessaire. Vous pouvez également modifier l'emplacement du dossier dans lequel vous souhaitez enregistrer le profil.
7. Cliquez sur **Suivant**.
8. Sélectionnez les colonnes sur lesquelles vous voulez exécuter la découverte de domaines de données et les domaines de données auxquels vous voulez comparer les colonnes.
9. Décidez d'exécuter la découverte de domaines de données sur les données de colonne, sur les noms de colonne ou sur les deux, et indiquez le pourcentage de conformité minimum.
10. Dans le panneau **Options d'échantillonnage**, configurez les options correspondantes.
11. Si vous avez sélectionné **Profilage de colonnes et découverte de domaines de données** comme option de profil, configurez les options de profilage de colonnes, y compris les options de filtre.
12. Cliquez sur **Enregistrer** pour créer le profil ou sur **Enregistrer et exécuter** pour créer le profil puis l'exécuter.

Modification d'un profil dans Informatica Analyst

Vous pouvez modifier les propriétés d'un profil après l'avoir exécuté. Les modifications incluent les propriétés de base telles que le nom, la description et le type de profil. Si vous avez exécuté un profil de colonne dans le cadre d'une découverte de domaines de données, vous pouvez également modifier les paramètres de profil de colonne.

1. Dans le navigateur, sélectionnez le projet ou le dossier qui contient le profil que vous souhaitez modifier.
2. Cliquez sur le profil pour l'ouvrir.
Le profil s'ouvre dans un onglet.
3. Cliquez sur **Actions > Éditer**.
Un menu contextuel s'affiche à l'écran.
4. Sélectionnez l'une des options de menu suivantes en fonction des modifications que vous voulez effectuer :
 - **Général**. Modifiez des propriétés de base telles que le nom, la description et le type de profil.
 - **Source de données**. Sélectionnez une autre source de données correspondante.
 - **Profilage de colonnes**. Sélectionnez les colonnes dans lesquelles vous voulez exécuter un profil de colonne, puis configurez les options d'échantillonnage et de développement requises.
 - **Filtres de profilage de colonnes**. Créez, modifiez et supprimez des filtres.
 - **Règles de profilage de colonnes**. Créez et modifiez les règles de profil de colonne.
 - **Découverte de domaines de données**. Configurez les options de découverte de domaines de données.
5. Cliquez sur **Enregistrer** pour enregistrer les modifications ou sur **Enregistrer et exécuter** pour enregistrer les modifications puis exécuter le profil.

Exécution d'un profil pour effectuer la découverte des domaines de données

Exécutez un profil dans le cadre de la découverte des domaines de données afin d'afficher les colonnes qui correspondent aux modèles de règles de domaine de données.

1. Dans le Navigateur, sélectionnez le projet ou le dossier qui contient le profil à exécuter.
2. Cliquez sur le profil pour l'ouvrir.
Le profil s'ouvre dans un onglet.
3. Cliquez sur **Actions > Exécuter le profil**.

Résultats de la découverte de domaines de données dans Informatica Analyst

Les résultats de la découverte de domaines de données affichent des statistiques sur les colonnes qui correspondent aux domaines de données, y compris le pourcentage de données de colonne correspondantes

et si chaque nom de colonne correspond au domaine de données. Vous pouvez développer les résultats pour analyse et vérifier les résultats sur toutes les lignes de la source de données.

Vous pouvez trier les résultats en fonction des domaines de données, des groupes de domaines de données et des colonnes.

Le tableau suivant décrit les résultats de la découverte de domaines de données :

Nom de colonne	Description
Nom	Nom du domaine de données, du groupe de domaines de données ou de la colonne selon si vous sélectionnez la vue Vue par domaines de données , Vue par groupe de domaines de données ou Vue par colonnes . Remarque: La vue par défaut répertorie tous les domaines de données.
% de conformité des données	Pourcentage des données de colonne qui correspondent à un domaine de données.
Conforme au nombre de lignes	Nombre total de lignes qui correspondent à un domaine de données.
% Null	Pourcentage de valeurs Null pour la colonne.
Lignes totales	Nombre total de lignes dans la source de données.
Correspondance de nom de colonne	Détermine si un nom de colonne correspond à un nom de domaine de données.
Groupe de domaines de données	Groupe de domaines de données auquel appartient le domaine de données.
Type de données consigné	Type de données déclaré pour la colonne de l'objet de profil.
Statut	Statut d'inférence du domaine de données. Les statuts peuvent être Inféré, Approuvé et Rejeté. <ul style="list-style-type: none">- Inféré. Indique le domaine de données inféré pour la colonne.- Approuvé. Indique un domaine de données approuvé pour la colonne. Lorsque vous approuvez un domaine de données, vous validez le domaine de données pour le référentiel modèle.- Rejeté. Indique un domaine de données rejeté pour la colonne.
Développement	Lorsqu'elle est sélectionnée, cette option développe les données de colonne.
Vérifié	Détermine si la validation de la correspondance du domaine de données sur toutes les lignes de la source de données est complète.
Heure de la dernière exécution de profil	Date et heure de la dernière exécution du profil.

Affichage par groupes de domaines de données dans Informatica Analyst

Vous pouvez afficher les résultats de domaines de données triés par groupes de domaines de données.

1. Exécutez le profil pour afficher ses résultats.
2. Cliquez sur **Découverte des domaines** si l'onglet n'est pas sélectionné par défaut.
3. Dans la liste au dessus du champ **Nom**, sélectionnez **Afficher par groupe de domaines de données**.

Vous pouvez afficher les résultats de colonne pour chaque groupe de domaines de données.

Affichage par colonnes dans Informatica Analyst

Vous pouvez afficher les résultats de domaines de données triés par colonnes dans la source de données.

1. Exécutez le profil pour afficher ses résultats.
2. Cliquez sur **Découverte des domaines** si l'onglet n'est pas sélectionné par défaut.
3. Dans la liste au dessus du champ **Nom**, sélectionnez **Afficher par colonnes**.

Vous pouvez afficher les résultats pour chaque colonne, puis triés par groupes de domaines de données associés avec la colonne.

Approbation des domaines de données

Vous pouvez approuver plusieurs domaines de données dans l'outil Analyst.

1. Dans le navigateur, sélectionnez le projet ou le dossier qui contient le profil.
2. Cliquez sur le profil pour l'ouvrir.
Le profil s'ouvre dans un onglet.
3. Dans la vue **Découverte de domaines de données**, sélectionnez une ligne.
4. Cliquez sur **Actions > Approuver**.
Le statut de la colonne ou du domaine de données devient Approuvé.
5. Pour restaurer le statut inféré de la colonne ou du domaine de données, sélectionnez la ligne, puis cliquez sur **Actions > Réinitialiser**.

Rejet des domaines de données

Lorsque vous ouvrez les résultats de profil, l'outil Analyst affiche les domaines de données approuvés par défaut. Vous pouvez afficher ou masquer les domaines de données rejetés.

1. Dans le navigateur, sélectionnez le projet ou le dossier qui contient le profil.
2. Cliquez sur le profil pour l'ouvrir.
Le profil s'ouvre dans un onglet.
3. Sélectionnez une ligne dans la vue **Découverte de domaines de données**.
4. Pour rejeter les domaines de données inférés, cliquez sur **Actions > Rejeter**.
L'outil Analyst supprime le domaine de données rejeté des résultats de la découverte de domaines de données.
5. Pour afficher les domaines de données rejetés, cliquez sur **Actions > Afficher les éléments rejetés**.
6. Pour masquer les domaines de données rejetés, cliquez sur **Actions > Masquer les éléments rejetés**.

Fichiers d'exportation de la découverte des domaines de données dans Informatica Analyst

Lorsque vous exportez les résultats de la découverte des domaines de données de l'outil Analyst, vous pouvez spécifier les valeurs de nom de fichier, format de fichier ainsi que la page de code. Vous pouvez exporter les résultats de la découverte des domaines de données vers un fichier Microsoft Excel.

Le fichier Microsoft Excel contient plusieurs feuilles de calcul séparant les résultats de découverte en fonction des colonnes, des domaines de données et des groupes de domaines de données. La feuille de calcul des propriétés affiche les propriétés du profil, telles que le nom, la description, le type, l'emplacement, les date et heure des dernières modifications apportées au profil ainsi qu'un lien vers le profil.

Résultats de la découverte des domaines de données dans Microsoft Excel

Lorsque vous exportez les résultats de la découverte des domaines de données vers Microsoft Excel, l'outil Analyst enregistre les noms de colonne, les noms de domaines de données correspondants ainsi que le pourcentage de conformité et de valeurs null. Le fichier Excel contient également les noms des groupes de domaine de données de chaque domaine de données ainsi que les types de données documentés des colonnes.

Le tableau suivant décrit chaque feuille de calcul dans le fichier d'exportation :

Tab	Description
Vue par colonnes	Résultats de la découverte des domaines de données triés par colonne de source de données.
Vue par domaines de données	Résultats de la découverte des domaines de données triés par domaine de données.
Vue par groupes de domaines de données	Résultats de la découverte des domaines de données triés par groupes de domaines de données.
Propriétés	Les propriétés du profil de base, telles que le nom, la description, le type, l'emplacement, les date et heure des dernières modifications apportées au profil ainsi qu'un lien vers le profil.

Exportation des résultats de la découverte des domaines de données depuis Informatica Analyst

Vous pouvez exporter les résultats de la découverte des domaines de données vers un fichier `.xlsx` afin de pouvoir afficher les données dans un fichier et distribuer celui-ci au sein de l'entreprise.

1. Exécuter un profil pour effectuer la découverte des domaines de données.
2. Cliquez sur l'onglet **Découverte des domaines de données**.
3. Cliquez sur **Actions > Exporter les données**.
La boîte de dialogue **Exporter les données vers un fichier** s'ouvre.
4. Entrez le nom de fichier. Utilisez éventuellement le nom de fichier par défaut.
5. Sélectionnez la page de codes du fichier.

6. Cliquez sur **OK**.

CHAPITRE 11

Découverte des données d'entreprise dans Informatica Analyst

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la découverte des données d'entreprise dans Informatica Analyst, 86](#)
- [Processus de découverte des données d'entreprise dans l'outil Analyst, 87](#)
- [Options de configuration pour la découverte des données d'entreprise, 87](#)
- [Exécution d'un profil de découverte des données d'entreprise des données dans l'outil Analyst, 89](#)
- [Modification des options de découverte des données d'entreprise, 90](#)

Présentation de la découverte des données d'entreprise dans Informatica Analyst

La découverte des données d'entreprise est le processus de découverte de métadonnées de colonne et de domaines de données dans plusieurs sources de données provenant de nombreux schémas et connexions relationnelles externes. Vous pouvez effectuer la découverte des données d'entreprise sur les sources de données que vous avez importées dans le référentiel modèle et sur celles provenant de connexions relationnelles externes.

En tant qu'analyste de données, vous pouvez effectuer la découverte des données d'entreprise dans l'outil Analyst pour inférer des caractéristiques de métadonnées spécifiques sur un grand nombre de sources de données. Vous pouvez également afficher les données source correspondant à des domaines de données prédéfinis. Vous avez alors la possibilité de conserver les résultats inférés de la découverte des données d'entreprise et de préparer les données aux initiatives de recherche de découverte et de qualité des données. Dans l'outil Analyst, la découverte des données d'entreprise génère un résumé consolidé des résultats de profil.

Les résultats de la découverte des données d'entreprise comprennent les statistiques de profil de colonne, telles que les modèles, les valeurs uniques et les colonnes avec les conflits de type de données. La découverte de domaines de données identifie les colonnes source qui correspondent aux domaines de données prédéfinis.

Processus de découverte des données d'entreprise dans l'outil Analyst

Vous pouvez créer, modifier et supprimer des profils de découverte des données d'entreprise. Vous pouvez exécuter un profil de découverte des données d'entreprise dans l'espace de travail Découverte. Vous devez configurer les options d'inférence du profil de colonne et de découverte de domaines de données avant d'exécuter le profil de découverte des données d'entreprise.

Suivez les étapes suivantes pour effectuer la découverte des données d'entreprise dans l'outil Analyst :

1. Configurez les propriétés générales du profil de découverte des données d'entreprise.
2. Dans le référentiel modèle, sélectionnez les objets de données que vous souhaitez inclure dans le profil de découverte des données d'entreprise.
3. Importez les sources de données relationnelles depuis des connexions de base de données externes.
4. Configurez les options d'inférence de données et de découverte pour le profil de découverte des données d'entreprise.
5. Enregistrez les modifications et exécutez le profil de découverte des données d'entreprise.
6. Surveillez l'exécution du profil et, si nécessaire, affichez les statuts des tâches de profil exécutées par l'outil Analyst.
7. Consultez le résumé des résultats de découverte des données d'entreprise. Les résultats s'affichent dans les volets **Résumé** et **Profils**.

Options de configuration pour la découverte des données d'entreprise

Les options de configuration pour la découverte des données d'entreprise comprennent des options de découverte de domaines de données, des options d'échantillonnage de profil de colonne et des propriétés générales de profil comme le nom et la description.

Vous pouvez choisir d'exécuter un profil de colonne ou un profil pour effectuer la découverte de domaines de données. Vous pouvez également choisir d'exécuter à la fois un profil de colonne et un profil pour effectuer la découverte de domaines de données dans le cadre de la configuration.

Paramètres de découverte de domaines de données

Parmi les paramètres de découverte de domaines de données, il y a l'option entre exécuter la découverte sur les données des colonnes, sur le nom des colonnes ou sur les deux. Vous pouvez choisir des domaines de données et indiquer si la découverte de domaines de données doit traiter toutes les lignes de la source de

données. Vous pouvez également définir un pourcentage de conformité minimum pour l'inférence de domaines de données.

Le tableau suivant expose les paramètres de découverte de domaines de données que vous pouvez configurer pour la découverte des données d'entreprise dans l'outil Analyst :

Option	Description
Activer la découverte de domaines de données	Exécute la découverte de domaines de données dans le cadre de la découverte des données d'entreprise.
Exécuter la découverte de domaines de données sur les données	Exécute la découverte de domaines de données sur les données des colonnes.
Exécutez la découverte de domaines de données sur le nom des colonnes	Exécute la découverte de domaines de données sur le nom de chaque colonne.
Toutes les lignes	Exécute la découverte de domaines de données sur toutes les lignes source.
Nombre maximum de lignes de la découverte de domaines de données	Nombre de lignes maximum sur lequel le profil peut s'exécuter. L'outil Analyst choisit les lignes à partir de la première ligne dans la source.
Pourcentage minimum de conformité	Pourcentage minimum de conformité des données de colonne éligibles pour la correspondance de domaine de données. Le pourcentage de conformité est le rapport entre le nombre de lignes correspondantes et le nombre total de lignes. Remarque: L'outil Analyst considère les valeurs Null comme des lignes qui ne correspondent pas.
Exclure les colonnes avec des domaines de données approuvés	Exclut les colonnes qui comprennent des domaines de données approuvés issus de l'inférence du domaine de données de l'exécution du profil.

Options d'échantillonnage de profil de colonne

Les options d'échantillonnage déterminent si l'outil Analyst exécute un profil de colonne sur toutes les lignes des sources de données ou sur un nombre limité de lignes.

Le tableau suivant décrit les options d'échantillonnage de profil de colonne que vous pouvez configurer pour la découverte des données d'entreprise :

Option	Description
Activer le profilage de colonnes	Exécute un profil de colonne dans le cadre de la découverte des données d'entreprise.
Toutes les lignes	Exécute un profil de colonne sur toutes les lignes de la source de données.
<nombre> première(s) ligne(s)	Nombre de lignes sur lequel vous voulez exécuter le profil de colonne. L'outil Analyst choisit les lignes à partir de la première ligne dans la source de données.
Exclure l'inférence du type de données pour les colonnes avec un type de données approuvé	Exclut les colonnes qui comprennent un type de données approuvé de l'inférence du type de données de l'exécution du profil.

Exécution d'un profil de découverte des données d'entreprise des données dans l'outil Analyst

Vous pouvez exécuter le profil de colonne et la découverte de domaines de données dans le cadre de la découverte des données d'entreprise dans l'outil Analyst.

1. Dans le panneau d'accueil **Découverte**, sélectionnez **Profil de découverte des données d'entreprise**.
L'assistant **Nouveau profil de découverte des données d'entreprise** s'affiche.
2. Entrez un nom de profil et éventuellement une description.
3. Vérifiez l'emplacement du projet. Si nécessaire, accédez à un autre emplacement.
4. Cliquez sur **Suivant**.
L'onglet **Sélectionner des objets de données** s'affiche.
5. Si vous avez sélectionné un objet de données dans le référentiel modèle, son nom s'affiche dans la section **Objets de données sélectionnés**. Cliquez sur **Choisir** pour sélectionner d'autres d'objets de données, si nécessaire.
6. Cliquez sur **Suivant** pour ouvrir l'onglet **Sélectionner des ressources**.
Vous pouvez importer des données provenant de plusieurs sources de données relationnelles.
7. Cliquez sur **Choisir** pour ouvrir la boîte de dialogue **Choisir les ressources**.
Le volet de gauche de la boîte de dialogue répertorie toutes les connexions internes et externes, les schémas, les tables et les vues du domaine Informatica.
8. Sélectionnez les connexions, les schémas, les tables et les vues que vous souhaitez inclure dans le profil de découverte des données d'entreprise.
9. Cliquez sur **Enregistrer** pour enregistrer les modifications et fermez la boîte de dialogue.
10. Cliquez sur **Suivant**.
L'onglet **Spécifier les paramètres de découverte** s'affiche. Vous pouvez configurer les options de profil de colonne et les options de découverte de domaines de données sur la page Spécifier les paramètres de découverte.
11. Sélectionnez **Activer la découverte de domaines de données** pour découvrir les domaines de données qui se trouvent dans les données source.
12. Cliquez sur **Choisir** pour sélectionner les domaines de données que vous souhaitez découvrir.
La boîte de dialogue **Choisir les domaines de données** s'affiche.
13. Sélectionnez les domaines de données requis, puis cliquez sur **OK**.
14. Configurez les options de profil de colonne, y compris le nombre de lignes dans lesquelles le profil s'exécute.
15. Cliquez sur **Enregistrer et exécuter** en haut de l'onglet **Spécifier les paramètres de découverte**.
Vous pouvez voir un message de statut indiquant que l'exécution du profil est en cours. Après l'exécution du profil, vous pouvez afficher les résultats de la découverte des données d'entreprise dans les onglets **Résumé** et **Profils**.

Modification des options de découverte des données d'entreprise

Après avoir effectué une découverte des données d'entreprise, vous pouvez en modifier les options. Vous pouvez renommer le profil et modifier la sélection des objets de données, celle des domaines de données et les options d'inférence.

1. Ouvrez un profil que vous avez exécuté pour effectuer la découverte des données d'entreprise.
Les résultats de profil s'affichent dans l'espace de travail **Découverte**.
2. Cliquez sur **Éditer le profil**.
La page **Spécifier les propriétés générales** s'affiche.
3. Modifiez les propriétés de profil selon les besoins.
4. Pour modifier la sélection des objets de données, cliquez sur **Sélectionner des objets de données**.
5. Pour modifier les sources de données externes de la découverte des données d'entreprise, cliquez sur **Sélectionner des ressources**.
6. Pour modifier les options d'inférence du domaine de données et les paramètres de profil de colonne, cliquez sur **Spécifier les paramètres de découverte**.
7. Pour appliquer les modifications de configuration à toutes les tâches de profil de domaine de données et aux tâches de profil de colonne dans le profil de découverte des données d'entreprise, sélectionnez **Utiliser les paramètres globaux pour tous les profils**. Si vous ne sélectionnez pas cette option, les modifications apportées aux paramètres de profil s'appliquent uniquement aux objets de données ou aux ressources nouvellement ajoutés au profil.

Par défaut, les modifications que vous effectuez s'appliquent aux objets de données nouvellement ajoutés au profil de découverte des données d'entreprise.
8. Cliquez sur **Annuler** pour annuler les modifications.
9. Cliquez sur **Enregistrer et exécuter** pour enregistrer les modifications et réexécuter le profil.

CHAPITRE 12

Résultats de la découverte des données d'entreprise dans Informatica Analyst

Ce chapitre comprend les rubriques suivantes :

- [Présentation des résultats de la découverte des données d'entreprise dans l'outil Analyst, 91](#)
- [Vue Résumé, 91](#)
- [Conflit de type de données, 94](#)
- [Vue Profils, 94](#)

Présentation des résultats de la découverte des données d'entreprise dans l'outil Analyst

Vous pouvez afficher les résultats de la découverte des données d'entreprise dans les vues **Résumé** et **Profils**.

La vue **Résumé** affiche les résultats de profil de colonne et ceux de la découverte de domaines de données. La section **Découverte de domaines de données** répertorie les domaines de données que vous avez inclus dans l'exécution du profil et le nombre de colonnes qui présentent une correspondance de domaines de données. La section **Profilage de colonnes** affiche les statistiques dans les colonnes source. Vous pouvez cliquer sur chaque ligne de résultats de profil pour afficher les détails dans le volet de droite de la vue **Résumé**.

Vue Résumé

La vue **Résumé** affiche un résumé des résultats de profil de colonne et de ceux de la découverte de domaines de données. Vous pouvez afficher les noms de domaines de données qui présentent des correspondances dans les colonnes et le nombre de colonnes qui présentent des correspondances dans les domaines de données. Les statistiques de colonne comprennent le nombre de colonnes des 10 premières correspondances de modèles, toutes les valeurs uniques et toutes les valeurs Null. Les statistiques de

colonne comprennent également le nombre de colonnes qui présentent des conflits de type de données entre les types de données inférés et consignés.

Résultats de profil de la vue Résumé

La vue Résumé affiche les résultats de la découverte des données d'entreprise dans les sections Découverte de domaines de données et Profilage de colonnes.

Découverte de domaines de données

Le tableau suivant décrit les colonnes des résultats de la découverte de domaines de données :

Nom de colonne	Description
Nom	Nom du domaine de données.
Trouvé dans les colonnes	Nombre total de colonnes qui présentent une correspondance de domaine de données.
Profil	Nom du profil qui contient la colonne de correspondance.
Nom de colonne	Nom de la colonne de correspondance.
Pourcentage de conformité des données	Pourcentage de conformité de la correspondance de domaine de données.
Nom de connexion	Nom de connexion de la base de données relationnelle.
Statut de l'inférence	Statut de l'inférence du domaine de données. Les statuts sont Accepté , Rejeté et Inféré .

Profilage de colonnes

Le tableau suivant décrit les colonnes des résultats de profil de colonne :

Nom de colonne	Description
Nom	Nom du type de résultats de profil, tel qu'un modèle, des données 100% Nulls et 100% Unique.
Trouvé dans les colonnes	Nombre total de colonnes qui présentent un type de résultats correspondant au profil.
Profil	Nom du profil qui contient la colonne de correspondance.
Connexion	Nom de connexion de la base de données relationnelle.
Source de données	Source de données du profil.
Nombre de colonnes	Nombre de colonnes du profil qui présentent un type de résultats correspondant au profil.

Affichage des résultats de la découverte de domaines de données dans l'outil Analyst

Vous pouvez cliquer sur un nom de domaine de données pour en afficher les résultats de la découverte de domaines de données. Vous pouvez ouvrir des profils spécifiques à partir des résultats de la découverte de domaines de données.

1. Exécutez un profil pour effectuer la découverte des données d'entreprise.
2. Vérifiez que vous êtes dans la vue **Résumé**.
3. Cliquez sur un domaine de données dans la section **Découverte de domaines de données** pour en afficher les résultats de découverte.

Une liste de profils contenant le domaine de données s'affiche dans le volet de droite.

4. Sélectionnez une ligne dans le volet de droite, si nécessaire.

Les liens hypertexte menant au profil s'affichent en bleu.

5. Cliquez sur le lien du nom de profil ou celui du nom de colonne pour ouvrir le profil.

Le profil s'ouvre et affiche les résultats de la découverte de domaines de données. L'outil Analyst met en évidence la ligne qui comprend le domaine de données dans les résultats. Si nécessaire, vous pouvez conserver les résultats de profil pour une utilisation ultérieure plus efficace, telle que la recherche de découverte.

6. Pour revenir à la vue **Résumé**, cliquez sur le dossier ou le lien du nom de projet dans le coin supérieur gauche de l'espace de travail **Découvertes**.

Affichage des résultats de profil de colonne

Vous pouvez afficher les résultats de profil de colonne de la découverte des données d'entreprise dans la vue **Résumé**. Vous pouvez ouvrir des profils spécifiques à partir des résultats de la découverte de domaines de données.

1. Exécutez un profil pour effectuer la découverte des données d'entreprise.
2. Vérifiez que vous êtes dans la vue **Résumé**.
3. Pour afficher les détails des modèles inférés, cliquez sur les 10 premiers modèles de la section **Profilage de colonnes**.

Une liste de profils contenant les résultats du modèle inféré s'affiche dans le volet de droite.

4. Pour afficher les informations telles que toutes les valeurs Null, toutes les valeurs uniques ou les conflits de type de données, cliquez sur **100% Nulls**, **100% Unique** ou **Types de données inférés et consignés en conflit**.

La liste des profils correspondants s'affiche dans le volet de droite.

5. Cliquez sur le lien du nom de profil ou celui du nom de colonne pour ouvrir le profil.

Le profil s'ouvre et affiche les résultats de profil de colonne.

6. Pour revenir à la vue **Résumé**, cliquez sur le dossier ou le lien du nom de projet dans le coin supérieur gauche de l'espace de travail **Découvertes**.

Conflit de type de données

La découverte des données d'entreprise identifie les conflits de type de données dans les colonnes. Un conflit de type de données est une différence des types de données inférés et consignés dans une colonne après exécution de la découverte des données d'entreprise. Un type de données inféré est le type de données provenant de l'outil Analyst pour une colonne de source de données en fonction des données de cette dernière. Un type de données consigné est le type de données déclaré pour une colonne dans la base de données source.

La découverte des données d'entreprise peut inférer un autre type de données pour une colonne en fonction des données de la colonne par rapport au type de données consigné d'une colonne. Par exemple, la découverte des données d'entreprise peut inférer une colonne qui comprend un type de données String documenté en tant que type de données Date. Vous pouvez examiner le conflit de type de données, choisir la date de type de données la plus appropriée pour la colonne et l'approuver.

Affichage des conflits de type de données

Lorsque vous ouvrez un profil présentant des conflits de type de données dans la vue **Résumé**, l'outil Analyst met en évidence les conflits de type de données en rouge.

1. Exécutez un profil pour effectuer la découverte des données d'entreprise.
2. Vérifiez que vous êtes dans la vue **Résumé**.
3. Dans la section **Profilage de colonnes**, cliquez sur **Types de données inférés et consignés en conflit** pour afficher les conflits de type de données dans les résultats de profil de colonne.
Une liste de profils contenant des colonnes avec des conflits de type de données s'affichent dans le volet de droite.
4. Sélectionnez une ligne dans le volet de droite, si nécessaire.
Les liens hypertexte menant au profil s'affichent en bleu.
5. Cliquez sur le lien du nom de profil ou celui du nom de colonne pour ouvrir le profil.
Le profil s'ouvre et affiche les conflits de type de données en rouge. Vous pouvez conserver les types de données inférés pour résoudre le conflit de type de données.
6. Pour conserver les types de données, sélectionnez une ligne avec les types de données en conflit, puis cliquez sur **Types de données**.
7. Cliquez sur **Actions**, puis sélectionnez **Approuver** ou **Rejeter**.
8. Pour revenir à la vue **Résumé**, cliquez sur le dossier ou le lien du nom de projet dans le coin supérieur gauche de l'espace de travail **Découvertes**.

Vue Profils

La vue **Profils** affiche une liste de tous les profils d'objet de données uniques exécutés par l'outil Analyst dans le cadre de la découverte des données d'entreprise. La liste de profils affiche également le statut

d'exécution de chaque profil. Vous pouvez ouvrir chaque profil pour afficher les résultats de profil de colonne et ceux de la découverte de domaines de données.

Affichage des propriétés de profil

Vous pouvez afficher la liste des profils qui font partie de la découverte des données d'entreprise dans la vue **Profils**. Vous pouvez ouvrir chaque profil et conserver les résultats de profil, si nécessaire.

1. Exécutez un profil pour effectuer la découverte des données d'entreprise.
2. Vérifiez que vous êtes dans la vue **Profils**.
3. Pour afficher les propriétés d'un profil, cliquez sur le nom du profil.

Les propriétés du profil s'affichent dans le volet de droite. Elles comprennent le nom des objets de données source, le nom de connexion et le nombre de lignes.

4. Pour afficher les résultats de profil, cliquez sur **Ouvrir le profil**.
Le profil affiche les résultats de profil de colonne.
5. Pour revenir à la vue **Profils**, cliquez sur le dossier ou le lien du nom de projet dans le coin supérieur gauche de l'espace de travail **Découvertes**.

CHAPITRE 13

Recherche de découverte dans Informatica Analyst

Ce chapitre comprend les rubriques suivantes :

- [Recherche de découverte dans Informatica Analyst - Présentation, 96](#)
- [Prérequis de la recherche de découverte, 97](#)
- [Processus de recherche de découverte dans Informatica Analyst, 98](#)
- [Options de recherche de découverte, 98](#)
- [Résultats de la recherche de découverte dans Informatica Analyst, 100](#)
- [Types de correspondances, 102](#)
- [Ressources liées, 103](#)
- [Forum Aux Questions \(FAQ\), 105](#)

Recherche de découverte dans Informatica Analyst - Présentation

La recherche de découverte recherche des ressources et identifie les relations avec d'autres ressources dans les bases de données et les schémas de l'entreprise. Les utilisateurs peuvent utiliser cette recherche pour localiser les données et les métadonnées dans l'entreprise. Vous pouvez rechercher des ressources spécifiques, telles que des objets de données, des règles et des profils.

Si vous effectuez une recherche globale, l'outil Analyst effectue une recherche basée sur du texte pour les objets de données, les sources de données et les dossiers. Si vous effectuez une recherche de découverte, outre les correspondances de texte, les résultats incluent les objets qui ont des relations avec les objets répondant aux critères de recherche. La recherche de découverte inclut également des correspondances basées sur les métadonnées de profil, telles que les types et les modèles de données. Par exemple, vous pouvez trouver des objets qui contiennent un modèle de données spécifique et dont le nom comporte un mot-clé donné.

Les résultats de la recherche de découverte incluent les types d'informations suivants :

Objets du référentiel modèle

La recherche porte sur les objets principaux associés aux objets qui répondent aux critères. Par exemple, lorsque vous recherchez un profil, les résultats incluent l'objet de données du profil.

Résultats de l'entrepôt de profils

Incluent les résultats d'inférence des profils, tels qu'un domaine de données ou un modèle de données.

Termes de glossaire métier

Selon la licence, la recherche inclut les métadonnées, telles qu'un terme métier associé à une règle.

Exemple de recherche de découverte

Vous êtes gestionnaire de données et à ce titre vous êtes chargé de veiller à ce que les données sensibles de l'entreprise soient correctement masquées. Vous souhaitez peut-être identifier des informations personnelles identifiables (PII) dans les schémas et les bases de données sur lesquels vous ou l'architecte de données avez exécuté la découverte des données d'entreprise. Vous avez pu créer des domaines de données pour identifier des données importantes qui n'ont pas été découvertes dans les sources de données. Vous effectuez une recherche sur la chaîne « SSN » (numéro de sécurité sociale). L'outil Analyst affiche le domaine de données de sécurité sociale et toutes les colonnes correspondantes des sources de données. La chaîne de recherche de découverte peut également identifier des colonnes ou des tables supplémentaires dont la description ou le nom contient la chaîne « SSN ». Pour affiner votre recherche, vous pouvez filtrer les spécifications de mappage de manière à afficher celles qui font référence aux objets de données correspondants. Vous pouvez appliquer des filtres supplémentaires pour filtrer d'autres spécifications de mappage en fonction des projets ou des utilisateurs. Vous pouvez ensuite ouvrir les spécifications de mappage dans les résultats pour vérifier si elles sont conformes aux stratégies de confidentialité de l'entreprise.

Prérequis de la recherche de découverte

Avant de pouvoir effectuer une recherche de découverte efficace dans les bases de données de l'entreprise, procédez à une découverte des données d'entreprise sur les bases de données et les schémas de l'entreprise.

Une fois la découverte des données d'entreprise effectuée, l'outil Analyst stocke tous les résultats de profil dans l'entrepôt de profilage. Vérifiez que toutes les sources de données requises sont dans le référentiel modèle. Vérifiez éventuellement que des termes métier sont associés aux ressources appropriées dans le référentiel modèle. Lorsque vous effectuez une recherche de découverte, le service de recherche récupère les informations de l'index de recherche en fonction des ressources du référentiel modèle et des résultats de l'entrepôt de profilage. Le service de recherche utilise ensuite les informations indexées pour afficher les résultats de la recherche en fonction des métadonnées et des relations appropriées de l'objet.

Processus de recherche de découverte dans Informatica Analyst

Vous pouvez rechercher des ressources en fonction de critères comme le texte, les modèles et les types de données dans les résultats de profil. La recherche renvoie une liste de ressources liées à la chaîne de recherche.

Suivez la procédure ci-après pour effectuer une recherche de découverte dans l'outil Analyst :

1. Effectuez la découverte des données d'entreprise et exécutez les profils d'objet de données uniques requis sur les sources de données de l'entreprise. Lorsque vous effectuez la recherche de découverte, l'outil Analyst recherche les informations dans les résultats de profil et les objets du référentiel modèle.
2. Choisissez le type d'informations à rechercher. Par exemple, vous pouvez rechercher toutes les ressources associées à une définition de domaine de données pour des données sensibles ou un modèle de données spécifique.
3. Effectuez la recherche.
4. Analysez les résultats de la recherche pour identifier les ressources et leurs relations avec d'autres ressources.
5. Si nécessaire, vérifiez que les données découvertes sont conformes aux spécifications de l'entreprise.

Options de recherche de découverte

Vous pouvez effectuer une recherche globale ou une recherche de découverte pour trouver des ressources et identifier des relations avec d'autres ressources. Les résultats de la recherche globale proviennent du référentiel modèle et éventuellement du glossaire métier. Les résultats de la recherche de découverte proviennent du référentiel modèle et du glossaire métier, mais aussi des profils, selon les résultats de profil de l'entrepôt de profilage.

Vous pouvez rechercher des ressources, telles que des objets de données, des profils et des spécifications de mappage. Entrez une chaîne de recherche pour rechercher des ressources qui correspondent à la chaîne de recherche et présentent une association avec cette dernière. Vous pouvez utiliser des caractères génériques pour rechercher des ressources.

Les caractères génériques suivants sont disponibles pour la recherche de ressources :

*** (astérisque)**

Ajoutez ce caractère à la fin de la chaîne de recherche pour rechercher tous les noms de ressources qui commencent par cette chaîne. Par exemple, pour rechercher tous les noms de ressource qui commencent par la chaîne « emp », vous pouvez saisir « emp* » dans le champ de recherche.

? (point d'interrogation)

Ajoutez ce caractère dans la chaîne de recherche pour représenter un caractère alphanumérique.

Remarque: Pour la recherche de ressources, la chaîne de recherche ne peut pas commencer par un caractère générique. La recherche n'est pas sensible à la casse.

Pour rechercher deux mots ou plus formant une expression, placez ces mots entre guillemets doubles. Utilisez le caractère + pour représenter l'opérateur AND et rechercher un terme qui doit apparaître dans les résultats de la recherche. Par exemple, si la chaîne de recherche est +données +sensibles, le service de recherche trouvera les métadonnées qui contiennent les deux termes. Utilisez un espace vide pour

représenter l'opérateur OR. Par exemple, si la chaîne de recherche est `données sensibles`, le service de recherche trouve les métadonnées qui contiennent l'un ou l'autre de ces termes.

Si la chaîne de recherche contient un tiret (-), un trait de soulignement (_) ou une casse mixte, le service de recherche trouve le mot entier et les différentes parties du mot séparées par le délimiteur. Par exemple, si vous recherchez `Profile_Customer`, le moteur de recherche trouve `Profil`, `Client` et `Profile_Customer` dans les référentiels. Pour inclure des caractères spéciaux, tels que * et ?, dans votre chaîne de recherche, placez la chaîne contenant ces caractères entre guillemets doubles.

Vous pouvez exécuter une recherche de découverte qui comprend une recherche par mot-clé et un filtre de découverte. Supposons que vous recherchez des colonnes d'ID d'employé qui utilisent le format `<InitialPrénom><InitialNom>-<NuméroSécuritéSociale>` pour identifier les risques en matière de sécurité des données. Pour rechercher les colonnes, saisissez `ID employé` dans le panneau **Recherche** de l'espace de travail de bibliothèque et définissez le filtre de modèle sur `XX-999999999 <= 100%`.

Critères de recherche de découverte

Utilisez les critères de recherche de découverte pour rechercher des informations en fonction de certains critères, tels que les modèles, les types de données, des valeurs uniques et des valeurs Null. Vous pouvez utiliser les opérateurs conditionnels `=`, `>=` ou `<=` dans la recherche.

Le tableau suivant décrit les critères que vous pouvez utiliser pour la recherche de découverte :

Option	Description
Rechercher	Expression à rechercher.
Effacer	Efface la chaîne de recherche et tous les autres critères de recherche précédemment sélectionnés.
Modèle de	Modèle de colonne et pourcentage à inclure dans la recherche. Remarque: L'option n'accepte pas les caractères de contrôle dans un modèle.
Type de données de	Type de données des colonnes et pourcentage à inclure dans la recherche.
Valeurs uniques	Pourcentage de valeurs uniques dans les colonnes à inclure dans la recherche.
Valeurs Null	Pourcentage de valeurs Null dans les colonnes à inclure dans la recherche.

Recherche d'une ressource

Vous pouvez rechercher une ressource dans l'espace de travail **Bibliothèque**. Les résultats de la recherche incluent les ressources créées à la fois dans les outils Developer et Analyst.

1. Ouvrez l'espace de travail **Bibliothèque**.
2. Vérifiez que vous êtes dans la section **Recherche de découverte**.
3. Dans le champ **Recherche**, saisissez la chaîne à rechercher.
4. Configurez les filtres de recherche pour réduire votre recherche.
Les filtres incluent les modèles, les types de données, les valeurs uniques et les valeurs Null.
5. Cliquez sur l'icône **Rechercher**.

Résultats de la recherche de découverte dans Informatica Analyst

La recherche de découverte identifie des ressources dans tous les référentiels dotés d'une licence pour ce type de recherche, telles que le référentiel modèle et l'entrepôt de profilage.

Les résultats de la recherche de découverte incluent le nombre total de correspondances et la liste des correspondances. Vous pouvez développer chaque correspondance pour afficher ses propriétés, les informations de correspondance directe, les informations de correspondance indirecte et le nombre total de ressources liées, le cas échéant. Une correspondance directe est une correspondance avec une partie ou l'ensemble des métadonnées de la ressource qui répond aux critères de recherche. Une correspondance indirecte est une correspondance liée à la ressource qui répond directement aux critères de recherche.

L'ordre des résultats de recherche dépend des facteurs suivants :

- Propriété de l'objet qui répond aux critères de recherche. Le nom de l'objet est prioritaire sur sa description. La description de l'objet est prioritaire sur d'autres propriétés.
- Type d'objet. Les domaines et groupes de domaines de données n'ont pas priorité sur d'autres objets.
- Conservation. Les résultats de profil conservés sont prioritaires sur les résultats de profil non conservés.
- Nombre de correspondances des objets aux critères de recherche, correspondances directes et indirectes incluses.
- Fréquence relative du mot-clé.

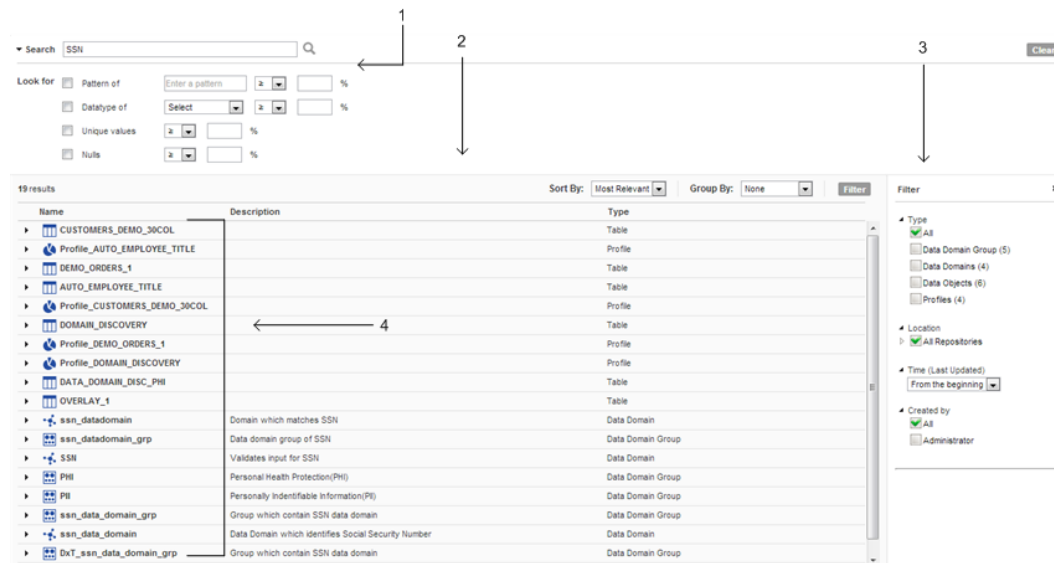
Pour afficher les résultats de la recherche, vous devez disposer des autorisations appropriées sur les projets qui contiennent les correspondances directes et indirectes.

Panneau de résultats de la recherche de découverte

Les résultats de la recherche de découverte incluent le nom, le type et la description de la ressource. Utilisez des filtres pour réduire les résultats.

Les résultats de la recherche s'affichent dans la grille de résultats de l'espace de travail de **bibliothèque**. Vous pouvez trier les résultats par pertinence. Vous pouvez également regrouper les résultats en fonction du type de ressource, de l'emplacement du référentiel, de l'heure et de l'utilisateur qui a créé les ressources.

La figure suivante présente l'interface des résultats de la recherche de découverte :



1. Critères de recherche
2. Grille de résultats
3. Filtre
4. Résultats de la recherche

Le panneau de résultats de la recherche de découverte affiche par défaut les sections suivantes :

Critères de recherche

Affiche les champs de recherche et notamment les filtres de recherche qui dépendent des métadonnées de profil que vous pouvez définir pour affiner la recherche. Les champs de recherche s'affichent dans la partie supérieure de l'outil Analyst.

Grille de résultats

Affiche le nombre total de correspondances et la liste de correspondances selon les critères sélectionnés dans les champs de critères de recherche. La grille contient également la description et le type de l'objet, ainsi que des listes déroulantes permettant de trier et de regrouper les résultats.

Filtre

Affiche les filtres que vous pouvez définir pour affiner la recherche. La section **Filtre** s'affiche sur la droite de l'outil Analyst.

Résultats de la recherche

Affiche les résultats correspondant à la chaîne de recherche et notamment les objets qui répondent aux critères de recherche. La section de résultats de la recherche contient les propriétés de correspondance et les sections **Correspondances directes** et **Correspondances indirectes** qui s'affichent lorsque vous développez une correspondance. Vous pouvez également afficher le nombre total de ressources liées à la correspondance dans la grille de résultats.

Filtrage des résultats de la recherche de découverte

Vous pouvez filtrer les résultats de la recherche en fonction du type de ressource, de l'emplacement du référentiel, de l'heure et de l'utilisateur ayant créé la ressource. Si le glossaire métier est installé, vous pouvez

également utiliser des filtres spécifiques des ressources pour les termes métier, les catégories et les stratégies du glossaire métier.

1. Effectuez la recherche globale ou la recherche de découverte dans la section **Recherche de découverte** de l'espace de travail **Bibliothèque**.
2. Cliquez sur **Filtre** dans la grille de résultats pour ouvrir la section **Filtre**.
3. Dans la section **Filtre**, sélectionnez les filtres requis et les paramètres associés.
4. Les résultats de la recherche révisés s'affichent dans la grille de résultats en fonction des paramètres de filtre que vous avez sélectionnés.
5. Pour effacer tous les paramètres de filtre, cliquez sur **Tout effacer** en haut de la grille de résultats.

Types de correspondances

Les résultats de la recherche de découverte incluent les correspondances directes et indirectes. Une correspondance directe est une correspondance avec tout ou partie des métadonnées de la ressource qui correspond à la requête de recherche. Une correspondance indirecte est une correspondance de ressource liée à une autre ressource qui correspond directement à la requête de recherche.

Si une requête de recherche contient plusieurs critères de recherche, les résultats de la recherche peuvent répondre aux critères de recherche directement, indirectement, ou les deux. Vous pouvez ouvrir des correspondances directes et indirectes à partir des résultats de la recherche en lecture seule ou en mode de modification.

Correspondance directe

Une correspondance directe est une correspondance avec tout ou partie des métadonnées de la ressource qui correspond à la requête de recherche. Par exemple, si vous recherchez toutes les ressources avec le nom « Client », l'outil Analyst peut répertorier les objets de données et les profils avec le nom « Client » en tant que correspondances directes. Une fois que vous avez effectué une recherche de découverte, la liste de correspondance que vous voyez contient des liens vers certains des objets.

Vous pouvez développer une ressource dans les résultats de la recherche pour afficher davantage d'informations sur la correspondance directe, par exemple les propriétés de la ressource.

Correspondance indirecte

Une correspondance indirecte est une correspondance liée à la correspondance directe. Par exemple, une fiche d'évaluation utilise une règle, qui contient le mot-clé recherché. La recherche de découverte renvoie la règle comme une correspondance directe et la fiche d'évaluation comme une correspondance indirecte. La fiche d'évaluation est une correspondance indirecte, car elle fait référence à la règle.

Utilisez les informations de correspondance indirecte pour identifier les relations masquées entre les objets et mieux comprendre les relations d'objets. Les résultats de correspondance indirecte permettent également de comprendre pourquoi la recherche de découverte renvoie un objet.

Affichage des informations de correspondance

Après avoir effectué une recherche de découverte, vous pouvez afficher les informations de correspondance, y compris les correspondances directes et indirectes. Vous pouvez également afficher les propriétés de la

ressource, telles son type, sa description et les ressources liées. Vous pouvez ouvrir des ressources à partir des résultats de la recherche et les modifier, si nécessaire.

1. Effectuez la recherche globale ou la recherche de découverte dans la section **Recherche** de l'espace de travail **Bibliothèque**.
2. Dans la grille de résultats, cliquez sur l'icône « Développer » au début d'un nom de ressource.
Les propriétés de la ressource et les informations de correspondance s'affichent dans une section sous le nom de la ressource.
3. Vérifiez les informations de correspondance directe et indirecte.
Vous pouvez voir les relations de ressource et d'autres informations comme le nombre total de ressources liées. Les relations de ressource comprennent les correspondances directes et indirectes.
4. Si les informations de ressource contiennent des liens hypertexte, cliquez sur ces liens pour ouvrir les ressources dans un autre espace de travail.
5. Cliquez sur l'icône « Développer » pour fermer la section des informations de correspondance.

Ouverture des ressources à partir des résultats de la recherche de découverte

Vous devez disposer des autorisations requises sur le projet, les ressources et la licence pour afficher les ressources à partir des résultats de la recherche.

1. Effectuez la recherche globale ou la recherche de découverte dans la section **Recherche** de l'espace de travail **Bibliothèque**.
2. Dans la grille de résultats, cliquez avec le bouton droit de la souris sur un nom de ressource.
Un menu contextuel s'affiche à l'écran.
3. Pour afficher la ressource en lecture seule dans son espace de travail, sélectionnez **Ouvrir**.
4. Pour modifier la ressource dans son espace de travail, sélectionnez **Modifier**.
5. Pour supprimer la ressource à partir des résultats de la recherche, sélectionnez **Supprimer**.
Lorsque vous supprimez une ressource dans les résultats de la recherche, l'outil Analyst la supprime du référentiel modèle.
6. Pour revenir à l'espace de travail **Bibliothèque**, cliquez sur **Bibliothèque**.

Ressources liées

Vous pouvez afficher les ressources liées à une ressource dans les résultats de la recherche. Une ressource liée est une ressource du référentiel modèle ou du glossaire métier qui est associée à une ressource sélectionnée dans les résultats de la recherche. La ressource liée partage certaines métadonnées avec la ressource des résultats de la recherche. Les ressources liées d'une source de données peuvent être des profils, des domaines de données inférés et des mappages.

Par exemple, un profil peut faire partie des résultats de la recherche. Vous pouvez afficher les ressources liées au profil, telles que les règles et les sources de données. Vous pouvez afficher les ressources liées dans l'espace de travail **Ressources liées**. Les ressources liées affichées diffèrent selon le type de ressource. Par exemple, lorsque vous affichez les ressources liées à une règle, vous pouvez afficher des ressources telles que le terme métier, la spécification de mappage et le profil.

Ressources liées pour chaque type de ressource

Les ressources liées affichées pour une ressource dans l'outil Analyst dépendent du type de ressource que vous recherchez.

Le tableau suivant décrit les ressources liées pour chaque ressource :

Type de ressource	Ressources liées
Terme métier	Domaine de données, objet de données, mapplet et règle.
Domaine de données	Terme métier, groupe de domaines de données, objet de données et profil.
Groupe de domaines de données	Domaine de données, objet de données et profil.
Objet de données	Terme métier, domaine de données, groupe de domaines de données, spécification de mappage, profil, fiche d'évaluation, mappage et mapplet.
Profil de découverte des données d'entreprise	Objet de données et profil.
Mappage Remarque: Vous ouvrez cet objet dans l'outil Developer.	Objet de données, spécification de mappage, mapplet et règle.
Spécification de mappage	Objet de données, profil, fiche d'évaluation, mappage, mapplet et règle.
Mapplet Remarque: Vous ouvrez cet objet dans l'outil Developer.	Terme métier, objet de données, spécification de mappage, mappage, mapplet et règle.
Profil	Domaine de données, groupe de domaines de données, objet de données, règle et spécification de mappage. Remarque: Une fiche d'évaluation n'est pas incluse dans les ressources liées pour un profil.
Règle	Objet de données, règle, terme métier, spécification de mappage, profil, fiche d'évaluation et mappage.

Affichage des ressources liées

Vous pouvez afficher le nombre total de ressources liées lorsque vous affichez les informations de correspondance de ressource dans les résultats de la recherche.

1. Effectuez une recherche de découverte dans l'espace de travail **Bibliothèque**.
2. Dans la grille de résultats, cliquez sur l'icône de développement, puis sur le lien du nombre de ressources liées ou cliquez avec le bouton droit et sélectionnez **Afficher les ressources liées**.

La liste de toutes les ressources liées s'affiche dans l'espace de travail **Ressources liées**.

3. Pour afficher les détails d'une ressource, cliquez sur son nom ou cliquez avec le bouton droit et sélectionnez **Ouvrir**.
4. Pour afficher les ressources liées d'une ressource liée, cliquez avec le bouton droit de la souris sur le nom de la ressource et sélectionnez **Afficher les ressources liées**.

Les informations relatives à la ressource liée s'affichent dans l'espace de travail.

5. Pour naviguer entre plusieurs espaces de travail de ressources liées, dans l'espace de travail **Ressources liées**, sélectionnez l'une des ressources récemment ouvertes.

Forum Aux Questions (FAQ)

Pourquoi certains résultats de recherche attendus ne s'affichent-ils pas ?

Les résultats de recherche peuvent ne pas s'afficher pour diverses raisons. Vérifiez que les critères de recherche respectent les instructions suivantes :

- Les ressources qui s'affichent dans les résultats de recherche dépendent des autorisations de projets.
- Les résultats de la recherche de découverte n'incluent pas de fréquences de valeur des résultats de profil de colonne.
- Les résultats de recherche n'incluent pas les résultats de profil rejetés lors de la conservation des résultats de profil.
- Les résultats de recherche affichés dépendent de l'intervalle d'extraction de l'index de recherche et de la disponibilité des ressources dans l'index de recherche.

Puis-je enregistrer les résultats de la recherche de découverte en vue d'une utilisation ultérieure ou les partager avec un autre utilisateur ?

Non. Vous ne pouvez pas enregistrer ou partager les résultats de la recherche de découverte.

Pourquoi certains résultats de la recherche de découverte s'affichent-ils en haut de la liste et d'autres en bas ?

L'ordre dans lequel l'outil Analyst affiche les résultats de la recherche dépend de plusieurs facteurs : type d'objet, résultats de profil conservés, propriété de l'objet répondant principalement aux critères de recherche, rang de recherche interne de chaque objet, etc.

Puis-je exporter les résultats de la recherche de découverte ?

Non. Vous ne pouvez pas exporter les résultats de la recherche.

CHAPITRE 14

Bureau de glossaire métier dans Informatica Analyst

Ce chapitre comprend les rubriques suivantes :

- [Termes métier, 106](#)
- [Gestion des termes métier dans le glossaire métier du gestionnaire de métadonnées, 107](#)
- [Recherche d'un terme métier dans le Bureau de glossaire métier, 107](#)

Termes métier

Vous pouvez rechercher des termes métier dans le Bureau de glossaire métier. Vous pouvez afficher des termes métier et effectuer des tâches associées en fonction de la licence du gestionnaire de métadonnées.

Un glossaire métier est un ensemble de termes qui utilisent un langage métier spécifique pour définir des concepts pour les utilisateurs métier. Un terme métier fournit la définition et l'utilisation métier d'un concept.

Le Bureau de glossaire métier est un client qui se connecte au Service du gestionnaire de métadonnées qui héberge le glossaire métier. Le Bureau de glossaire métier doit être ouvert pour que vous puissiez rechercher un nom d'objet de l'outil Analyst. Vous pouvez rechercher la signification d'un nom d'objet de l'outil Analyst en tant que terme métier dans le Bureau de glossaire métier pour comprendre son exigence métier et son implémentation en cours.

Le gestionnaire de métadonnées héberge des glossaires métier. Vous devez associer un Service du gestionnaire de métadonnées au Service Analyst pour parcourir un glossaire métier du gestionnaire de métadonnées à partir de l'outil Analyst. Vous pouvez afficher les termes métier dans un glossaire métier ou en les regroupant par catégorie. Vous pouvez éditer les termes métier Metadata Manager.

Vous pouvez rechercher des objets Metadata Manager par terme métier Metadata Manager dans le référentiel Metadata Manager. Vous pouvez sélectionner les objets du gestionnaire de métadonnées dans les résultats de la recherche et les importer en tant qu'objets de données dans l'outil Analyst. Vous ne pouvez pas ajouter un terme métier Metadata Manager au glossaire métier Metadata Manager.

Gestion des termes métier dans le glossaire métier du gestionnaire de métadonnées

Vous pouvez accéder au glossaire métier Metadata Manager à partir de l'outil Analyst pour gérer les termes métier Metadata Manager.

1. Dans l'en-tête de l'outil Analyst, cliquez sur **Gérer > Gérer les termes**.
Metadata Manager et le glossaire métier Metadata Manager s'ouvrent dans un onglet distinct. Les termes métier Metadata Manager s'affichent dans la vue **Glossaire** de Metadata Manager.
2. Pour choisir un glossaire métier, sélectionnez-le dans la liste Afficher.
3. Pour afficher les termes métier par catégorie, cliquez sur **Actions > Afficher > Catégories**.
4. Pour afficher tous les termes d'un glossaire métier dans l'ordre alphabétique, cliquez sur **Actions > Afficher > Alphabet**.
5. Pour afficher les termes métier commençant par une lettre spécifique, cliquez sur la lettre.
6. Pour éditer un terme métier, sélectionnez le terme métier et cliquez sur **Actions > Éditer les propriétés**.

Recherche d'un terme métier dans le Bureau de glossaire métier

Recherchez un nom d'objet de l'outil Analyst dans le Bureau de glossaire métier en tant que terme métier pour comprendre son exigence métier et son implémentation en cours.

Le Bureau de glossaire métier doit être installé sur votre machine.

1. Mettez le nom d'un objet en surbrillance.
2. Utilisez la combinaison de touches de raccourci pour rechercher le nom de l'objet en tant que terme métier dans le Bureau de glossaire métier.

La combinaison de touches de raccourci par défaut est **MAJ + ALT + Q**.

Partie III : Découverte de données avec Informatica Developer

Cette partie contient les chapitres suivants :

- [Profils de Informatica Developer, 109](#)
- [Profils d'objet de données, 112](#)
- [Résultats de profil de colonne dans Informatica Developer, 121](#)
- [Règles dans Informatica Developer, 127](#)
- [Fiches d'évaluation dans Informatica Developer, 129](#)
- [Profilage de mapplet et de mappage, 132](#)
- [Découverte des domaines de données dans Informatica Developer, 135](#)
- [Découverte des données d'entreprise dans Informatica Developer, 147](#)
- [Résultats de la découverte des données d'entreprise, 161](#)
- [Bureau de glossaire métier dans Informatica Developer, 172](#)

CHAPITRE 15

Profils de Informatica Developer

Ce chapitre comprend les rubriques suivantes :

- [Présentation des profils de Informatica Developer, 109](#)
- [Vues des profils d'Informatica DeveloperVues des profils, 111](#)

Présentation des profils de Informatica Developer

Créez et exécutez des profils dans Informatica Developer pour découvrir les problèmes de qualité des données dans un ensemble de données et pour comprendre les relations entre les colonnes dans un ensemble de données.

Vous pouvez créer des profils pour les types suivants d'analyse de données :

- Profilage de colonne
- Découverte de clé principale
- Découverte de dépendance fonctionnelle
- Découverte de clé étrangère
- Analyse de jointure
- Découverte du chevauchement
- Découverte de domaines de données
- Découverte des données d'entreprise

Vous devez créer des profils dans l'outil Developer à l'aide d'un assistant. L'assistant Création de profils propose les options **Profil**, **Profils multiples** et **Profil de découverte des données d'entreprise** pour créer des profils.

Profil

Créez un profil pour un seul objet de données. Pour un seul profil, vous devez définir des filtres, des règles et des options de développement pour le profilage de colonnes. Vous pouvez également sélectionner des options avancées pour créer un profil de colonne, un profil de clé principale, un profil de dépendance fonctionnelle, et pour la découverte des domaines de données. Les résultats affichent le profil de colonne, l'inférence de clé primaire, la dépendance fonctionnelle et l'inférence de domaine de données.

Profils multiples

Créez un ensemble de profils pour plusieurs objets. L'outil Developer crée un profil pour chaque objet et exécute les profils simultanément. Quand vous créez plusieurs profils à la fois, vous ne pouvez pas analyser les données dans les objets.

Profil de découverte des données d'entreprise

Générez un modèle de données à partir de plusieurs objets de données et créez un profil qui analyse les données dans les objets. Créez un profil de découverte des données d'entreprise et ajoutez-lui des objets de données physiques à profiler ensemble. Vous pouvez créer un profil d'objet de données, un profil de clé étrangère et un profil de jointure. Pour chaque objet de données du profil de découverte des données d'entreprise, vous pouvez configurer des propriétés générales, des colonnes à profiler, des clés et des relations. Vous pouvez découvrir des données de chevauchement dans une source de données ou dans des sources de données multiples.

Vous pouvez également exécuter une découverte des données d'entreprise qui crée et exécute des tâches de découverte de données, telles que le profil de colonne, la découverte de domaines de données, le profil de clé primaire et le profil de clé étrangère. La découverte des données d'entreprise s'exécute sur un grand nombre de sources de données sur plusieurs connexions.

La table suivante indique les opérations que vous pouvez effectuer avec chaque type de profil :

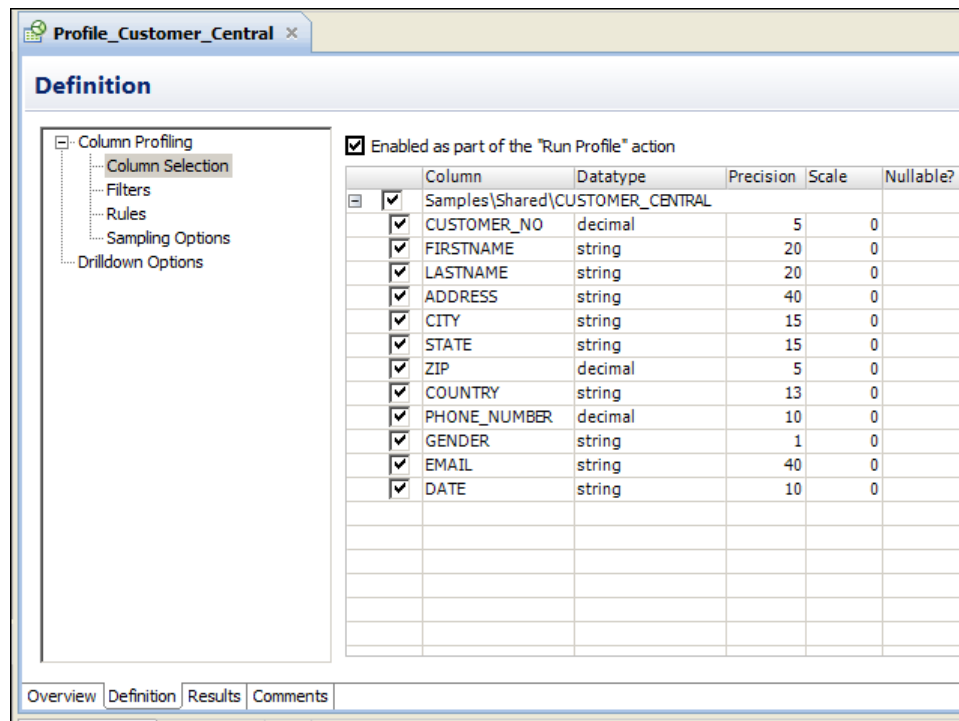
Option de profilage	Opérations de profilage
Profil	<ul style="list-style-type: none">- Exécuter un profil de colonne dans un seul ensemble de données.- Rechercher des clés principales.- Rechercher des dépendances fonctionnelles.- Identifier des domaines de données
Profils multiples	Créer et exécuter des profils de colonne sur plusieurs objets en même temps.
Profil de découverte des données d'entreprise	<ul style="list-style-type: none">- Exécuter un profil de colonne dans un seul ensemble de données.- Rechercher des clés principales.- Rechercher des clés étrangères.- Rechercher des dépendances fonctionnelles.- Effectuer une analyse de jointure.- Découvrir le chevauchement entre deux colonnes.- Exécuter la découverte des données d'entreprise.

Vues des profils d'Informatica Developer

Vues des profils

Vous pouvez afficher et ajouter des informations sur un profil dans Informatica Developer à l'aide des vues **Présentation**, **Définition**, **Commentaires** et **Résultats**.

La figure suivante montre les vues des profils dans l'éditeur :



Lorsque vous ouvrez un profil dans la vue **Explorateur d'objets**, l'éditeur dans le volet droit affiche les informations du profil dans les vues suivantes :

Présentation

Affichez et indiquez des informations générales sur le profil, tels que le nom, la description et l'emplacement.

Définition

Affichez et configurez la définition du profil.

Ces informations comprennent la liste de filtres et de règles que vous affectez au profil, les options de développement et les fonctions de profil activées lors de l'exécution du profil.

Ces informations comprennent la liste de filtres et de règles affectés au profil.

Résultats

Affiche les résultats de l'exécution du profil. Vous pouvez exporter les résultats après avoir exécuté un profil.

Commentaires

Affichez et ajoutez des commentaires au profil.

CHAPITRE 16

Profils d'objet de données

Ce chapitre comprend les rubriques suivantes :

- [Présentation de profils d'objet de données, 112](#)
- [Profils de colonne dans Informatica Developer, 113](#)
- [Découverte de clé primaire, 114](#)
- [Découverte de dépendance fonctionnelle, 116](#)
- [Création d'un profil d'objet de données unique](#)[Création d'un profil, 118](#)
- [Création de plusieurs profils d'objets de données, 119](#)
- [Synchronisation d'un objet de données fichier plat, 119](#)
- [Synchronisation d'un objet de données relationnel, 119](#)

Présentation de profils d'objet de données

Un profil d'objet de données découvre les informations sur les données et les métadonnées de colonne dans une source de données. Vous pouvez exécuter un profil dans un seul objet de données ou plusieurs objets de données dans l'outil Developer. Un profil d'objet de données unique analyse une seule source de données. Des profils d'objets à plusieurs données analysent plusieurs sources de données. Lorsque vous créez plusieurs profils d'objet de données, vous pouvez exécuter un profil de colonne sur ces profils.

Le tableau suivant décrit les tâches de découverte de données que vous pouvez effectuer pour un profil d'objet de données unique :

Tâche	Description
Profilage de colonne	Découvre les caractéristiques de données, telles que les fréquences, pourcentages et formes. Vous pouvez ajouter des filtres pour déterminer les lignes que le profil lit lors de l'exécution. Le profil ne traite pas les lignes qui ne répondent pas aux critères de filtrage.
Découverte de clé principale	Découvre les colonnes avec des valeurs qui peuvent identifier de façon unique les lignes dans une source de données.
Découverte de dépendance fonctionnelle	Découvre les dépendances entre les paires de colonnes dans une source de données.
Découverte de domaines de données	Identifie tous les domaines de données d'une colonne en fonction de sa valeur ou de son nom.

Le tableau suivant décrit les tâches de découverte de données que vous pouvez effectuer sur plusieurs objets de données lorsque vous créez un modèle de données au moyen de l'option **Profil de découverte des données d'entreprise** :

Tâche	Description
Découverte de clé étrangère	Découvre les colonnes qui comportent des valeurs qui correspondent aux valeurs de clé principale dans une autre source de données.
Analyse de jointure	Découvre le degré de jointures potentielles entre les données dans deux colonnes d'une source de données ou entre deux sources de données.
Découverte du chevauchement	Découvre le pourcentage de chevauchement des données entre des paires de colonnes d'une source de données ou de plusieurs sources de données.
Découverte des données d'entreprise	Découvre les statistiques de profil de colonne, les domaines de colonne, les clés principales et les clés étrangères dans un grand nombre de sources de données réparties sur plusieurs connexions ou schémas.

Profils de colonne dans Informatica Developer

Un profil de colonne permet d'analyser les caractéristiques des colonnes dans un ensemble de données, telles que les pourcentages et les modèles de valeurs. Vous pouvez ajouter des filtres pour déterminer les lignes que le profil lit lors de l'exécution. Le profil ne traite pas les lignes qui ne répondent pas aux critères de filtrage.

Vous pouvez détecter les types suivants d'informations concernant les colonnes sur lesquelles vous exécutez un profil :

- Le nombre de fois où une valeur s'affiche dans une colonne.
- La fréquence d'exécution de chaque valeur dans une colonne, exprimée en pourcentage.
- Les modèles de caractères des valeurs d'une colonne.
- Des statistiques, telles que la longueur maximale et la longueur minimale des valeurs dans une colonne, ainsi que la première valeur et la dernière.
- Les types de données inférés, la fréquence, le pourcentage de conformité et le statut de l'inférence du type de données.

Vous pouvez définir un profil de colonne pour un objet de données dans un mappage ou une mapplet, ou encore un objet dans le référentiel Modèle. L'objet dans le référentiel peut être dans un profil d'objet de données unique, plusieurs profils d'objet de données ou un profil de découverte des données d'entreprise.

Vous pouvez ajouter des règles à un profil de colonne. Utilisez les règles pour définir la logique métier que vous pouvez appliquer aux données source. Vous pouvez également modifier les options de développement pour les profils de colonne afin de déterminer si la tâche de développement lit à partir des données stockées ou en direct.

Options de filtrage

Vous pouvez ajouter des filtres pour déterminer les lignes qu'un profil de colonne utilise lors de l'exécution des opérations de profilage. Le profil ne traite pas les lignes qui ne répondent pas aux critères de filtrage.

1. Créez ou ouvrez un profil de colonne.

2. Sélectionnez la vue **Filtre**.
3. Cliquez sur **Ajouter**.
4. Sélectionnez un type de filtre et cliquez sur **Suivant**.
5. Entrez un nom pour le filtre. Entrez éventuellement une description textuelle du filtre.
6. Sélectionnez **Définir comme actif** pour appliquer le filtre au profil. Cliquez sur **Suivant**.
7. Définissez les critères de filtrage.
8. Cliquez sur **Terminer**.

Propriétés d'échantillonnage

Configurez les propriétés d'échantillonnage pour déterminer le nombre de lignes que le profil lit lors d'une opération de profilage.

La table suivante décrit les propriétés d'échantillonnage.

Propriété	Description
Toutes les lignes	Lit toutes les lignes à partir de la source. La valeur par défaut est activée.
Premier	Lit depuis la première ligne jusqu'à la ligne que vous indiquez.
Échantillon aléatoire de	Lit un échantillon aléatoire à partir du nombre de lignes que vous indiquez.
Échantillon aléatoire (Auto)	Lit à partir d'un échantillon aléatoire de lignes.
Exclure l'inférence du type de données pour les colonnes avec un type de données approuvé	Exclut les colonnes qui comprennent un type de données approuvé de l'inférence du type de données de l'exécution du profil.

Propriété	Description
Toutes les lignes	Lit toutes les lignes à partir de la source. La valeur par défaut est activée.
Premier	Lit depuis la première ligne jusqu'à la ligne que vous indiquez.

Découverte de clé primaire

La découverte de clé primaire génère des candidats de clés primaires provenant des colonnes que vous indiquez.

Une clé primaire est une colonne ou une combinaison de colonnes qui identifie de façon unique une ligne dans une source de données. La découverte de clé primaire identifie les colonnes et les combinaisons de colonnes qui répondent à un niveau de confiance spécifique. Vous pouvez éditer le niveau de confiance, ainsi que le nombre maximum de colonnes à combiner pour l'identification de clé primaire.

La découverte de clé primaire peut souligner des problèmes potentiels de qualité de données en identifiant les lignes non uniques dans un candidat de clé primaire. Ceci est particulièrement utile dans les cas où la découverte de clé primaire combine plusieurs colonnes, étant donné que des enregistrements non conformes sont susceptibles de contenir des informations dupliquées.

Propriétés d'inférence de clés primaires

Quand vous créez un profil d'objet de données unique, vous pouvez utiliser la vue **Profilage de clé primaire** pour configurer les propriétés d'inférence de clés primaires.

La table suivante décrit les propriétés d'inférence de clés primaires dans la vue **Profilage de clé primaire** :

Propriété	Description
Remplacer les options d'inférence par défaut	Permet de configurer les paramètres personnalisés pour l'inférence de clé principale.
Nombre max. de clés	Nombre maximum de colonnes qui peuvent former une clé principale.
Nombre max. de lignes	Le nombre de lignes à profiler.
Critères de conformité	Le pourcentage minimum ou le nombre maximum de lignes pour les violations de clés que le profil autorise lors de la définition des clés primaires.
Exclure les objets de données avec une clé documentée, définie par l'utilisateur	Exclut les objets de données avec des clés primaires documentées ou définies par l'utilisateur.
Exclure les objets de données avec une clé approuvée	Exclut les objets de données avec des clés primaires approuvées.

Propriétés de clés primaires inférées

Après avoir exécuté un profil d'objet à donnée unique, vous pouvez utiliser la vue **Profilage de clé primaire** pour afficher les détails des clés primaires inférées dans la source de données.

La table suivante décrit les propriétés de clés primaires inférées dans la vue **Profilage de clé primaire** :

Propriété	Description
Colonne	Le nom de la colonne dans le profil.
% de conformité	Pourcentage de valeurs uniques dans la colonne.
% de doublons	Le pourcentage de valeurs dupliquées pour la colonne.
% null	Pourcentage de valeurs null pour la colonne.
Vérifié	Détermine si la colonne est une colonne de clé primaire.
Statut de l'inférence	Statut d'inférence de la colonne.
Heure de la dernière exécution	La date et l'heure de la dernière exécution du profil de clé primaire.

Propriétés de violations de clés

Après avoir exécuté un profil d'objet à donnée unique, vous pouvez utiliser la vue **Profilage de clé primaire** pour afficher les détails des violations de clés primaires dans la source de données.

La table suivante décrit les propriétés des violations de clés dans la vue **Profilage de clé primaire** :

Propriété	Description
Colonne(s)	Le nom de la ou des colonnes d'où le profil infère une clé primaire candidate.
Nombre de violations de clés	Le nombre de violations de clés dans le candidat de clé primaire.

Découverte de dépendance fonctionnelle

La découverte de dépendance fonctionnelle fournit des informations sur les dépendances entre les paires de colonnes dans une source de données.

Deux colonnes sont fonctionnellement dépendantes si les valeurs d'une colonne peuvent prédire de manière fiable les valeurs dans une autre colonne. Par exemple, si un ensemble de données contient une colonne Identifiant d'employé et une colonne Date de naissance, la date de naissance doit être identique dans toutes les lignes qui contiennent un identifiant d'employé donné.

Les dépendances fonctionnelles peuvent identifier les problèmes potentiels de qualité des données en identifiant les enregistrements non conformes à une dépendance fonctionnelle de colonne. Par exemple, si 99,8 % de lignes dans une source de données sont fonctionnellement dépendantes, il est fort probable que les lignes restantes contiennent des informations inexactes.

Propriétés d'inférence de dépendances fonctionnelles

La vue **Profilage de dépendance fonctionnelle** fournit des informations sur les dépendances fonctionnelles entre les colonnes.

La table suivante décrit les propriétés d'inférence de dépendances fonctionnelles dans la vue **Profilage de dépendance fonctionnelle** :

Propriété	Description
Remplacer les options d'inférence par défaut	Permet de configurer les paramètres personnalisés pour l'inférence de dépendance fonctionnelle.
Nombre max. de colonnes dans le déterminant	Le nombre de colonnes que le profil peut combiner pour trouver un déterminant.
Nombre max. de lignes	Le nombre de lignes à profiler.
Dépendances renvoyées	Le nombre de dépendances que le profil affiche. La valeur par défaut est Couverture minimum , qui affiche le plus petit ensemble de dépendances où chaque colonne apparaît au moins une fois dans une dépendance ou un déterminant.

Propriété	Description
Nombre max. de dépendances renvoyées	Le nombre maximum de dépendances que le profil affiche.
Critères de conformité	Le pourcentage minimum ou le nombre maximum de lignes pour les violations de dépendances que le profil autorise lors de la définition des dépendances fonctionnelles.

Propriétés des dépendances fonctionnelles inférées

Après avoir exécuté un profil d'objet à donnée unique, vous pouvez utiliser la vue **Inférence de dépendance fonctionnelle** pour afficher les détails des dépendances fonctionnelles inférées dans la source de données.

La table suivante décrit les propriétés des dépendances fonctionnelles inférées dans la vue **Inférence de dépendance fonctionnelle** :

Propriété	Description
Colonnes déterminantes	Nom de la colonne analysée pour les dépendances fonctionnelles.
Colonnes dépendantes	Nom de la colonne dépendante de la colonne déterminante.
% Null	Pourcentage de valeurs null pour la colonne.
% de conformité	Pourcentage de correspondance de dépendance fonctionnelle.
Vérifié	Détermine si oui ou non les colonnes sont fonctionnellement dépendantes.
Heure de la dernière exécution	La date et heure de la dernière exécution du profil de dépendance fonctionnelle.

Propriétés de violations de dépendances fonctionnelles

La vue contient des informations sur les dépendances fonctionnelles entre les colonnes. Après avoir exécuté un profil d'objet à donnée unique, vous pouvez utiliser la vue **Inférence de dépendance fonctionnelle** pour afficher les détails des violations de dépendances fonctionnelles dans la source de données.

La table suivante décrit les propriétés des violations de dépendances fonctionnelles dans la vue **Profilage de dépendance fonctionnelle** :

Propriété	Description
Colonne déterminante	Nom de la colonne analysée pour les dépendances fonctionnelles.
Dépendances distinctes	Le nombre de dépendances fonctionnelles uniques.

Création d'un profil d'objet de données unique

Création d'un profil

Vous pouvez créer un profil d'objet de données unique pour une ou plusieurs colonnes dans un objet de données et stocker l'objet de profil dans le référentiel modèle. Vous pouvez créer un profil pour une ou plusieurs colonnes dans un objet de données et stocker l'objet de profil dans le référentiel modèle.

1. Dans la vue **Explorateur d'objets**, sélectionnez l'objet de données à profiler.
 2. Cliquez sur **Fichier > Nouveau > Profil** pour ouvrir l'assistant Création de profil.
 3. Sélectionnez **Profil** et cliquez sur **Suivant**.
 4. Entrez un nom pour le profil et vérifiez l'emplacement du projet. Si nécessaire, accédez à un nouvel emplacement.
 5. Entrez éventuellement une description textuelle du profil.
 6. Vérifiez que le nom de l'objet de données que vous avez sélectionné s'affiche dans la section **Objets de données**.
 7. Cliquez sur **Suivant**.
 8. Configurez les opérations de profilage à effectuer. Vous pouvez configurer les opérations suivantes :
 - Profilage de colonne
 - Découverte de clé principale
 - Découverte de dépendance fonctionnelle
 - Découverte de domaines de données
- Remarque:** Pour activer une opération de profilage, sélectionnez **Activée dans le cadre de l'action « Exécuter le profil »** pour cette opération. Le profilage de colonne est activé par défaut.
9. Configurez les options de profil de colonne telles que la sélection, le filtre, l'échantillonnage et l'exploration des colonnes.
 10. Vérifiez les options pour votre profil.

Vous pouvez modifier la sélection de colonne pour tous les types de profil. Vérifier les options de filtre et d'échantillonnage pour les profils de colonne. Vous pouvez vérifier les options d'inférence pour clé principale, la dépendance fonctionnelle et la découverte du domaine de données. Vous pouvez également vérifier la sélection de domaine de données pour la découverte de domaine de données.
 11. Vérifiez les options d'exploration et modifiez-les si nécessaire. L'option **Activer le zoom avant sur la ligne** est sélectionnée par défaut. Vous pouvez modifier les options d'exploration des profils de colonne. Les options déterminent si les opérations d'exploration lisent à partir de la source de données ou à partir des données intermédiaires, et si le profil stocke les données de résultat provenant des exécutions précédentes de profil.
 12. Dans la section **Paramètres d'exécution**, configurez les paramètres de validation et d'exécution. Vous pouvez sélectionner la connexion native ou Hive.
 13. Cliquez sur **Terminer**.

Création de plusieurs profils d'objets de données

Les opérations de profilage pour plusieurs objets de données utilisent des options de profilage de colonne par défaut pour générer des profils de colonne pour un ou plusieurs objets de données.

1. Dans la vue **Explorateur d'objets**, sélectionnez les objets de données à profiler.
2. Cliquez sur **Fichier > Nouveau > Profil** pour ouvrir l'assistant **Nouveau profil**.
3. Sélectionnez l'option **Profils multiples** et cliquez sur **Suivant**.
4. Sélectionnez l'emplacement dans lequel enregistrer les profils. Vous pouvez créer chaque profil aux mêmes emplacements que son objet profilé, ou vous pouvez indiquer un emplacement commun pour les profils.
5. Vérifiez que le nom des objets de données que vous avez sélectionné s'affiche dans la section **Objets de données**.
Cliquez éventuellement sur **Ajouter** pour ajouter un autre objet de données.
6. Indiquez éventuellement le nombre de lignes à profiler et choisissez si vous souhaitez exécuter le profil à la fin de l'exécution de l'assistant.
7. Cliquez sur **Terminer**.
8. Entrez éventuellement les chaînes de préfixe et de suffixe à ajouter aux noms de profils.
9. Cliquez sur **OK**.

Synchronisation d'un objet de données fichier plat

Vous pouvez synchroniser les modifications apportées à une source de données de fichier plat externe avec son objet de données dans Informatica Developer. Utilisez l'assistant **Synchronisation d'un fichier plat** pour synchroniser les objets de données.

1. Dans la vue **Explorateur d'objets**, sélectionnez un objet de données fichier plat.
2. Cliquez avec le bouton droit et sélectionnez **Synchroniser**.
L'assistant **Synchronisation d'un objet de données fichier plat** s'affiche.
3. Vérifiez le chemin d'accès au fichier plat dans le champ **Sélectionner le fichier plat existant**.
4. Cliquez sur **Suivant**.
5. Sélectionnez éventuellement les propriétés des éléments suivants : page de code, format, format délimité et colonne.
6. Cliquez sur **Terminer** puis sur **OK**.

Synchronisation d'un objet de données relationnel

Vous pouvez synchroniser les modifications d'une source de données externe avec son objet de données dans Informatica Developer. Les modifications d'une source de données externe incluent l'ajout, la modification et la suppression de colonnes, ainsi que les modifications apportées aux règles.

1. Dans la vue **Explorateur d'objets**, sélectionnez un objet de données relationnel.

2. Cliquez avec le bouton droit et sélectionnez **Synchroniser**.
Un message vous invite à confirmer l'action.
3. Pour valider le processus de synchronisation, cliquez sur **OK**. Cliquez sur **Annuler** pour annuler le processus.
Si vous cliquez sur **OK**, un message d'état de processus de synchronisation s'affiche.
4. Quand le message **Synchronisation terminée** s'affiche, cliquez sur **OK**.
Ce message affiche un résumé des modifications de métadonnées apportées à l'objet de données.

CHAPITRE 17

Résultats de profil de colonne dans Informatica Developer

Ce chapitre comprend les rubriques suivantes :

- [Résultats de profil de colonne dans Informatica Developer](#)[Résultats de profil de colonne, 121](#)
- [Propriétés des valeurs de colonne, 122](#)
- [Propriétés des formes de colonne, 123](#)
- [Propriétés des statistiques de colonne, 123](#)
- [Propriétés des types de données de colonne, 124](#)
- [Conservation dans l'outil Developer](#)[Conservation dans Informatica Developer, 125](#)
- [Exportation des résultats de profil depuis Informatica Developer, 126](#)

Résultats de profil de colonne dans Informatica Developer

Résultats de profil de colonne

L'analyse de profil de colonne fournit des informations sur la qualité des données en mettant en surbrillance les fréquences de valeur, les modèles et les statistiques des données.

La table suivante décrit les résultats de profil pour chaque type d'analyse :

L'analyse du profilage de colonnes génère les résultats de profil suivants :

Type de profil	Résultats de profil
Profil de colonne	<ul style="list-style-type: none"> - Pourcentage et statistiques de comptage pour les valeurs uniques et null - Types de données inférés - Type de données que la source de données déclare pour les données - Valeurs maximum et minimum - Date et heure d'exécution du profil le plus récent - Pourcentage et statistiques de comptage pour chaque élément de données unique dans une colonne - Pourcentage et statistiques de comptage pour chaque forme de caractère unique dans une colonne
Profil de clé primaire	<ul style="list-style-type: none"> - Clés primaires inférées - Violations de clés
Profil de dépendance fonctionnelle	<ul style="list-style-type: none"> - Dépendances fonctionnelles inférées - Violations de dépendances fonctionnelles

- Pourcentage et statistiques de comptage pour les valeurs uniques et null
- Types de données inférés
- Type de données que la source de données déclare pour les données
- Valeurs maximum et minimum
- Date et heure d'exécution du profil le plus récent
- Pourcentage et statistiques de comptage pour chaque élément de données unique dans une colonne
- Pourcentage et statistiques de comptage pour chaque forme de caractère unique dans une colonne

La figure suivante montre les résultats de profil de colonne :

Column Profiling										Details		
All 1934 rows. Last run on: Mar 25, 2013 7:24:28 PM										Show: Values		
Column	Unique Values	% Unique	Nulls	% Null	Datatype	Documented Datatype	Max Value	Min Value	Last Profiled	Value	Frequency	Percent
CUSTOMER_CENTRAL												
CUSTOMER_NO	1833	94.78	3	0.16	Integer(5) [100.00]	decimal(5)	99999	2	Mar 25, 2013 7:24:28 PM IST	99999	7	0.36%
FIRSTNAME	1282	66.29	1	0.05	String(14) [100.00]	string(20)	ZYLIA	A	Mar 25, 2013 7:24:28 PM IST	6661	4	0.21%
LASTNAME	973	50.31	-	-	String(15) [100.00]	string(20)	ZUCATI	ABAUNZA	Mar 25, 2013 7:24:28 PM IST	5716	4	0.21%
ADDRESS	931	48.14	-	-	String(30) [100.00]	string(40)	Y	1HIGH...	Mar 25, 2013 7:24:28 PM IST	6489	3	0.16%
CITY	8	0.41	-	-	String(12) [100.00]	string(15)	Minneap...	AnnArb...	Mar 25, 2013 7:24:28 PM IST	6263	3	0.16%
STATE	5	0.26	-	-	String(9) [100.00]	string(15)	Texas	Illinois	Mar 25, 2013 7:24:28 PM IST	6216	3	0.16%
ZIP	194	10.03	28	1.45	Integer(5) [100.00]	decimal(5)	98199	0	Mar 25, 2013 7:24:28 PM IST	6126	3	0.16%
COUNTRY	1	0.05	-	-	Fixed Length String(13)...	string(13)	United S...	United S...	Mar 25, 2013 7:24:28 PM IST	6100	3	0.16%
PHONE_NUMBER	1832	94.73	-	-	Integer(10) [100.00]	decimal(10)	9417575...	89	Mar 25, 2013 7:24:28 PM IST	6096	3	0.16%
GENDER	3	0.16	-	-	Fixed Length String(1)...	string(1)	U	F	Mar 25, 2013 7:24:28 PM IST	4587	3	0.16%
EMAIL	1664	86.04	118	6.1	String(27) [100.00]	string(40)	zxenia@...	aachess...	Mar 25, 2013 7:24:28 PM IST	3139	3	0.16%
DATE	1932	99.90	-	-	Date [100.00]	string(10)	9/9/1999	1/1/1995	Mar 25, 2013 7:24:28 PM IST	2422	3	0.16%
										NULL	3	0.16%
										729	2	0.10%

Propriétés des valeurs de colonne

Les propriétés des valeurs de colonne indiquent les valeurs dans les colonnes profilées et la fréquence d'affichage de chaque valeur dans chaque colonne. Les fréquences sont indiquées en pourcentage, en nombre et sous la forme d'un graphique à barres.

Pour afficher les propriétés de valeurs de colonne, sélectionnez Valeurs dans le menu **Affichage**. Double-cliquez sur une valeur de colonne pour développer les lignes qui contiennent la valeur.

La table suivante décrit les propriétés des valeurs de colonne :

Propriété	Description
Valeurs	Liste de toutes les valeurs de la colonne dans le profil.
Fréquence	Le nombre de fois où une valeur s'affiche dans une colonne.
Pourcentage	Le nombre de fois où une valeur s'affiche dans une colonne, exprimé en pourcentage de toutes les valeurs de la colonne.
Graphique	Le graphique à barres pour le pourcentage.

Propriétés des formes de colonne

Les propriétés des formes de colonne indiquent les formes de données dans les colonnes profilées et la fréquence d'affichage des formes dans chaque colonne. Les fréquences sont indiquées en nombre, en pourcentage et sous la forme d'un graphique à barres.

Pour afficher des informations sur les formes, sélectionnez **Modèles** dans la liste **Affichage**. Double-cliquez sur une forme pour développer les lignes qui contiennent la forme.

La table suivante décrit les propriétés des formes de valeurs de colonne :

Propriété	Description
Formes	La forme pour la colonne sélectionnée.
Fréquence	Le nombre de fois où une forme s'affiche dans une colonne.
Pourcentage	Le nombre de fois où une forme s'affiche dans une colonne, exprimé en pourcentage de toutes les valeurs de la colonne.
Graphique	Le graphique à barres pour le pourcentage.

Propriétés des statistiques de colonne

Les statistiques de colonne comprennent les propriétés, telles que les longueurs maximum et minimum des valeurs et les premières et dernières valeurs.

Pour afficher des informations statistiques, sélectionnez **Statistiques** dans la liste **Affichage**.

Le tableau suivant décrit les propriétés des statistiques de colonne :

Propriété	Description
Longueur maximale	La longueur de la valeur la plus longue dans la colonne.
longueur minimale	La longueur de la valeur la plus courte dans la colonne.
Inférieur	Les cinq dernières valeurs dans la colonne.
Supérieur	Les cinq premières valeurs dans la colonne.
Somme	Somme de toutes les valeurs dans la colonne avec un type de données numérique.

Remarque: Le profil indique également les statistiques de moyenne et d'écart-type pour les colonnes de type Nombre entier.

Propriétés des types de données de colonne

Les types de données de colonne incluent tous les types de données inférés de chaque colonne dans les résultats de profil.

Pour afficher les informations relatives au type de données, sélectionnez **Types de données** dans la liste **Afficher**. Double-cliquez sur un type de données pour développer les lignes qui contiennent ce type de données.

Le tableau suivant décrit les propriétés des types de données de colonne :

Propriété	Description
Type de données	Liste de tous les types de données inférés de la colonne dans le profil.
Fréquence	Nombre de fois où un type de données s'affiche pour une colonne, exprimé en nombre.
% de conformité	Pourcentage d'affichage d'un type de données pour une colonne.
Statut	<p>Indique le statut du type de données. Les statuts peuvent être Inféré, Approuvé ou Rejeté.</p> <p>Inféré</p> <p>Indique le type de données de la colonne que l'outil Developer a inféré.</p> <p>Approuvé</p> <p>Indique que le type de données de la colonne est approuvé. Lorsque vous approuvez un type de données, ce dernier est validé pour le référentiel modèle.</p> <p>Rejeté</p> <p>Indique que le type de données de la colonne est rejeté.</p>

Conservation dans l'outil Developer

Conservation dans Informatica Developer

La conservation est le processus de validation et de gestion des métadonnées d'une source de données découvertes, de sorte à les préparer afin d'être utilisées et de figurer dans des rapports. Lorsque vous conservez des métadonnées dans l'outil Developer, vous pouvez approuver, rejeter et rétablir les types de données ou domaines de données inférés dans les résultats de profil.

Vous pouvez approuver un type de données ou domaine de données par colonne et masquer les types de données ou domaines de données rejetés dans cette dernière. Après avoir approuvé ou rejeté un type de données ou domaine de données inféré, vous pouvez le rétablir afin de restaurer le statut inféré.

Approbation des types de données dans l'outil Developer

Approbation des types de données dans Informatica Developer

Les résultats de profil comprennent les types de données inférés, la fréquence, le pourcentage de conformité et le statut de l'inférence de chaque colonne dans la source de données. Vous ne pouvez choisir et approuver qu'un seul type de données pour chaque colonne dans l'outil Analyst.

1. Dans la vue **Explorateur d'objets**, sélectionnez et ouvrez un profil.
2. Vérifiez que vous êtes dans l'onglet **Résultats**.
3. Dans la vue **Profilage de colonnes**, sélectionnez une colonne pour en afficher les fréquences de valeur, les modèles, les types de données et les statistiques dans le panneau de droite.
4. Dans le panneau **Détails**, sélectionnez **Types de données** dans la liste **Afficher**.
Les types de données inférés de la colonne s'affichent.
5. Cliquez avec le bouton droit de la souris sur la colonne à approuver, puis cliquez sur **Approuver**.
Le type de données prend le statut **Approuvé**.
6. Pour restaurer le statut inféré du type de données, cliquez avec le bouton droit sur le type de données, puis cliquez sur **Réinitialiser**.

Rejet des types de données dans l'outil Developer

Rejet des types de données dans Informatica Developer

Lorsque vous affichez les résultats de profil, l'outil Developer affiche les types de données inférés par défaut. Vous pouvez rejeter des types de données inférés ou approuvés. Vous pouvez afficher ou masquer les types de données rejetés.

1. Dans la vue **Explorateur d'objets**, sélectionnez un profil.
2. Double-cliquez sur le profil pour l'ouvrir.
Le profil s'ouvre dans un onglet.
3. Sélectionnez une ligne dans la vue **Profilage de colonnes**.
4. Sélectionnez la vue **Types de données** dans le panneau de droite pour rejeter les types de données de colonnes inférés. Sélectionnez le type de données inféré que vous souhaitez rejeter, cliquez avec le bouton droit sur la ligne, puis sélectionnez **Rejeter**.

L'outil Developer grise le type de données rejeté dans la liste des types de données.

5. Pour masquer les types de données rejetés, cliquez avec le bouton droit sur la ligne, puis sélectionnez **Masquer les éléments rejetés**.
6. Pour afficher les types de données rejetés, cliquez avec le bouton droit sur l'une des lignes, puis sélectionnez **Afficher les éléments rejetés**.

Exportation des résultats de profil depuis Informatica Developer

Vous pouvez exporter les résultats de profil de colonne dans un fichier .csv ou un fichier Microsoft Excel. Lorsque vous exportez les résultats de profil dans un fichier Microsoft Excel, l'outil Developer enregistre les informations dans un fichier .xlsx.

1. Ouvrez un profil dans la vue **Explorateur d'objets**.
2. Exécutez éventuellement le profil pour mettre à jour les résultats de profil.
3. Sélectionnez la vue **Résultats**.
4. Sélectionnez une colonne.
5. Sous **Détails**, sélectionnez **Valeurs**, **Modèles** ou **Types de données** et cliquez sur l'icône **Exporter**.
La boîte de dialogue **Exporter les données vers un fichier** s'ouvre.
6. Acceptez ou modifiez le nom de fichier par défaut.
7. Sélectionnez le type de données à exporter. Vous pouvez sélectionner **Valeurs pour la colonne sélectionnée**, **Formes pour la colonne sélectionnée**, **Types de données pour la colonne sélectionnée** ou **Tout (résumé, valeurs, modèles, types de données, statistiques, propriétés)**.
8. Cliquez sur **Parcourir** pour sélectionner un emplacement et enregistrer le fichier localement sur votre ordinateur.
9. Pour ne pas exporter les noms de champs comme première ligne, décochez la case **Exporter les noms de champs comme première ligne**.
10. Cliquez sur **OK**.

CHAPITRE 18

Règles dans Informatica Developer

Ce chapitre comprend les rubriques suivantes :

- [Présentation des règles dans Informatica Developer](#)[Instructions pour les règles, 127](#)
- [Création d'une règle dans Informatica Developer, 128](#)
- [Application d'une règle dans Informatica Developer](#)[Application d'une règle, 128](#)

Présentation des règles dans Informatica Developer

Instructions pour les règles

Une règle est une logique métier qui définit les conditions appliquées aux données source lorsque vous exécutez un profil. Vous pouvez créer des règles réutilisables à partir de mapplets dans l'outil Developer. Vous pouvez réutiliser ces règles dans les profils de l'outil Analyst pour valider les données source. Vous pouvez créer des règles réutilisables à partir de mapplets dans l'outil Developer. Vous pouvez utiliser ces règles dans les profils pour valider les données source.

Créez un mapplet et validez-la en tant que règle. Cette règle s'affiche comme règle réutilisable dans l'outil Analyst. Vous pouvez appliquer la règle à un profil de colonne dans l'outil Developer ou Analyst.

Une règle doit répondre aux exigences suivantes :

- Elle doit contenir une transformation d'entrée et de sortie. Vous ne pouvez pas utiliser de sources de données dans une règle.
- Elle peut contenir des transformations Expression, Recherche et de qualité des données passives. Elle ne peut pas contenir d'autre type de transformation. Par exemple, une règle ne peut pas contenir une transformation de correspondance, car c'est une transformation active.
- Elle ne spécifie pas de cardinalité entre les groupes d'entrée.

Une règle doit répondre aux exigences suivantes :

- Elle doit contenir une transformation d'entrée et de sortie. Vous ne pouvez pas utiliser de sources de données dans une règle.
- Elle peut contenir des transformations Expression, Recherche et passives. Elle ne peut pas contenir d'autre type de transformation. Par exemple, une règle ne peut pas contenir de transformation Trieur car il s'agit d'une transformation active.
- Elle ne spécifie pas de cardinalité entre les groupes d'entrée.

Création d'une règle dans Informatica Developer

Vous devez valider une mapplet en tant que règle pour créer une règle dans l'outil Developer.

Créez une mapplet dans l'outil Developer.

1. Cliquez avec le bouton droit sur l'éditeur de la mapplet.
2. Sélectionnez **Valider en tant que > Règle**.

Application d'une règle dans Informatica Developer

Vous pouvez ajouter une règle à un profil de colonne enregistré. Vous ne pouvez pas ajouter une règle à un profil configuré pour l'analyse de jointure. Vous pouvez ajouter une règle à un profil de colonne enregistré.

1. Parcourez la vue **Explorateur d'objets** et recherchez le profil requis.
2. Faites un clic droit sur le profil et sélectionnez **Ouvrir**.
Le profil s'ouvre dans l'éditeur.
3. Cliquez sur l'onglet **Définition** et sélectionnez Règles.
4. Cliquez sur **Ajouter**.
La boîte de dialogue **Appliquer la règle** s'ouvre.
5. Cliquez sur **Parcourir** pour rechercher la règle à appliquer.
Sélectionnez une règle dans un projet de référentiel, et cliquez sur **OK**.
6. Cliquez dans la colonne **Valeur** sous **Valeurs d'entrée** pour sélectionner un port d'entrée pour la règle.
7. Cliquez éventuellement dans la colonne **Valeur** dans **Valeurs de sortie** pour éditer le nom du port de sortie de la règle.
La règle s'affiche dans l'onglet **Définition**.

CHAPITRE 19

Fiches d'évaluation dans Informatica Developer

Ce chapitre comprend les rubriques suivantes :

- [Présentation des fiches d'évaluation dans Informatica Developer, 129](#)
- [Création d'une fiche d'évaluation, 129](#)
- [Exportation d'un fichier de ressource pour le lignage des fiches d'évaluation, 130](#)
- [Affichage du lignage des fiches d'évaluation dans Informatica Developer, 130](#)

Présentation des fiches d'évaluation dans Informatica Developer

Une fiche d'évaluation est une représentation graphique des mesures de qualité dans un profil. Vous pouvez consulter les fiches d'évaluation dans l'outil Developer. Après avoir créé une fiche d'évaluation dans l'outil Developer, vous pouvez vous connecter à l'outil Analyst pour ouvrir la fiche d'évaluation et la modifier. Exécutez la fiche d'évaluation sur les données actuelles dans l'objet de données ou sur les données stockées dans l'entrepôt de profilage.

Vous pouvez éditer une fiche d'évaluation, exécutez la fiche d'évaluation et afficher le lignage des fiches d'évaluation pour une métrologie ou un groupe de métrologies dans l'outil Analyst.

Création d'une fiche d'évaluation

Créez une fiche d'évaluation et ajoutez des colonnes depuis un profil vers la fiche d'évaluation. Vous devez exécuter un profil avant d'ajouter des colonnes à la fiche d'évaluation.

1. Dans la vue **Explorateur d'objets**, sélectionnez le projet ou le dossier où créer la fiche d'évaluation.
2. Cliquez sur **Fichier > Nouveau > Fiche d'évaluation**.

La boîte de dialogue **Nouvelle fiche d'évaluation** s'affiche.

3. Cliquez sur **Ajouter**.

La boîte de dialogue **Sélectionner le profil** s'affiche. Sélectionnez le profil qui contient les colonnes à ajouter.

4. Cliquez sur **OK**, puis sur **Suivant**.
5. Sélectionnez les colonnes à ajouter à la fiche d'évaluation.
Par défaut, l'assistant Fiche d'évaluation sélectionne les colonnes et règles définies dans le profil. Vous ne pouvez pas ajouter de colonnes non comprises dans le profil.
6. Cliquez sur **Terminer**.
L'outil Developer crée la fiche d'évaluation.
7. Cliquez éventuellement sur **Ouvrir avec Informatica Analyst** pour vous connecter à l'outil Analyst et ouvrir la fiche d'évaluation dans cet outil.

Exportation d'un fichier de ressource pour le lignage des fiches d'évaluation

Vous pouvez exporter un projet contenant des fiches d'évaluation et les objets qui en dépendent en tant que fichier de ressource pour Metadata Manager. Utilisez le fichier de ressource exporté au format XML pour créer et charger une ressource pour le lignage des fiches d'évaluation dans Metadata Manager.

1. Pour ouvrir l'assistant **Exportation**, cliquez sur **Fichier > Exporter**.
2. Sélectionnez **Informatica > Fichier de ressource pour Metadata Manager**.
3. Cliquez sur **Suivant**.
4. Cliquez sur **Parcourir** pour sélectionner un projet qui contient les objets de fiche d'évaluation et le lignage à exporter.
5. Cliquez sur **Suivant**.
6. Sélectionnez les objets de fiche d'évaluation que vous voulez exporter.
7. Entrez le nom et l'emplacement du fichier d'exportation.
8. Pour afficher les objets dépendants que l'assistant **Exportation** exporte avec les objets que vous avez sélectionnés, cliquez sur **Suivant**.
L'assistant **Exportation** affiche les objets dépendants.
9. Cliquez sur **Terminer**.
L'outil Developer exporte les objets dans le fichier XML.

Affichage du lignage des fiches d'évaluation dans Informatica Developer

Pour afficher le lignage des fiches d'évaluation pour une métrologie ou un groupe de métrologies depuis l'outil Developer, lancez l'outil Analyst.

1. Dans la vue **Explorateur d'objets**, sélectionnez le projet ou dossier qui contient la fiche d'évaluation.
2. Double-cliquez sur la fiche d'évaluation pour l'ouvrir.
La fiche d'évaluation s'affiche dans un onglet.

3. Cliquez sur **Ouvrir avec Informatica Analyst**.
L'outil Analyst s'ouvre dans la fenêtre de navigateur.
4. Dans la vue **Fiche d'évaluation** de l'outil Analyst, sélectionnez une métrologie ou un groupe de métrologies.
5. Faites un clic droit et sélectionnez **Afficher le lignage**.
Le diagramme du lignage des fiches d'évaluation s'affiche dans une boîte de dialogue.

CHAPITRE 20

Profilage de mapplet et de mappage

Ce chapitre comprend les rubriques suivantes :

- [Présentation du profilage de mapplet et de mappage](#)[Profils de mapplet et de mappage, 132](#)
- [Exécution d'un profil sur un mapplet ou un objet de mappage, 132](#)
- [Comparaison des profils pour les objets de mappage ou de mapplet, 133](#)
- [Génération d'un mappage à partir d'un profil, 133](#)

Présentation du profilage de mapplet et de mappage

Vous pouvez définir un profil de colonne pour un objet dans une mapplet ou dans un mappage. Exécutez un profil sur une mapplet ou un objet de mappage si vous souhaitez vérifier la conception du mappage ou de la mapplet sans enregistrer les résultats de profil. Vous pouvez également générer un mappage depuis un profil.

Exécution d'un profil sur un mapplet ou un objet de mappage

Lorsque vous exécutez un profil sur un mapplet ou un objet de mappage, il s'exécute sur toutes les colonnes de données et active les opérations de développement sur les données stockées pour l'objet de données. Vous pouvez exécuter un profil sur un mapplet ou un objet de mappage avec plusieurs ports de sortie. Lorsque vous exécutez un profil sur un mapplet ou un objet de mappage, il s'exécute sur toutes les colonnes de données et active les opérations de développement sur les données. Vous pouvez exécuter un profil sur un mapplet ou un objet de mappage avec plusieurs ports de sortie.

Le profil permet le suivi des données source à travers le mappage aux ports de sortie de l'objet que vous avez sélectionné. Le profil analyse les données qui s'affichent dans ces ports si vous avez exécuté le mappage.

1. Ouvrez un mapplet ou un mappage.

2. Vérifiez que le mapplet ou le mappage est valide.
3. Faites un clic droit sur un objet de données ou une transformation et sélectionnez **Profiler maintenant**.
Si la transformation a plusieurs groupes de sortie, la boîte de dialogue **Sélectionner le groupe de sortie** s'ouvre. Si la transformation a un seul groupe de sortie, les résultats de profil s'affichent dans l'onglet **Résultats** du profil.
4. Si la transformation a plusieurs groupes de sortie, sélectionnez les groupes en fonction des besoins.
5. Cliquez sur **OK**.
Les résultats de profil s'affichent dans l'onglet **Résultats** du profil.

Comparaison des profils pour les objets de mappage ou de mapplet

Vous pouvez créer un profil qui analyse deux objets dans une mapplet ou un mappage et compare les résultats des profils de colonne pour ces objets.

Comme les profils des objets à mappage ou mapplet unique, les comparaisons de profils sont exécutées dans toutes les colonnes de données et activent les opérations de développement dans les données stockées pour les objets de données.

Comme les profils d'objets à mappage ou mapplet unique, les comparaisons de profils sont exécutées dans toutes les colonnes de données.

1. Ouvrez un mapplet ou un mappage.
2. Vérifiez que le mapplet ou le mappage est valide.
3. Appuyez sur la touche **CTRL** et cliquez sur deux objets dans l'éditeur.
4. Faites un clic droit sur l'un des objets et sélectionnez **Comparer les profils**.
5. Configurez éventuellement la comparaison de profil pour faire correspondre les colonnes d'un objet à l'autre.
6. Faites correspondre éventuellement des colonnes en cliquant sur une colonne dans un objet et en la faisant glisser vers une colonne dans l'autre objet.
7. Indiquez éventuellement si le profil analyse toutes les colonnes ou les colonnes correspondantes uniquement.
8. Cliquez sur **OK**.

Génération d'un mappage à partir d'un profil

Vous pouvez créer un objet de mappage depuis un profil. Utilisez l'objet de mappage que vous créez pour développer un mappage valide. Le mappage que vous créez comprend une source de données basée sur l'objet profilé et peut contenir des transformations basées sur la logique de règle de profil. Après avoir créé le mappage, ajoutez des objets pour le compléter.

1. Dans la vue **Explorateur d'objets**, recherchez le profil dans lequel créer le mappage.
2. Cliquez avec le bouton droit sur le nom du profil et sélectionnez **Générer le mappage**.

La boîte de dialogue **Générer le mappage** s'affiche.

3. Entrez un nom de mappage. Entrez éventuellement une description du mappage.
4. Confirmez l'emplacement du dossier pour le mappage.

Par défaut, Developer Tool crée le mappage dans le dossier **Mappages** dans le même projet que le profil. Cliquez sur **Parcourir** pour sélectionner un emplacement différent pour le mappage.

5. Confirmez la définition de profil que Developer Tool utilise pour créer le mappage. Pour utiliser un autre profil, cliquez sur **Sélectionner le profil**.
6. Cliquez sur **Terminer**.

Le mappage s'affiche dans l'**Explorateur d'objets**.

Ajoutez des objets au mappage pour le compléter.

CHAPITRE 21

Découverte des domaines de données dans Informatica Developer

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la découverte de domaines de données dans Informatica Developer, 135](#)
- [Glossaire de domaine de données dans Informatica Developer, 136](#)
- [Options de découverte des domaines de données dans Informatica Developer, 140](#)
- [Création d'un profil pour découvrir des domaines de données dans Informatica Developer, 142](#)
- [Modification d'un profil dans Informatica Developer, 142](#)
- [Exécution d'un profil pour effectuer une découverte des domaines de données dans Informatica Developer, 143](#)
- [Résultats de la découverte des domaines de données dans Informatica Developer, 143](#)

Présentation de la découverte de domaines de données dans Informatica Developer

Utilisez le glossaire de domaine de données pour gérer les domaines de données. Pour créer un domaine de données, vous pouvez utiliser des règles de données et des règles de noms de colonne prédéfinies. Vous pouvez également générer le domaine de données à partir de valeurs ou de modèles spécifiques dans les résultats de profil de colonne.

Vous pouvez sélectionner les colonnes source, les domaines de données auxquels vous voulez comparer les données de colonne et le nom de colonne, et les options d'échantillonnage. Les options d'échantillonnage vous permettent d'indiquer si vous voulez exécuter la découverte de domaines de données sur les données de colonne, le nom de colonne, ou les deux. Vous pouvez également choisir le nombre maximum de lignes sur lesquelles exécuter la découverte de domaines de données et les critères de pourcentage de conformité minimum. Après avoir exécuté un profil, vous pouvez vérifier, conserver et développer les résultats. Vous pouvez également ajouter les résultats à un modèle de données dans l'éditeur de l'outil Developer.

Glossaire de domaine de données dans Informatica Developer

Vous gérez les domaines de données et les groupes de domaines de données dans le glossaire de domaine de données. Vous pouvez ajouter, modifier et supprimer les domaines de données et les groupes de domaines de données. Vous pouvez également rechercher des domaines de données et groupes de domaines de données spécifiques.

Vous pouvez exporter les domaines de données depuis le glossaire de domaine de données dans un fichier XML. Vous pouvez également importer des domaines de données depuis un fichier XML dans le glossaire de domaine de données. Vous créez un groupe de domaines de données pour organiser les domaines de données en groupes spécifiques comme par exemple Informations médicales personnelles (PHI), Informations personnelles identifiables (PII) ou tout autre groupe conceptuel pertinent pour le projet. Vous pouvez avoir un même domaine de données dans plusieurs groupes de domaines de données. Par exemple, le numéro de téléphone peut appartenir à la fois aux groupes de domaines de données PII et PHI.

Création d'un groupe de domaines de données dans Informatica Developer

Vous pouvez ajouter des domaines de données aux groupes de domaines de données pour une analyse des données de colonne effective.

1. Cliquez sur **Fenêtre > Préférences**.

La boîte de dialogue **Préférences** s'affiche.

2. Dans la boîte de dialogue **Préférences**, sélectionnez **Informatica > Glossaire de domaine de données**.

L'outil Developer affiche une liste de tous les domaines de données dans le panneau **Glossaire de domaine de données**.

3. Dans le champ **Afficher**, sélectionnez **Groupes de domaines de données**.

Le panneau **Glossaire de domaine de données** réorganise la liste des domaines de données en fonction des groupes de domaines de données.

4. Dans le panneau **Glossaire de domaine de données**, sélectionnez **Groupes de domaines de données**.

5. Cliquez sur **Ajouter**.

La boîte de dialogue **Groupe de domaines de données** s'affiche.

6. Entrez un nom et une description.

7. Cliquez sur **Suivant**.

8. Cliquez sur **Sélectionner** pour ouvrir la boîte de dialogue **Sélectionner des domaines de données**.

9. Sélectionnez les domaines de données que vous voulez ajouter au groupe de domaines de données et cliquez sur **OK**.

L'outil Developer répertorie les domaines de données sélectionnés dans le panneau **Domaines de données sélectionnés**.

10. Cliquez sur **Terminer**.

L'outil Developer ajoute le groupe de domaines de données au glossaire de domaine de données.

Création d'un domaine de données dans Informatica Developer

Vous pouvez créer des domaines de données et les ajouter au glossaire de domaine de données. Vous pouvez également ajouter des domaines de données dans un ou plusieurs groupes de domaines de données.

1. Cliquez sur **Fenêtre > Préférences**.

La boîte de dialogue **Préférences** s'affiche.

2. Dans la boîte de dialogue **Préférences**, sélectionnez **Informatica > Glossaire de domaine de données**.

L'outil Developer affiche une liste de tous les domaines de données dans le panneau **Glossaire de domaine de données**.

3. Dans le panneau **Glossaire de domaine de données**, sélectionnez **Domaines de données**.

4. Cliquez sur **Ajouter**.

La boîte de dialogue **Domaine de données** s'affiche.

5. Entrez un nom et une description.

6. Cliquez sur **Utiliser la règle de données** pour découvrir les domaines de données en fonction des données de colonne. Vous pouvez également sélectionner **Utiliser la règle de nom de colonne** pour découvrir les domaines de données en fonction des noms de colonne dans la source de données.

Le bouton **Parcourir** est activé.

7. Cliquez sur **Parcourir** pour ouvrir la boîte de dialogue **Sélectionner un emplacement**.

8. Sélectionnez les règles appropriées et cliquez sur **OK**.

Lorsque vous créez un domaine de données, l'outil Developer copie les règles et autres objets dépendants associés au domaine de données dans le glossaire de domaine de données. Pour modifier une règle associée à un domaine de données, vous devez accéder à la règle d'origine et la modifier. Vous pouvez ensuite associer à nouveau la règle modifiée au domaine de données.

Les règles que vous avez sélectionnées s'affichent dans les champs **Règle de données** et **Règle de nom de colonne**.

9. Cliquez sur **Suivant**.

10. Cliquez sur **Sélectionner** pour ouvrir la boîte de dialogue **Sélectionner des groupes de domaines de données**.

11. Sélectionnez les groupes de domaines de données que vous voulez inclure au domaine de données et cliquez sur **OK**.

L'outil Developer affiche les groupes de domaines de données sélectionnés dans le volet **Affecter à des groupes de domaines de données**.

12. Cliquez sur **Terminer**.

L'outil Developer ajoute le domaine de données au glossaire de domaine de données.

Création d'un domaine de données depuis les résultats de profil dans Informatica Developer

Après avoir exécuté un profil de colonne, vous pouvez afficher les valeurs et les modèles des données source. Vous pouvez alors créer un domaine de données.

1. Exécutez un profil de colonne pour afficher ses résultats.
2. Sélectionnez les valeurs ou les modèles en fonction de ce que vous voulez utiliser pour créer un domaine de données.

Les valeurs, les modèles et les statistiques s'affichent dans la vue **Résultats**.

3. Effectuez un clic droit sur les valeurs ou les modèles, puis sélectionnez **Envoyer à > Nouveau domaine de données**.
La boîte de dialogue **Domaine de données** s'ouvre.
4. Entrez le nom du domaine de données et éventuellement une description.
L'emplacement est défini sur le glossaire de domaine de données par défaut.
5. Cliquez sur **Terminer**.
Le domaine de données est ajouté au glossaire de domaine de données.

Rechercher des domaines de données dans Informatica Developer

Le glossaire de domaine de données affiche tous les domaines de données par défaut. Vous pouvez rechercher des domaines de données et des groupes de domaines de données spécifiques.

Le volet **Glossaire de domaine de données** de la boîte de dialogue **Préférences** affiche tous les domaines de données et les groupes de domaines de données. Vous pouvez rechercher et afficher de plus amples informations sur les domaines de données et les groupes de domaines de données des manières suivantes :

Rechercher des domaines de données et des groupes de domaines de données.

Saisissez une partie du nom de domaine ou du nom du groupe de domaines de données dans le champ en haut du panneau **Glossaire de domaine de données**. Si vous êtes dans la **Vue Groupes de domaines de données**, l'outil Developer répertorie les domaines de données dont le nom contient la chaîne de recherche ainsi que les groupes de domaines de données qui y sont associés. Si vous êtes dans la **Vue Domaine de données**, l'outil Developer répertorie tous les domaines de données dont le nom contient la chaîne de recherche.

Afficher les groupes de domaines de données et les domaines de données qu'ils contiennent.

Dans le champ **Afficher**, sélectionnez **Groupes de domaines de données**.

Affichez tous les domaines de données.

Dans le champ **Afficher**, sélectionnez **Domaines de données**.

Affichez les propriétés d'un domaine de données.

Cliquez sur un nom de domaine de données pour afficher ses propriétés sous le panneau **Glossaire de domaine de données**. Vous pouvez afficher la description et les règles associées. Pour afficher les groupes de domaines auquel appartient un domaine de données, cliquez sur **Afficher dans les groupes de domaines de données**.

Afficher les propriétés d'un groupe de domaines de données.

Cliquez sur un nom de groupe de domaines de données pour afficher sa description sous le panneau **Glossaire de domaine de données**.

Importation des domaines de données

Vous pouvez importer des domaines de données depuis un fichier XML source dans le glossaire de domaine de données de l'outil Developer. Vous devez vérifier que le fichier contient des informations sur les domaines de données que vous avez besoin d'importer.

1. Ouvrez le glossaire de domaine de données.
2. Vérifiez que **Domaines de données** ou **Groupes de domaines de données** est sélectionné.
3. Cliquez sur **Importer**.
La boîte de dialogue **Importer** s'affiche.

4. Dans le champ **Nom de fichier**, entrez le nom du fichier XML depuis lequel vous voulez importer les domaines de données.
Cliquez sur **Parcourir** pour sélectionner le fichier.
5. Cliquez sur **Suivant**.
Le volet **Sélectionner les objets à importer** dans lequel vous pouvez indiquer la source et la cible s'affiche.
6. Dans le panneau **Source**, sélectionnez les domaines de données que vous voulez importer.
Remarque: Pour sélectionner plusieurs domaines de données, maintenez la touche Maj enfoncée.
7. Cliquez sur **Correspondance auto. à la cible** pour déplacer les domaines de données vers le panneau **Cible**.
L'outil Developer tente de faire correspondre individuellement les descendants de la sélection source actuelle par nom, type et hiérarchie parente dans la sélection cible, et ajoute les objets correspondants.
8. Cliquez sur **Résolution** pour indiquer le mode de traitement des objets dupliqués.
Vous pouvez renommer l'objet importé, remplacer l'objet existant par l'objet importé ou réutiliser l'objet existant. L'outil Developer renomme tous les objets dupliqués par défaut.
9. Cliquez sur **Suivant**.
L'outil Developer récapitule les paramètres d'importation pour que vous les vérifiiez. Vous pouvez spécifier des paramètres d'importation supplémentaires dans le volet **Paramètres d'importation supplémentaires**.
10. Cliquez sur **Terminer**.

Exportation des domaines de données

Vous pouvez exporter des domaines de données et des règles de domaine de données depuis le glossaire de domaine de données dans l'outil Developer vers un fichier XML.

1. Ouvrez le glossaire de domaine de données.
2. Vérifiez que **Domaines de données** ou **Groupes de domaines de données** est sélectionné.
3. Cliquez sur **Exporter**.
La boîte de dialogue **Exporter** s'affiche.
4. Pour exporter des domaines de données, sélectionnez **Exporter des domaines de données**. Sélectionnez **Exporter des règles de domaine de données** pour exporter les règles de domaines de données.
5. Cliquez sur **Suivant**.
6. Dans le panneau **Exporter dans le fichier**, sélectionnez les domaines de données ou les règles de domaine de données que vous voulez exporter.
7. Pour exporter des domaines de données, cliquez sur **Parcourir** pour choisir le fichier d'exportation et son emplacement. Pour exporter des règles de domaines de données vers un autre projet dans le service de référentiel modèle, sélectionnez **Copier vers le projet** et sélectionnez le projet vers lequel copier les règles de domaine de données.
8. Cliquez sur **Suivant**.
Le volet **Dépendances** affiche une liste des objets dépendants.
9. Cliquez sur **Suivant**.
Le volet **Paramètres d'exportation du contenu** s'affiche. Vous pouvez sélectionner les tables de référence associées pour l'exportation.

10. Cliquez sur **Terminer**.

Si vous associez une règle qui utilise les tables de référence avec un domaine de données, vous ne pourrez peut-être pas exporter les tables de référence dans la même session de l'outil Developer que vous utilisez pour créer le domaine de données. Après avoir cliqué sur **Exporter** dans le glossaire de domaine de données, déconnectez-vous du Model Repository Service et reconnectez-vous avant d'importer la règle qui utilise les tables de référence.

Options de découverte des domaines de données dans Informatica Developer

Vous pouvez sélectionner les colonnes source, les domaines de données et les options d'inférence lorsque vous créez un profil pour effectuer une découverte de domaines de données. Vous pouvez également choisir d'omettre des colonnes de la découverte des domaines de données en fonction de leurs types de données et de la longueur des données.

Sélection de domaines de données dans Informatica Developer

Les options **Sélection de domaines de données** répertorient tous les domaines du glossaire de domaine de données. Vous pouvez rechercher des domaines de données spécifiques et les sélectionner avant de les exécuter dans le cadre d'une découverte de domaines de données.

Le tableau suivant décrit les options **Sélection de domaines de données** pour la découverte de domaines de données :

Option	Description
Activé dans le cadre de l'action « Exécuter le profil »	Inclut les options de découverte de domaines de données lorsque vous exécutez le profil.
Nom	Nom du domaine de données.
Description	Description du domaine de données.
Groupe de domaines de données	Nom des groupes de domaines de données auquel le domaine de données appartient.
Afficher le groupe de domaines de données dans la hiérarchie	Répertorie tous les groupes de domaines de données, avec les domaines de données regroupés sous chaque groupe de domaines de données.

Sélection de colonne de domaine de données dans Informatica Developer

Vous utilisez les options **Sélection de colonne** pour sélectionner les colonnes que vous souhaitez exécuter dans le cadre d'une découverte de domaines de données.

Le tableau suivant décrit les options **Sélection de colonne** pour la découverte de domaines de données :

Option	Description
Colonne	Nom de la colonne.
Type de données	Type de données de la colonne.
Précision	Précision maximum de la colonne.
Échelle	Échelle de la colonne.
Valeur Null autorisée	Indique une colonne pouvant posséder des valeurs null.
Description	Description de la colonne.

Options d'inférence des domaines de données dans Informatica Developer

Les options d'inférence déterminent si la découverte des domaines doit s'exécuter sur les données de colonne, les noms de colonne ou les deux. Vous pouvez également indiquer si le profil doit traiter toutes les lignes dans la source de données et définir un pourcentage minimum de conformité.

Le tableau suivant décrit les options d'**inférence** pour la découverte de domaine de données :

Option	Description
Remplacer les options d'inférence par défaut	Vous permet de modifier les options d'inférence prédéfinies.
Données	Le profil s'exécute sur les données de colonne.
Nom de colonne	Le profil s'exécute sur les titres de colonne.
Nom des colonnes et des données	Le profil s'exécute sur les données de colonne et les titres de colonne.
Toutes les lignes	Le profil s'exécute sur toutes les lignes de la source de données.
Nombre maximum de lignes à profiler	Nombre de lignes maximum sur lequel le profil peut s'exécuter. L'outil Developer choisit les lignes à partir de la première ligne dans la source.
Pourcentage minimum de conformité	Pourcentage de conformité minimum des données de colonne éligibles pour la correspondance de domaine de données. Le pourcentage de conformité est le rapport entre le nombre de lignes correspondantes et le nombre total de lignes. Remarque: L'outil Developer considère les valeurs null comme des lignes qui ne correspondent pas. Les colonnes contenant un grand nombre de valeurs null ne peuvent pas entraîner une inférence de domaine de données sauf si vous spécifiez une valeur faible pour le pourcentage de conformité minimum.

Création d'un profil pour découvrir des domaines de données dans Informatica Developer

Vous pouvez découvrir des domaines de données dans une source de données dans le cadre d'un profil d'objet de données unique ou d'un profil de découverte des données d'entreprise. Après avoir effectué la découverte de domaines de données, vous pouvez vérifier, développer les résultats et les ajouter à un modèle de données à partir de l'éditeur dans l'outil Developer.

1. Dans la vue **Explorateur d'objets**, sélectionnez le projet contenant l'objet de données pour le profil.
2. Cliquez avec le bouton droit de la souris sur l'objet de données et sélectionnez **Profil**.
L'assistant **Nouveau** s'affiche.
3. Sélectionnez **Profil**.
4. Cliquez sur **Suivant**.
L'outil Developer affiche un autre volet dans lequel vous pouvez configurer les propriétés générales du profil.
5. Modifiez le nom et la description du profil, si nécessaire. Vous pouvez également ajouter ou supprimer des objets de données.
6. Cliquez sur **Suivant**.
7. Sélectionnez les colonnes sur lesquelles vous voulez exécuter la découverte de domaines de données et les domaines de données auxquels vous voulez comparer les colonnes.
8. Modifiez les options d'inférence par défaut, si nécessaire.
9. Cliquez sur **Terminer** pour créer le profil.

Modification d'un profil dans Informatica Developer

Vous pouvez modifier un profil après l'avoir configuré pour la découverte de domaines de données. Vous pouvez exclure des colonnes avec des types de données spécifiques, modifier la sélection de colonne, la sélection de domaine de données et les options d'inférence.

1. Dans la vue **Explorateur d'objets**, sélectionnez le projet ou le dossier qui contient le profil que vous voulez modifier.
2. Double-cliquez sur le profil pour l'ouvrir.
La définition de profil s'affiche dans un onglet.
3. Modifiez la sélection de colonne, la sélection de domaine de colonne et les options d'inférence si nécessaire.
4. Dans la section **Sélection de colonne**, vous pouvez cliquer sur **Exclure les colonnes** pour configurer les options d'exclusion en fonction des types de données.
La boîte de dialogue **Exclure les colonnes** s'affiche.
5. Enregistrez les modifications.

Exécution d'un profil pour effectuer une découverte des domaines de données dans Informatica Developer

Vous pouvez choisir d'exécuter le profil immédiatement après l'avoir créé. Vous pouvez également exécuter un profil manuellement après l'avoir créé.

1. Dans la vue **Explorateur d'objets**, sélectionnez le projet ou le dossier qui contient le profil que vous voulez exécuter.
Pour exécuter un profil automatiquement, sélectionnez **Exécuter le profil à la fin** dans l'assistant **Nouveau profil** lorsque vous créez le profil.
2. Double-cliquez sur le profil pour l'ouvrir.
La définition de profil s'affiche dans un onglet.
3. Cliquez avec le bouton droit de la souris sur le profil et sélectionnez **Exécuter le profil**.
La boîte de dialogue **Exécuter le profil** indiquant le statut d'exécution du profil s'affiche.

Résultats de la découverte des domaines de données dans Informatica Developer

Les résultats de la découverte des domaines de données affichent les statistiques sur les colonnes qui correspondent aux domaines de données, y compris le pourcentage de données de colonne correspondantes et si les noms de colonne correspondent aux domaines de données.

Vous pouvez développer les résultats pour les analyser. Vous pouvez également vérifier les résultats sur toutes les lignes de la source de données et ajouter les résultats à un modèle de données à partir de l'éditeur dans l'outil Developer. Vous pouvez trier les résultats en fonction des domaines de données, des groupes de domaines de données et des colonnes. Vous pouvez exporter les résultats de la découverte des domaines de données vers un fichier Microsoft Excel.

Le tableau suivant décrit les résultats de la découverte des domaines de données :

Nom de colonne	Description
Nom	Nom du domaine de données, du groupe de domaines de données ou de la colonne selon si vous sélectionnez la vue Domaine de données , Groupe de domaines de données ou Colonnes .
Connexion	Nom de la connexion.
Statut	Le statut de l'inférence de la colonne.
% de conformité des données	Pourcentage des données de colonne qui correspond à un domaine de données.
Conforme au nombre de lignes	Nombre total de lignes correspondant aux critères de domaine de données.
% Null	Pourcentage de valeurs null pour la colonne.

Nom de colonne	Description
Lignes totales	Nombre total de lignes.
Correspondance de nom de colonne	Détermine si le nom de colonne correspond à un nom de domaine de données.
Groupes de domaines de données	Groupe de domaines de données auquel appartient le domaine de données.
Type de données consigné	Type de données déclaré pour la colonne de l'objet de profil.
Développement	Lorsqu'elle est sélectionnée, cette option développe les données de colonne.
Vérifié	Indique la validation de la correspondance du domaine de données sur toutes les lignes de la source de données.
Heure de la dernière exécution	Date et heure de la dernière exécution du profil.

Affichage par groupes de domaines de données dans Informatica Developer

Vous pouvez afficher les résultats de la découverte des domaines de données triés par groupes de domaines de données.

1. Exécutez le profil pour afficher ses résultats.
2. Cliquez sur **Résultats**.
3. Cliquez sur **Découverte de domaines de données**.

Vous pouvez afficher les résultats de la découverte des domaines de données dans le panneau de droite.

4. Vérifiez que l'option **Domaine de données** est sélectionnée dans le champ **Afficher**.
5. Sélectionnez **Afficher la hiérarchie du groupe de domaines de données** pour afficher les résultats par groupes de domaines de données.

Affichage par colonnes dans Informatica Developer

Vous pouvez afficher les résultats de la découverte des domaines de données triés par colonnes et qui correspondent aux domaines de données.

1. Exécutez le profil pour afficher ses résultats.
2. Cliquez sur **Résultats**.
3. Cliquez sur **Découverte de domaines de données**.

Vous pouvez afficher les résultats de la découverte des domaines de données dans le panneau de droite.

4. Sélectionnez **Colonnes** pour afficher les résultats triés par colonnes source et qui correspondent aux domaines de données.

Vérification des résultats dans Informatica Developer

Lorsque vous exécutez un profil, il analyse un échantillon de la source de données pour inférer les résultats de profil. Vous pouvez exécuter le profil sur toutes les lignes de la source de données pour vérifier les résultats d'inférence.

1. Exécutez le profil pour en afficher les résultats.
2. Cliquez sur **Résultats**.
3. Cliquez sur **Découverte de domaines de données**.
Vous pouvez afficher les résultats de la découverte des domaines de données dans le panneau de droite.
4. Sélectionnez une colonne que vous souhaitez vérifier dans le panneau de droite.
5. Cliquez avec le bouton droit de la souris sur la colonne, puis sélectionnez **Vérifier** pour exécuter le profil sur toutes les lignes de la source de données.
Vous pouvez voir une modification de la valeur **% de conformité des données** après avoir vérifié les résultats.
6. Pour vérifier les résultats d'inférence de plusieurs colonnes, sélectionnez plusieurs colonnes. Vous pouvez alors cliquer avec le bouton droit et sélectionner **Vérifier tout**.

Approbation des domaines de données dans l'outil Developer

Si vous exécutez la découverte de domaines de données dans un profil d'objet de données unique, vous pouvez approuver les domaines de données inférés de plusieurs colonnes à la fois. Si vous exécutez la découverte de domaines de données dans le cadre de la découverte des données d'entreprise, vous pouvez approuver le domaine de données d'une seule colonne source à la fois. Pour valider les domaines de données de plusieurs colonnes après la découverte des données d'entreprise, vous pouvez ouvrir les tâches individuelles de profil d'objet de données et approuver les domaines de données.

1. Dans la vue **Explorateur d'objets**, sélectionnez un profil.
2. Double-cliquez sur le profil pour l'ouvrir.
Le profil s'ouvre dans un onglet.
3. Si vous avez exécuté un profil d'objet de données unique, sélectionnez la vue **Découverte de domaines de données**, puis sélectionnez une ligne. La ligne contient les résultats de découverte de domaine de données de chaque colonne.
4. Cliquez avec le bouton droit sur la ligne, puis sélectionnez **Accepter**.
Le statut de l'inférence du domaine de données devient **Accepté**.
5. Si vous avez exécuté la découverte des données d'entreprise, choisissez la vue **Domaines de données**, puis sélectionnez un domaine de données.
Les colonnes qui correspondent au domaine de données s'affichent dans le panneau de droite.
6. Cliquez avec le bouton droit de la souris sur la colonne à approuver, puis sélectionnez **Accepter**. Vous pouvez également sélectionner plusieurs colonnes rejetées et les approuver selon les besoins.
Le statut de l'inférence du domaine de données devient **Accepté**.
7. Pour restaurer le statut inféré du domaine de données, cliquez avec le bouton droit sur la ligne, puis cliquez sur **Réinitialiser**.

Rejet des domaines de données dans l'outil Developer

Lorsque vous affichez les résultats de profil, l'outil Developer affiche les domaines de données inférés par défaut. Vous pouvez rejeter des domaines de données inférés ou approuvés. Vous pouvez afficher ou masquer les domaines de données rejetés.

1. Dans la vue **Explorateur d'objets**, sélectionnez un profil.
2. Double-cliquez sur le profil pour l'ouvrir.
Le profil s'ouvre dans un onglet.
3. Sélectionnez une ligne dans la vue **Découverte de domaines de données** ou **Domaines de données**.
4. Pour rejeter les domaines de données inférés, cliquez avec le bouton droit sur la ligne, puis sélectionnez **Rejeter**.
L'outil Developer grise le domaine de données rejeté dans les résultats de la découverte de domaines de données.
5. Pour masquer les domaines de données rejetés, cliquez avec le bouton droit sur la ligne, puis sélectionnez **Masquer les éléments rejetés**.
6. Pour afficher les domaines de données rejetés, cliquez avec le bouton droit sur l'une des lignes, puis sélectionnez **Afficher les éléments rejetés**.

Exportation des résultats de la découverte des domaines de données depuis Informatica Developer

Quand vous exportez les résultats de la découverte des domaines de données vers un fichier `.xlsx` depuis l'outil Developer, vous pouvez enregistrer le fichier sur le serveur ou un emplacement spécifique sur la machine du client.

1. Exécuter un profil pour effectuer la découverte des domaines de données.
2. Cliquez sur la vue **Résultats**.
3. Cliquez sur l'icône **Exporter les résultats vers le fichier**.
La boîte de dialogue **Exporter les données vers un fichier** s'ouvre.
4. Entrez le nom de fichier. Utilisez éventuellement le nom de fichier par défaut.
5. Sous **Enregistrer**, choisissez **Enregistrer sur le client** et cliquez sur **Parcourir** pour sélectionner un emplacement et enregistrer le fichier localement sur votre ordinateur. Par défaut, Informatica Developer enregistre le fichier dans un emplacement du serveur défini dans les propriétés du service d'intégration de données d'Informatica Administrator.
6. Cliquez sur **OK**.

CHAPITRE 22

Découverte des données d'entreprise dans Informatica Developer

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la découverte des données d'entreprise dans Informatica Developer, 147](#)
- [Processus de découverte des données d'entreprise, 148](#)
- [Options de profil pour la découverte des données d'entreprise, 149](#)
- [Création d'un profil de découverte des données d'entreprise dans Informatica Developer, 152](#)
- [Modification d'un profil, 153](#)
- [Exécution d'un profil de découverte des données d'entreprise, 153](#)
- [Découverte de clé étrangère, 154](#)
- [Analyse de jointure, 156](#)
- [Découverte du chevauchement, 158](#)
- [Fichiers de script DDL, 160](#)

Présentation de la découverte des données d'entreprise dans Informatica Developer

La découverte des données d'entreprise est le processus de découverte des statistiques de profil de colonne, des domaines de données, des clés principales et des clés étrangères dans un grand nombre de sources de données. Vous pouvez effectuer la découverte des données d'entreprise sur plusieurs connexions ou schémas.

En tant qu'analyste de données d'une entreprise, vous aurez peut-être à découvrir les caractéristiques des données importantes sur un grand nombre de sources de données. Les conditions requises peuvent inclure l'identification des ressources des données relationnelles, les exécutions du profil de colonne sur les ressources des données découvertes, la découverte des caractéristiques de données critiques au sein de l'entreprise, des clés principales et des clés candidates. Vous voudrez peut-être également afficher les relations de clé étrangère qui existent dans les sources de données afin de pouvoir dériver un modèle de données en fonction des relations découvertes.

La découverte des données d'entreprise recherche les problèmes, schémas, tendances et caractéristiques de données critiques dans les ressources de votre entreprise. Vous pouvez choisir les sources de données que vous avez importées dans le référentiel modèle et les sources de données provenant de connexions relationnelles externes. Le processus de découverte des données inclut la découverte des statistiques de profil de colonne, l'analyse des domaines de données, les structures d'objet de données, y compris les clés candidates et les relations d'objet de données incluant les clés étrangères. Vous exécutez la découverte des données d'entreprise dans l'outil Developer et il effectue les tâches suivantes sur chaque source de données :

- Exécuter un profil de colonne.
- Découvrir les domaines de données.
- Inférer les clés principales.

Après l'exécution des profils de colonne, de la découverte des domaines de données et des profils de clé principale, l'outil Developer exécute un profil de clé étrangère sur toutes les sources de données. Une fois que l'outil Developer a terminé les tâches de profilage et de découverte, il génère un récapitulatif des résultats consolidés aux formats graphique et tabulaire.

Processus de découverte des données d'entreprise

Vous pouvez exécuter le profil de découverte des données d'entreprise pour effectuer la découverte des données d'entreprise dans l'outil Developer. Vous devez configurer les options de découverte de données pour différents types de profil avant d'exécuter le profil.

L'outil Developer crée des objets de données pour les sources de données sélectionnées et des tâches de profil pour chaque objet de données. Il exécute ensuite les tâches de profil pour générer les résultats de profil.

Procédez comme suit pour effectuer la découverte des données d'entreprise :

1. Créez un profil de découverte des données d'entreprise en sélectionnant plusieurs objets de données importés dans le référentiel modèle et les sources de données sur plusieurs connexions relationnelles externes.
2. Définissez les paramètres de configuration pour la découverte des domaines de données, le profil de colonne, le profil de clé principale et le profil de clé étrangère.
3. Exécutez le profil de découverte des données d'entreprise.
4. Actualisez le service de référentiel modèle.

Remarque: Cette action doit être exécutée, car l'importation des métadonnées pour les connexions externes se produit dans le référentiel modèle. Vous devez actualiser le service de référentiel modèle pour que l'outil Developer reflète les modifications du référentiel modèle.

5. Surveillez l'exécution du profil et, si nécessaire, affichez les statuts des tâches de profil exécutées par l'outil Developer.
6. Consultez le résumé des résultats de découverte des données d'entreprise. Le résumé inclut une vue de l'interface utilisateur interactive et une vue tabulaire.

Options de profil pour la découverte des données d'entreprise

Configurez les options de profil avant d'exécuter un profil pour effectuer une découverte des données d'entreprise. Les options de profil incluent les options de découverte des domaines de données, les options d'échantillonnage de profil de colonne et les options d'inférence pour les clés principale et étrangères.

Vous pouvez choisir d'exécuter le profil de découverte des données d'entreprise après avoir configuré les options de profil. Vous pouvez également choisir de créer des tâches de profil après la configuration, sans exécuter le profil.

Sélection de domaine de données pour la découverte des données d'entreprise

Les options d'inférence déterminent si la découverte des domaines de données doit s'exécuter sur les données de colonne, les noms de colonne ou les deux. Vous pouvez également indiquer si le profil doit traiter toutes les lignes dans la source de données et définir un pourcentage minimum de conformité.

Le tableau suivant décrit les options d'inférence de domaine de données que vous configurez pour la découverte des données d'entreprise :

Option	Description
Remplacer les options d'inférence par défaut	Modifie les options d'inférence prédéfinies.
Données	Le profil s'exécute sur les données de colonne.
Nom de colonne	Le profil s'exécute sur les titres de colonne.
Nom des colonnes et des données	Le profil s'exécute sur les données de colonne et les titres de colonne.
Toutes les lignes	Le profil s'exécute sur toutes les lignes de la source de données.
Nombre maximum de lignes à profiler	Nombre de lignes maximum sur lequel le profil peut s'exécuter. L'outil Developer choisit des lignes à partir de la première ligne dans la source.
Pourcentage minimum de conformité	Pourcentage de conformité minimum des données de colonne éligibles pour la correspondance de domaine de données. Le pourcentage de conformité est le rapport entre le nombre de lignes correspondantes et le nombre total de lignes. Remarque: L'outil Developer considère les valeurs Null comme des lignes qui ne correspondent pas.
Exclure les colonnes avec des domaines de données approuvés	Exclut les colonnes qui comprennent des domaines de données approuvés issus de l'inférence du domaine de données de l'exécution du profil.

Options d'échantillonnage de profil de colonne pour la découverte des données d'entreprise

Les options d'échantillonnage déterminent si l'outil Developer exécute un profil de colonne sur toutes les lignes des sources de données ou sur un nombre limité de lignes.

Le tableau suivant décrit les options d'échantillonnage de profil de colonne que vous configurez pour la découverte des données d'entreprise :

Option	Description
Toutes les lignes	Choisit toutes les lignes dans la source de données.
<nombre> première(s) ligne(s)	Nombre de lignes sur lequel vous voulez exécuter le profil de colonne. L'outil Developer choisit les lignes en commençant par la première ligne dans la source de données.
Exclure l'inférence du type de données pour les colonnes avec un type de données approuvé	Exclut les colonnes disposant d'un type de données approuvé de l'inférence du type de données dans l'exécution du profil de colonne.

Options d'inférence de clé principale pour la découverte des données d'entreprise

Vous pouvez écraser les options d'inférence de clé principale par défaut pour la découverte des données d'entreprise. Les options incluent le nombre maximum de lignes sur lequel vous pouvez effectuer l'exécution et le pourcentage de conformité minimum.

Le tableau suivant décrit les options d'inférence de clé principale que vous configurez pour la découverte des données d'entreprise :

Options	Description
Remplacer les options d'inférence par défaut	Permet de configurer les paramètres personnalisés pour l'inférence de clé principale.
Nombre max. de clés	Nombre maximum de colonnes qui peuvent former une clé principale.
Nombre max. de lignes	Nombre de lignes maximum sur lequel vous voulez exécuter le profil.
Pourcentage minimum	Pourcentage de conformité minimum des données de colonne éligibles pour la correspondance de clé principale.
Rangées de violation maximum	Le nombre maximum de lignes avec violations de clés que le profil autorise lors de la définition des clés principales.

Options d'inférence de clé étrangère pour la découverte des données d'entreprise

Définissez les options d'inférence de clé étrangère pour définir les paramètres de colonne afin de découvrir les relations de clé étrangère entre les objets de données. Les résultats d'inférence de clé étrangère

dépendent des options d'inférence de clé principale que vous définissez pour la découverte des données d'entreprise, des clés principales documentées et des clés principales définies par l'utilisateur.

Le tableau suivant décrit les options d'inférence de clé étrangère que vous configurez pour la découverte des données d'entreprise :

Options	Description
Remplacer les options d'inférence par défaut	Modifie les options d'inférence prédéfinies.
Types de données utilisés dans les comparaisons	Type de données utilisé dans les comparaisons de clé primaire et de clé étrangère. Remarque: Cette option s'applique si vous exécutez un profil de colonne sur la source de données avant l'inférence de clé étrangère.
Sensibilité à la casse de la comparaison	Inclut la sensibilité à la casse lors de la comparaison des données de colonne.
Raccourcir les valeurs avant comparaison	Détermine si l'outil Developer inclut des espaces au début ou à la fin dans les données de colonne lors du traitement.
Clés primaires inférées utilisées dans les comparaisons Utilisez des clés de rang supérieur	Nombre de clés primaires de rang supérieur utilisées dans l'inférence de clé étrangère lorsque l'outil Developer exécute un profil de clé étrangère sur toutes les sources de données. L'outil Developer utilise la méthode de rang supérieur avec les clés primaires documentées et les clés primaires définies par l'utilisateur pour inférer les relations de clé étrangère. Le classement des clés inférées est basé sur le pourcentage de conformité décroissant arrondi à une seule décimale. Par exemple, l'outil Developer considère un pourcentage de conformité de 99,75 comme étant 99,8 et 99,74 comme étant 99,7. La valeur par défaut est 1. Définissez la valeur sur -1 si vous voulez que l'outil Developer utilise toutes les clés inférées dans l'inférence de clé étrangère. Remarque: Si les sources de données de clé primaire disposent de clés primaires approuvées, l'outil Developer n'utilise pas de clés primaires inférées pour l'inférence de clé étrangère.
Nombre maximal de clés étrangères entre les objets de données	Nombre maximal de colonnes inférées renvoyées par l'outil Developer après l'exécution du profil et qui sont éligibles pour la découverte des clés étrangères.
Pourcentage minimal de conformité	Valeur d'éligibilité minimale en pourcentage pour l'inclusion des colonnes dans les résultats de clé étrangère.
Régénérer la signature	Recharge les signatures de colonne si les données source changent.

Création d'un profil de découverte des données d'entreprise dans Informatica Developer

Vous pouvez créer un profil sur plusieurs sources de données sous plusieurs connexions. L'outil Developer crée des tâches de profil individuelles pour chaque source.

1. Dans la vue **Explorateur d'objets**, sélectionnez les objets de données sur lesquels vous voulez exécuter un profil.
2. Cliquez sur **Fichier > Nouveau > Profil** pour ouvrir l'assistant Création de profils.
3. Sélectionnez **Profil de découverte d'entreprise** et cliquez sur **Suivant**.
4. Entrez un nom pour le profil et vérifiez l'emplacement du projet. Si nécessaire, accédez à un nouvel emplacement.
5. Vérifiez que le nom des objets de données sélectionnés s'affiche dans la section **Objets de données**. Cliquez sur **Choisir** pour sélectionner d'autres d'objets de données, si nécessaire.
6. Cliquez sur **Suivant**.

Le volet **Ajouter les ressources à la définition de profil** s'affiche. Vous pouvez ajouter plusieurs connexions relationnelles externes et sources de données à partir de ce volet.

7. Cliquez sur **Choisir** pour ouvrir la boîte de dialogue **Sélectionner des ressources**.
Le volet **Ressources** répertorie toutes les connexions internes et externes et les objets de données sous le domaine Informatica.
8. Cliquez sur **OK** pour fermer la boîte de dialogue.
9. Cliquez sur **Suivant**.
10. Configurez les types de profil que vous souhaitez exécuter. Vous pouvez configurer les types de profil suivants :
 - Découverte de domaines de données
 - Profil de colonne
 - Profil de clé principale
 - Profil de clé étrangère

Remarque: Sélectionnez **Activé dans le cadre de l'action « Exécution du profil de découverte d'entreprise »** pour les types de profil à exécuter dans le cadre du profil de découverte des données d'entreprise. Le profilage de colonne est activé par défaut.

11. Vérifiez les options du profil.
Vous pouvez modifier les options d'échantillonnage pour les profils de colonne. Vous pouvez également modifier les options d'inférence pour les profils de domaine de données, de clé primaire et de clé étrangère.
12. Sélectionnez **Créer des profils**.
L'outil Developer crée des profils pour chaque source de données.
13. Sélectionnez **Exécuter le profil de découverte d'entreprise à la fin** pour exécuter le profil lorsque la configuration du profil est terminée. Si vous avez activé toutes les opérations de profilage, l'outil Developer exécute les profils de colonne, de domaine de données et de clé primaire sur toutes les sources de données sélectionnées. Il exécute ensuite un profil de clé étrangère sur toutes les sources de données.

14. Cliquez sur **Terminer**.

Après avoir exécuté un profil de découverte des données d'entreprise, vous devez actualiser le service de référentiel modèle avant de consulter les résultats. Cette étape est obligatoire, car l'importation de métadonnées pour les connexions externes s'effectue dans le référentiel modèle. Vous devez actualiser le service de référentiel modèle pour que l'outil Developer reflète les modifications du référentiel modèle.

Modification d'un profil

Vous pouvez apporter des modifications à un profil de découverte des données d'entreprise après l'avoir configuré. Vous pouvez exclure des colonnes avec des types de données spécifiques, modifier la sélection de colonne, la sélection de domaine de données et les options d'inférence.

1. Dans la vue **Explorateur d'objets**, sélectionnez le projet ou le dossier qui contient le profil que vous voulez modifier.
2. Double-cliquez sur le profil pour l'ouvrir.
3. Cliquez sur la vue **Propriétés**.
La vue Propriétés se trouve dans la vue Par défaut.
4. Cliquez sur **Profils** pour afficher les tâches de profil.
5. Sélectionnez une tâche de profil à modifier dans le volet de droite et cliquez sur **Ouvrir**.
La définition de profil s'affiche dans un onglet.
6. Pour modifier les paramètres généraux du profil de découverte des données d'entreprise, sélectionnez le profil en haut de la **Liste des tâches de profilage** et cliquez sur **Configurer**.
7. Apportez les modifications nécessaires aux options de définition de profil.
8. Enregistrez les modifications.

Exécution d'un profil de découverte des données d'entreprise

Vous pouvez exécuter un profil de découverte des données d'entreprise de plusieurs manières. Vous pouvez exécuter le profil depuis la vue **Explorateur d'objets** ou depuis l'onglet **Profils** de la fenêtre **Propriétés**. Vous pouvez choisir d'exécuter des tâches de profil individuelles et multiples qui forment une partie du profil de découverte des données d'entreprise.

1. Dans la vue **Explorateur d'objets**, sélectionnez le projet ou le dossier qui contient le profil que vous voulez exécuter.
Pour exécuter un profil automatiquement, sélectionnez **Exécuter le profil de découverte des données d'entreprise à la fin** dans l'assistant **Nouvelle découverte des données d'entreprise** lorsque vous créez le profil.
2. Double-cliquez sur le profil pour l'ouvrir.
Le profil s'ouvre dans un onglet.

3. Dans la vue **Explorateur d'objets**, cliquez avec le bouton droit de la souris sur le profil et sélectionnez **Exécuter un profil de découverte des données d'entreprise**.

Vous pouvez également sélectionner **Profils** dans la fenêtre **Propriétés**, sélectionner le nom du profil sous **Liste des tâches de profil**, puis cliquer sur **Exécuter**.

Remarque: Lorsque vous exécutez un profil de découverte des données d'entreprise, vous devez actualiser le service de référentiel modèle pour pouvoir afficher les résultats. Cette étape est obligatoire, car l'importation de métadonnées pour les connexions externes s'effectue dans le référentiel modèle. Vous devez actualiser le service de référentiel modèle pour que l'outil Developer reflète les modifications du référentiel modèle.

4. La boîte de dialogue **Exécuter** s'affiche. Vous pouvez modifier les paramètres généraux du profil dans cette boîte de dialogue.

Par défaut, les modifications que vous effectuez s'appliquent aux objets de données nouvellement ajoutés au profil de découverte des données d'entreprise.

5. Pour appliquer les modifications à toutes les tâches de profil d'objet de données et à la tâche de profil de clé étrangère générée dans le profil de découverte des données d'entreprise, sélectionnez **Utiliser les paramètres généraux pour les profils actuels**.

L'outil Developer met à jour toutes les tâches de profil d'objet de données et la tâche de profil de clé étrangère en fonction des paramètres modifiés.

6. Pour exécuter des tâches de profil individuelles, sélectionnez une tâche et cliquez sur **Exécuter**.
7. Pour exécuter plusieurs tâches de profil, cliquez sur **Exécuter plusieurs**.

La boîte de dialogue **Exécuter plusieurs** s'affiche.

Astuce: Si le chargement des résultats de la découverte d'entreprise est long, vous pouvez mettre à jour les statistiques de la base de données de l'entrepôt de profilage. Exécutée à plusieurs reprises, la découverte de profil des données d'entreprise peut entraîner d'importantes modifications du volume de données et des valeurs de colonne. Lorsque vous mettez à jour les statistiques, la base de données exécute un plan d'exécution pour les requêtes SQL basé sur les dernières statistiques et optimise les opérations de la base de données.

8. Toutes les tâches sont sélectionnées par défaut. Supprimez les tâches que vous ne voulez pas exécuter et cliquez sur **OK**.

Découverte de clé étrangère

Une colonne est une clé étrangère si ses valeurs de données correspondent aux valeurs de colonnes de clé primaire dans un autre objet de données.

Vous pouvez effectuer une découverte de clé étrangère dans des objets à plusieurs données dans l'outil Developer. Créez un profil de découverte des données d'entreprise pour sélectionner des objets de données et définir le profil.

Avant d'effectuer une découverte de clé étrangère, vous devez identifier les objets de données parent et enfant dans le profil de découverte des données d'entreprise. Le profil utilise une ou plusieurs clés dans l'objet parent, dont sa clé primaire, pour découvrir les clés étrangères dans l'objet enfant. Après avoir défini les objets parent et enfant, et identifié les clés dans l'objet parent, vous devez créer et exécuter le profil.

Définition des relations entre les objets parent et enfant

Pour rechercher des relations de clés étrangères entre deux objets de données, vous devez sélectionner un objet de données parent et indiquer la clé primaire dans cet objet.

1. Ouvrez un profil de découverte des données d'entreprise qui contient les objets de données que vous souhaitez analyser.
2. Sélectionnez l'objet parent.
3. Sélectionnez la clé primaire dans l'objet parent :
 - Cliquez sur l'onglet **Propriétés**, puis cliquez sur **Clés**.
 - Cliquez sur **Ajouter** et sélectionnez la colonne de clé primaire dans la boîte de dialogue Nouvelle clé.
 - Cliquez sur **OK** dans la boîte de dialogue **Nouvelle clé**. Vérifiez que la clé primaire est affichée dans le volet **Champs sélectionnés** et que l'option **Clé primaire** est cochée.

Créez un profil de clé étrangère pour analyser l'objet enfant pour les clés étrangères.

Découverte de relations de clés étrangères entre les objets de données

Utilisez un profil de découverte des données d'entreprise dans l'outil Developer pour rechercher des relations de clé entre deux objets de données

L'objet de données qui contient la clé primaire est l'objet parent, et l'objet de données qui contient la clé étrangère est l'objet enfant.

1. Ouvrez un profil de découverte des données d'entreprise qui contient les objets de données que vous souhaitez analyser.
2. Cliquez avec le bouton droit de la souris sur le nom d'un objet de données et sélectionnez **Profil de clé étrangère**.
3. Entrez un nom pour le profil et vérifiez l'emplacement du projet. Si nécessaire, accédez à un nouvel emplacement. Entrez éventuellement une description textuelle du profil.
4. Sélectionnez les clés dans l'objet parent que le profil utilise pour rechercher les clés étrangères dans l'objet enfant.
5. Enregistrez et exécutez le profil.

Résultats d'analyse de clé étrangère

Après avoir exécuté un profil de clé étrangère, cliquez sur le nom du profil au-dessous de l'éditeur de modélisation pour observer les résultats de l'analyse.

La vue de résultats indique les colonnes qui répondent aux critères d'inférence de clé primaire-étrangère que vous avez définis. Cliquez sur le bouton **Options** pour éditer les paramètres d'inférence. Cliquez sur un nom de colonne et sélectionnez **Valider** pour vérifier qu'une clé inférée est une clé valide pour les objets de données.

La table suivante décrit les propriétés d'analyse de clé étrangère :

Propriété	Description
Clé primaire parent	Une colonne de clé primaire dans l'objet de données parent que le profil utilise pour rechercher des clés étrangères dans un objet enfant.
Clé étrangère enfant	Une colonne que le profil considère comme une clé étrangère à la clé primaire parent dans cette ligne.
% d'inclusion	La quantité de valeurs de données qui correspondent entre la clé primaire et la clé étrangère, exprimée en pourcentage. Remarque: Il se peut que vous observiez un écart dans la valeur % d'inclusion pour une colonne inférée dans les résultats de clé étrangère et après validation. Pour une colonne inférée, le % d'inclusion est le nombre de valeurs uniques de colonnes de clé étrangère d'un objet enfant qui correspondent aux valeurs uniques de colonnes de clé primaire de l'objet parent. Après avoir validé une colonne inférée, c'est le nombre de valeurs de colonnes de clé étrangère d'un objet enfant qui correspondent aux valeurs de colonnes de clé primaire de l'objet parent.
Type de relation	Le type de relation défini pour les colonnes de clé primaire et étrangère avant d'exécuter le profil. Si vous définissez une relation avant d'exécuter le profil, ce dernier renvoie les données pour la relation même si le chiffre de pourcentage d'inclusion ne respecte pas le seuil de confiance défini pour le profil.
Vérifié	Indique qu'un utilisateur a validé la relation de clé primaire-étrangère.
Heure de la dernière exécution	La date et l'heure de la dernière exécution du profil.
Type de relation (dans le modèle)	Indique que le profil a vérifié la relation entre les colonnes.

Analyse de jointure

L'analyse de jointure décrit le degré de jointures potentielles entre deux colonnes de données. Utilisez un profil de jointure pour analyser des jointures de colonnes dans une ou plusieurs sources de données.

Un profil de jointure affiche des résultats sous la forme d'un diagramme de Venn et de valeurs numériques et de pourcentage. Vous créez et exécutez un profil de jointure à partir d'un profil de découverte des données d'entreprise.

Création d'un profil de jointure

Vous pouvez analyser des jointures potentielles entre les objets de données dans un profil de découverte des données d'entreprise. Le profil de jointure stocke l'analyse dans le référentiel modèle.

1. Créez ou ouvrez un profil de découverte des données d'entreprise.
2. Vérifiez que le profil de découverte des données d'entreprise contient les objets de données dont vous avez besoin.

Pour ajouter un objet de données au profil de jointure, faites-le glisser depuis la vue **Explorateur d'objets** vers l'éditeur de modélisation.

3. Sélectionnez les objets de données à profiler.
4. Cliquez avec le bouton droit de la souris sur les objets et sélectionnez **Profil de jointure**.
L'assistant Création de profils s'ouvre.
5. Entrez un nom pour le profil. Entrez éventuellement une description textuelle du profil.
6. Vérifiez que le nom des objets de données s'affiche dans **Objets de données** dans l'assistant.
7. Cochez ou désélectionnez l'option **Exécuter le profil à la fin**.
8. Cliquez sur **Suivant**.
9. Sélectionnez les colonnes de données à inclure dans le profil et cliquez sur **Suivant**.
Si nécessaire, faites défiler les objets de données pour afficher toutes les colonnes disponibles. Le profil est exécuté par défaut dans toutes les colonnes.
10. Cliquez sur **Ajouter**.
La boîte de dialogue **Condition de jointure** s'affiche.
11. Cliquez sur **Nouveau** pour activer les champs de sélection des colonnes.
12. Sélectionnez les objets de données et les colonnes à valider.
Vous devez définir une condition de jointure entre deux colonnes. Vous pouvez définir plusieurs conditions de jointure dans un ou plusieurs objets de données.
13. Cliquez sur **OK** pour créer la condition de jointure.
Cliquez éventuellement sur **Ajouter** pour définir d'autres conditions.
14. Vérifiez que les colonnes de jointure Gauche et Droite sont préfixées avec les noms d'objet de données corrects.
15. Cliquez sur **Terminer**.

Résultats d'analyse de jointure

L'onglet **Résultats** d'analyse de jointure contient des informations sur le nombre et le pourcentage de lignes orphelines parents, de lignes orphelines enfants et de lignes de jointure. Les résultats de l'analyse de jointure comprennent également les diagrammes de Venn qui indiquent les relations entre les colonnes.

La table suivante décrit les propriétés indiquées dans l'onglet **Résultats** :

Propriété	Description
Table de gauche	Nom de la table de gauche et colonnes utilisées dans l'analyse de jointure.
Table de droite	Nom de la table de droite et colonnes utilisées dans l'analyse de jointure.
Lignes de gauche seulement	Le nombre de lignes dans la table de gauche qui ne peuvent pas être jointes.
Lignes de droite seulement	Le nombre de lignes dans la table de droite qui ne peuvent pas être jointes.
Lignes de jointure	Le nombre de lignes comprises dans la jointure.

Sélectionnez une condition de jointure pour afficher un diagramme de Venn qui indique les relations entre les colonnes. La zone au-dessous du diagramme de Venn affiche également le nombre et pourcentage de valeurs orphelines, Null et jointes dans les colonnes.

Double-cliquez sur une section dans le diagramme de Venn pour afficher les enregistrements que la section représente. Ces enregistrements s'ouvrent dans la vue Visionneuse de données.

Remarque: Vous pouvez exporter la liste des enregistrements depuis la vue Visionneuse de données vers un fichier plat.

Exportation des résultats de profil de jointure dans un fichier

Vous pouvez exporter les lignes de données renvoyées pour une condition de jointure dans un fichier délimité. Exportez les lignes superposées entre les sources gauche ou droite ou les lignes orphelines dans une source.

1. Dans la vue **Explorateur d'objets**, ouvrez le profil de découverte des données d'entreprise qui contient l'analyse de jointure.
2. Exécutez le profil de jointure.
3. Sélectionnez la vue **Résultats de jointure**.
4. Dans l'onglet **Visionneuse de données**, cliquez sur l'icône **Exporter les résultats du développement dans un fichier**.
La boîte de dialogue **Exporter les données** s'affiche.
5. Entrez un nom de fichier et cliquez sur **Enregistrer**.

Découverte du chevauchement

La découverte de chevauchement fournit des informations sur le chevauchement des données entre des paires de colonnes d'une source de données ou de plusieurs sources de données. Vous pouvez rechercher des données de chevauchement à partir d'un profil de découverte des données d'entreprise. Vous pouvez valider les résultats de profils et les afficher dans un diagramme de Venn.

La découverte de chevauchement identifie le chevauchement des données selon les paramètres par défaut ou les paramètres que vous indiquez. Vous pouvez remplacer les paramètres par défaut et spécifier des options d'inférence, notamment le nombre maximum de paires que la découverte de chevauchement renvoie en fonction du pourcentage de chevauchement. Vous pouvez également indiquer un niveau de confiance qui détermine l'éligibilité de la découverte de chevauchement.

Résultats de découverte du chevauchement

L'onglet **Découverte du chevauchement** affiche des informations sur les colonnes concernées et la valeur du pourcentage de chevauchement. Les résultats de découverte de chevauchement incluent les diagrammes de Venn, qui représentent le chevauchement des données dans des paires de colonnes, et la date et heure de la dernière exécution de la découverte de chevauchement.

Vous pouvez cliquer sur une colonne et sélectionner **Vérifier** pour afficher les résultats sous la forme d'un diagramme Venn.

Le tableau suivant décrit les propriétés de découverte de chevauchement :

Propriété	Description
Colonne de gauche	Colonne principale à laquelle les colonnes restantes sont comparées pour l'analyse du chevauchement.
Colonne de droite	Colonne comparée à la colonne principale.

Propriété	Description
% du chevauchement	Pourcentage de chevauchement entre deux colonnes.
Vérifié	Indique que vous avez validé la ligne de résultats de chevauchement.
Heure de la dernière exécution	Date et heure de la dernière exécution de la découverte de chevauchement.

Informatica Developer affiche chaque paire de chevauchement deux fois dans les résultats de découverte du chevauchement. Examinez les éléments de sources de données et les commandes. Les éléments disposent des colonnes « m » et « n » et les commandes, des colonnes « p » et « q ».

Le tableau suivant affiche les résultats de la découverte du chevauchement des éléments et des commandes :

Colonne de gauche	Colonne de droite
Éléments	-
m	Orders.p
m	Orders.q
n	Orders.p
n	Orders.q
Commandes	-
p	Items.m
p	Items.n
q	Items.m
q	Items.m

Découverte de données de chevauchement

Vous pouvez déterminer les données se chevauchant entre des paires de colonnes dans un profil de découverte des données d'entreprise. L'analyse des chevauchements est basée sur des valeurs uniques dans les colonnes et ne prend pas les valeurs nulles en considération.

1. Créez ou ouvrez un profil de découverte des données d'entreprise qui contient les objets de données.
2. Sélectionnez les objets de données sur lesquels vous souhaitez trouver des données de chevauchement.
Vous pouvez sélectionner un objet de données unique pour trouver des données se chevauchant dans des paires de colonnes ou dans plusieurs objets de données.
3. Cliquez avec le bouton droit de la souris sur les objets et sélectionnez **Découverte du chevauchement**.
La boîte de dialogue **Nouvelle découverte du chevauchement** apparaît.
4. Entrez un nom.
5. Saisissez éventuellement une description textuelle pour l'analyse de chevauchement.

6. Vérifiez que le nom des objets de données s'affiche dans **Objets de données** dans l'assistant.
7. Sélectionnez éventuellement **Exécuter le profil à la fin** pour exécuter le profil une fois la configuration des paramètres terminée.
8. Cliquez sur **Suivant**.
9. Sélectionnez les colonnes pour la découverte du chevauchement.
10. Cliquez sur **Suivant**.
Les options d'inférence par défaut apparaissent dans la boîte de dialogue.
11. Spécifiez éventuellement les options d'inférence pour la découverte du chevauchement afin de remplacer les paramètres par défaut.
12. Cliquez sur **Terminer**.

Fichiers de script DDL

Les fichiers de script du langage de définition de données (DDL) contiennent les instructions SQL `Créer`, `Modifier` et `Abandonner`.

Vous pouvez spécifier un nom de fichier, un emplacement ainsi qu'un type de base de données cible lorsque vous générez les fichiers de script. L'outil Developer ajoute les libellés « `_créer` » et « `_abandonner` » aux noms des fichiers de script. Les colonnes virtuelles ne font pas partie des fichiers de script DDL.

Création de scripts DDL à partir d'un profil de découverte des données d'entreprise

Lorsque vous générez des fichiers de script DDL à partir d'un profil de découverte des données d'entreprise, vous pouvez choisir l'emplacement auquel enregistrer ces fichiers. Vous pouvez également choisir le type de base de données sur lesquelles vous souhaitez exécuter les scripts. Veillez à vérifier et valider toutes les modifications nécessaires dans le profil de découverte des données d'entreprise avant de générer les scripts DDL.

1. Dans la vue **Explorateur d'objets**, sélectionnez un profil de découverte des données d'entreprise.
2. Cliquez avec le bouton droit de la souris sur le profil et sélectionnez **Générer DDL**.
La boîte de dialogue **Générer DDL** s'ouvre.
3. Cliquez sur **Parcourir** pour ouvrir la boîte de dialogue **Enregistrer sous**.
L'extension du fichier par défaut est `.sql`.
4. Choisissez l'emplacement du fichier et entrez un nom de fichier.
5. Sélectionnez le type de base de données cible.
6. Cliquez sur **OK**.

L'outil Developer génère les fichiers de script DDL à l'emplacement spécifié.

CHAPITRE 23

Résultats de la découverte des données d'entreprise

Ce chapitre comprend les rubriques suivantes :

- [Présentation des résultats de la découverte des données d'entreprise, 161](#)
- [Vue Relations, 162](#)
- [Vue Profilage de clé étrangère, 163](#)
- [Vue tabulaire, 166](#)
- [Vue Domaines de données, 167](#)
- [Vue Profil de colonne, 169](#)
- [Affichage des résultats de profil de colonne lors de l'exécution de la découverte des données d'entreprise, 170](#)
- [Affichage des résultats de la découverte des domaines de données lors de l'exécution de la découverte des données d'entreprise, 170](#)
- [Affichage du statut d'exécution de la découverte des données d'entreprise, 170](#)
- [Fichiers d'exportation de la découverte des données d'entreprise, 171](#)

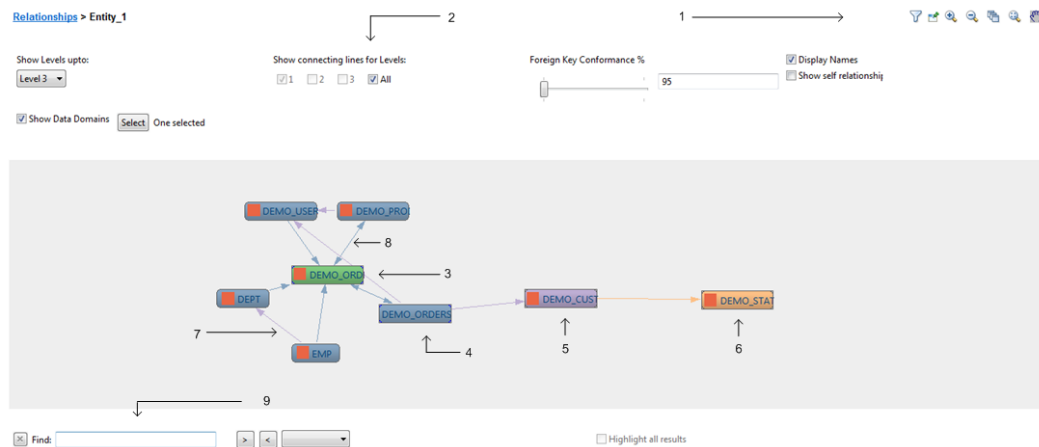
Présentation des résultats de la découverte des données d'entreprise

Vous pouvez afficher les résultats de la découverte des données d'entreprise dans plusieurs vues. Les vues sont **Relations**, **Domaines de données**, **Profil de colonne**, **Profil de jointure** et **Découverte du chevauchement**.

La vue **Relations** affiche les groupes d'objets de données sous forme de cercles. Vous pouvez lancer les résultats du profil de clé étrangère à partir de cette vue. Vous pouvez afficher les résultats du profil de clé étrangère dans les vues graphique et tabulaire. La vue **Domaines de données** affiche les résultats de la découverte des domaines de données. La vue **Profil de colonne** affiche les résultats de profil de colonne de chaque objet de données. La vue **Profil de jointure** affiche le nombre de lignes orphelines parents, de lignes orphelines enfants et de lignes comprises dans la jointure. La vue **Découverte du chevauchement** affiche des informations sur les colonnes concernées et la valeur du pourcentage de chevauchement.

Les objets de données peuvent avoir plusieurs relations entre elles. La vue graphique des résultats de clé étrangère affiche les relations d'objet de données présentant le pourcentage de conformité le plus élevé.

La figure suivante montre la représentation graphique de quelques exemples d'objets de données des résultats de la découverte des données d'entreprise :



1. Icônes de la barre d'outils incluant les icônes Filtre, Épingler l'objet de données, Zoom avant et Zoom arrière, Arranger tout, Ajuster à la fenêtre et Panoramique.
2. Options de filtre, telles que l'affichage de différents niveaux de relations d'objet de données, l'affichage des lignes de connexion et l'affichage des objets de données liés à eux-mêmes.
3. Objet de données sélectionné en fonction duquel l'éditeur visuel affiche le reste des relations entre les objets de données.
4. Premier niveau des relations de l'objet de données.
5. Deuxième niveau des relations de l'objet de données.
6. Troisième niveau des relations de l'objet de données.
7. Connecteur entre les objets de données. Une seule tête de flèche indique une relation clé principale à clé étrangère entre deux objets de données. La tête de flèche pointe vers l'objet de données avec la clé principale.
8. Connecteur entre les objets de données. Un connecteur à double tête de flèche indique une relation clé principale à clé principale entre deux objets de données. Survolez le connecteur pour afficher les colonnes de jointure avec la conformité maximum pour la relation inférée.
9. Appuyez sur CTRL + F pour afficher le champ Rechercher et utilisez un astérisque (*) en tant que caractère de remplacement pour rechercher des objets de données dans la vue graphique.

Vue Relations

Vous pouvez afficher un récapitulatif des résultats de découverte des données d'entreprise dans la vue **Relations**, y compris les entités. Les entités sont des groupes d'objets de données représentés sous forme de cercles. Les entités incluent les objets de données liés, liés à eux-mêmes et non liés des nombreux connexions et schémas des bases de données source.

Un objet de données lié à lui-même présente des colonnes dans l'objet de données qui possèdent des relations. Un objet de données non lié ne possède pas de relation avec les autres objets de données de la base de données source, ni de relation entre les colonnes dans l'objet de données. Le diagramme de relation des entités des objets de données dans les résultats de la découverte des données d'entreprise est basé sur les relations inférées et non les relations documentées dans les sources de données.

Rechercher un objet de données

Vous pouvez rechercher un objet de données dans la vue **Relations** ou la vue **Profilage de clé étrangère**. Vous pouvez utiliser un astérisque (*) comme caractère générique pour rechercher les objets de données.

1. Vérifiez que vous êtes dans la vue **Relations** ou **Profilage de clé étrangère**.

2. Entrez une partie du nom de l'objet de données que vous recherchez et ajoutez le caractère générique * au début ou à la fin de la chaîne de recherche en fonction des besoins de la recherche. Par exemple, pour rechercher tous les objets de données commençant par la chaîne « CA », entrez « CA* » et appuyez sur la touche **Entrée**. Pour rechercher tous les objets de données contenant la chaîne « CODE » dans leurs noms, entrez « *CODE* ».

La recherche est sensible à la casse.

Navigation dans la vue Profilage de clé étrangère

La vue **Profilage de clé étrangère** affiche une vue consolidée des relations de clé étrangère dans plusieurs objets de données sur lesquels vous avez exécuté le profil. Les cercles de la vue représentent les entités, les objets liés à eux-mêmes et les objets non référencés.

1. Vérifiez que vous êtes dans l'onglet **Relations**.
Vous pouvez afficher le lien du profil de clé étrangère dans le volet de droite.
2. Cliquez sur **Profil de clé étrangère** pour ouvrir la vue.
La vue s'affiche dans un nouvel onglet. La vue s'affiche dans les objets de données dans différents ensembles de cercles en fonction du type de relation. Vous pouvez également voir le nombre total d'objets de données faisant partie de la vue de clé étrangère consolidée.
3. Vous pouvez éventuellement cliquer sur le lien **Relations** pour revenir à la vue **Relations**.

Vue Profilage de clé étrangère

Vous pouvez afficher un résumé des résultats de la découverte des données d'entreprise dans un format graphique dans la vue **Profilage de clé étrangère**. Vous pouvez ouvrir les résultats de profil et les relations au niveau des colonnes pour un objet de données dans un format tabulaire à partir de la vue.

Les résultats de profil pour un objet de données incluent les résultats de profil de colonne, d'inférence de clé primaire, d'inférence de dépendance fonctionnelle et de découverte de domaines de données. Lorsque vous ouvrez les relations au niveau des colonnes pour un objet de données, vous pouvez vérifier et conserver les relations de données. Lorsque vous vérifiez une relation de données, l'outil Developer exécute le profil sur toutes les lignes de la source de données pour vérifier les résultats d'inférence. Vous pouvez approuver, rejeter et réinitialiser des relations de données dans la vue **Profilage de clé étrangère**.

Affichage des relations d'objets de données

Vous pouvez afficher les relations entre les objets de données au format graphique. Double-cliquez sur un cercle d'entité pour afficher les tables et leurs relations.

1. Vérifiez que vous êtes dans la vue **Profilage de clé étrangère**.
2. Pour inclure les domaines de données dans la vue graphique consolidée, sélectionnez **Afficher les domaines de données**.
Le bouton **Sélectionner** est activé.
3. Cliquez sur **Sélectionner** pour choisir les domaines de données que vous voulez inclure dans la vue graphique.
La boîte de dialogue **Sélectionner des domaines de données** s'affiche.

4. Sélectionnez les domaines de données requis, puis cliquez sur **OK**.
L'outil Developer met en surbrillance les cercles d'entité qui incluent les domaines de données que vous avez sélectionnés.
5. Double-cliquez sur un cercle d'entité pour afficher la représentation visuelle des relations de table dans l'entité. L'outil Developer affiche les tables dans un format graphique qui représente la relation que chaque objet de données possède avec les autres objets de données dans l'entité.
L'objet de données ayant le nombre maximum de relations avec d'autres objets de données ou l'objet de données depuis lequel vous démarrez la navigation est mis en surbrillance en vert. Si vous avez inclus des domaines de données, l'outil Developer met en surbrillance la sélection de domaine de données à gauche de la représentation visuelle de chaque objet de données.
6. Vérifiez les informations de relation directe et de domaine de données dans le volet de droite de la vue **Profilage de clé étrangère**.
7. Vous pouvez éventuellement cliquer sur le lien **Relations** pour revenir à la vue **Relations**.

Zoom avant et zoom arrière dans la vue

Vous pouvez zoomer dans la représentation graphique des relations d'objet de données dans la vue **Profilage de clé étrangère** pour plus de clarté. Lorsque vous effectuez un zoom avant, l'outil Developer augmente le niveau d'agrandissement de l'image. Faites un zoom arrière pour réduire le niveau d'agrandissement.

1. Vérifiez que vous êtes dans la vue **Profilage de clé étrangère**.
2. Effectuez un clic droit sur la vue et sélectionnez **Zoom avant** pour augmenter le niveau d'agrandissement de l'image.
3. Pour réduire le niveau d'agrandissement de la présentation graphique, cliquez avec le bouton droit de la souris sur la vue et sélectionnez **Zoom arrière**.

Recherche d'un objet de données

Vous pouvez rechercher et trouver un objet de données dans la représentation graphique des résultats de clé étrangère. Utilisez un astérisque (*) comme caractère générique pour rechercher des objets de données.

1. Vérifiez que vous êtes dans la vue **Profilage de clé étrangère**.
Assurez-vous d'ouvrir la représentation graphique des résultats de clé étrangère dans la vue.
2. Appuyez sur CTRL + F pour afficher le champ **Trouver**.
3. Dans le champ **Trouver**, entrez une partie du nom de l'objet de données que vous recherchez et ajoutez le caractère générique * au début ou à la fin de la chaîne de recherche en fonction de la condition de recherche. Par exemple, pour rechercher tous les objets de données commençant par la chaîne « EMP », entrez « EMP* » et appuyez sur la touche **Entrée**. Pour rechercher tous les objets de données contenant la chaîne « CODE » dans leurs noms, entrez « *CODE* ».
4. Cliquez sur le bouton **Correspondance suivante** pour vous déplacer vers l'objet de données correspondant suivant.
Utilisez le bouton **Correspondance précédente** pour vous déplacer vers l'objet de données correspondant précédent.
5. Sélectionnez **Mettre en surbrillance tous les résultats** pour mettre en surbrillance toutes les correspondances de l'objet de données.
6. Pour effacer la chaîne de recherche du champ **Trouver**, cliquez sur le bouton **Effacer** en regard du champ.

Affichage des relations de colonne

Vous pouvez afficher les relations de chaque colonne d'un objet de données avec les colonnes des objets de données liés. Vous pouvez également vérifier et valider la relation de l'objet de données pour le modèle de données.

1. Vérifiez que vous êtes dans la vue **Profilage de clé étrangère**.
2. Cliquez avec le bouton droit de la souris sur un objet de données et sélectionnez **Afficher les relations de colonne**.

Les relations de la colonne s'affichent dans une vue tabulaire. La vue affiche les informations de relation, par exemple l'objet de données source, l'objet de données associé et les colonnes qu'il contient.

3. Vérifiez le statut de l'inférence, le statut de vérification et le statut de conservation.
4. Sélectionnez **Tous les objets de données dans le groupe** pour afficher tous les objets de données dans l'entité parent et leurs informations de relation de colonne.

Par défaut, la vue affiche les informations de relation pour l'objet de données sélectionné dans la vue.

5. Vous pouvez également cliquer sur le lien **entité** en haut de la vue pour revenir à la représentation graphique des objets de données.

Enregistrement du diagramme de la relation de l'entité sous la forme d'une image

Vous pouvez enregistrer le diagramme de la relation de l'entité des objets de données depuis les résultats de la découverte des données d'entreprise au format ".png".

1. Exécutez un profil pour effectuer la découverte des données d'entreprise.
2. Vérifiez que vous êtes dans la vue **Profilage de clé étrangère**.
3. Passer au format graphique des relations de l'objet de données depuis la vue.
4. Effectuez un clic droit et sélectionnez **Enregistrer en tant qu'image**.

La boîte de dialogue **Enregistrer sous** s'ouvre. Enregistrez l'image au format ".png" par défaut.

5. Choisissez l'emplacement du fichier et entrez un nom de fichier.
6. Cliquez sur **Enregistrer**.

Affichage des résultats de profil d'objet de données depuis la vue Profilage de clé étrangère

Vous pouvez afficher les résultats de profil de colonne, de clé principale et de découverte des domaines de données d'un objet de données sélectionné depuis la vue **Profilage de clé étrangère**. Vérifiez que vous avez choisi l'objet de données en l'épinglant comme table sélectionnée dans le canevas.

1. Vérifiez que vous êtes dans la vue **Profilage de clé étrangère**.
2. Cliquez avec le bouton droit de la souris sur un objet de données et sélectionnez **Épingler l'objet de données comme cible** pour sélectionner la table.

Vous pouvez également utiliser l'icône **Épingle** pour sélectionner un objet de données.

3. Cliquez avec le bouton droit de la souris dans le canevas et sélectionnez **Afficher le profil d'objet de données**.

Les résultats du profil d'objet de données s'affichent dans un onglet.

Vue tabulaire

Lorsque vous ouvrez la vue **Profilage de clé étrangère**, l'outil Developer affiche par défaut une vue graphique des résultats. Passez en vue tabulaire pour afficher les tables et les détails de leurs relations au format tabulaire.

Vous pouvez afficher le nombre d'objets de données dans l'entité, les noms des tables associées, les informations de connexion et le nombre de relations entre les deux objets de données. Vous pouvez également vérifier les relations de colonne et les ajouter à un modèle de données.

Volet Détails de la table

Vous pouvez afficher les détails de l'objet de données dans les vues graphique et tabulaire des résultats de la découverte des données d'entreprise. Dans la vue graphique, le volet des détails de la table affiche le nombre d'objets de données qui possèdent une relation directe avec l'objet de données sélectionné et les noms d'objet de données.

Le tableau suivant décrit les colonnes du volet des détails de la table dans la vue tabulaire :

Nom de colonne	Description
Nom de la table	Nom de l'objet de données qui possède une relation directe avec l'objet de données sélectionné dans le volet de gauche.
Connexion	Nom de la connexion pour l'objet de données associé.
Relations	Nombre de relations entre l'objet de données dans le volet de gauche et l'objet de données associé dans le volet des détails de la table.

Vérification des résultats de la découverte des données d'entreprise

Lorsque vous vérifiez les résultats de la découverte des données d'entreprise, l'outil Developer exécute le profil sur toutes les lignes de la source de données. La valeur du pourcentage de conformité peut varier après la vérification en fonction des valeurs de colonne sur toutes les lignes de la source de données.

1. Ouvrez un profil après l'avoir exécuté.
2. Vérifiez que vous êtes dans la vue **Profilage de clé étrangère**.
3. Cliquez sur l'icône **Vue tabulaire** en haut de la vue.
La vue tabulaire affiche les entités dans le volet de gauche.
4. Cliquez avec le bouton droit de la souris sur un objet de données dans le volet de gauche et sélectionnez **Afficher les relations de colonne**.
Vous pouvez afficher les relations des colonnes de l'objet de données sélectionné avec les colonnes des autres objets de données. Déplacez-vous vers la droite pour afficher les détails comme le type de relation, le pourcentage de conformité, l'état de la vérification et le statut de la validation.
5. Cliquez avec le bouton droit de la souris sur une ligne et sélectionnez **Vérifier**.
La boîte de dialogue **Exécuter profil** s'affiche. Une fois la vérification terminée, sélectionnez la ligne pour afficher le chevauchement des relations de clé principale et de clé étrangère dans un diagramme de Venn.

Conservation des relations de colonne dans l'outil Developer

Vous pouvez approuver, rejeter et réinitialiser des relations de données dans la vue **Profilage de clé étrangère**.

1. Ouvrez un profil après l'avoir exécuté.
2. Vérifiez que vous êtes dans la vue **Profilage de clé étrangère**.
3. Dans la vue tabulaire, cliquez avec le bouton droit de la souris sur un objet de données et sélectionnez **Afficher les relations de colonne**.
4. Sélectionnez la relation de l'objet de données que vous voulez conserver.
5. Pour approuver la relation de colonne, cliquez avec le bouton droit de la souris et cliquez sur **Approuver**.
Le statut de la ligne devient **Approuvé**.
6. Pour restaurer le statut inféré de la relation de colonne, cliquez avec le bouton droit de la souris et cliquez sur **Réinitialiser**.
7. Pour rejeter une relation de colonne inférée, sélectionnez la ligne, cliquez avec le bouton droit de la souris et sélectionnez **Rejeter**.
8. Pour afficher les relations de colonne rejetées, cliquez avec le bouton droit de la souris sur l'une des lignes, puis sélectionnez **Afficher les éléments rejetés**.
9. Pour masquer les types de données rejetés, cliquez avec le bouton droit de la souris sur l'une des lignes, puis sélectionnez **Masquer les éléments rejetés**.

Validation des résultats dans le référentiel modèle

Après avoir exécuté un profil, vous pouvez enregistrer les relations de colonne entre les objets de données dans le référentiel modèle. Vous pouvez valider les relations dans le référentiel modèle à partir de la vue tabulaire de la vue **Profilage de clé étrangère**.

1. Ouvrez un profil après l'avoir exécuté.
2. Vérifiez que vous êtes dans la vue **Profilage de clé étrangère**.
3. Cliquez avec le bouton droit de la souris sur un objet de données dans le volet de gauche et sélectionnez **Afficher les relations de colonne**.
Vous pouvez afficher les relations des colonnes de l'objet de données sélectionné avec les colonnes des autres objets de données.
4. Cliquez sur une ligne avec le bouton droit de la souris et sélectionnez **Approuver**.

Vue Domaines de données

La vue **Domaines de données** répertorie les domaines de données et les statistiques de correspondance de colonne découvertes par l'outil Developer dans le cadre de la découverte des données d'entreprise. Vous

pouvez vérifier les colonnes, développer les lignes et afficher les résultats du profil d'objet de données depuis la vue **Domaines de données**.

Affichage des résultats de la découverte des domaines de données

Vous pouvez afficher les résultats de découverte des domaines de données dans l'onglet **Domaines de données**. Vous pouvez rechercher les domaines de données et afficher les domaines de données triés par groupes de domaines de données.

1. Exécutez un profil pour effectuer la découverte des données d'entreprise.
2. Ouvrez le profil.
3. Cliquez sur l'onglet **Domaines de données** pour afficher les résultats de la découverte des domaines de données.

Les résultats du profil d'objet de données s'affichent dans le volet de droite.

4. Entrez une partie d'un nom de domaine de données dans le champ de recherche pour trouver un domaine de données spécifique.

Sélectionnez **Afficher la hiérarchie du groupe de domaines de données** pour trier la liste des domaines de données par groupes de domaines de données.

Vérification des résultats de la découverte des domaines de données

Lorsque vous exécutez un profil, il analyse un échantillon de la source de données pour inférer les résultats de profil. Vous pouvez exécuter le profil sur toutes les lignes de la source de données pour vérifier les résultats d'inférence.

1. Ouvrez un profil après l'avoir exécuté.
2. Cliquez sur l'onglet **Domaines de données** pour afficher les résultats.

Vous pouvez afficher les résultats de la découverte des domaines de données dans le panneau de droite.

3. Sélectionnez une colonne que vous voulez vérifier dans le panneau de droite.
4. Cliquez avec le bouton droit de la souris sur la colonne et cliquez sur **Vérifier** pour exécuter le profil sur toutes les lignes de la source de données.

Vous pouvez voir une modification de la valeur **% de conformité des données** après avoir vérifié les résultats.

Développement des lignes

Vous pouvez développer les résultats de la découverte des domaines de données pour une analyse approfondie des données.

1. Ouvrez un profil après l'avoir exécuté.
2. Cliquez sur l'onglet **Domaines de données** pour afficher les résultats.

Vous pouvez afficher les résultats de la découverte des domaines de données dans le panneau de droite.

3. Sélectionnez une ligne que vous voulez développer dans le panneau de droite.

4. Cliquez avec le bouton droit de la souris sur la colonne et sélectionnez **Développer** pour développer les lignes de la source.

Affichage des résultats de profil d'objet de données depuis la vue Domaines de données

Vous pouvez afficher les résultats du profil d'objet de données d'un objet de données sélectionné dans la vue **Domaines de données**.

1. Vérifiez que vous êtes dans la vue **Domaines de données**.
2. Sélectionnez un domaine de données dans le volet **Domaines profilés**.
3. Dans le volet **Colonnes** à droite, sélectionnez une colonne.
4. Cliquez avec le bouton droit de la souris sur une colonne et sélectionnez **Ouvrir un profil d'objet de données**.

Les résultats du profil d'objet de données s'affichent dans un onglet.

Vue Profil de colonne

La vue **Profil de colonne** affiche un résumé des résultats de profil de colonne pour les profils d'objet de données uniques que l'outil Developer exécute comme élément de la découverte des données d'entreprise. Vous pouvez afficher les statistiques de colonne comme les valeurs uniques, les valeurs null, les types de données et les valeurs minimum et maximum pour chaque colonne dans les objets de données.

Affichage des résultats de profil d'objet de données

La découverte des données d'entreprise inclut l'exécution d'un profil d'objet de données pour découvrir les statistiques de données de colonne, les clés principales et candidates ainsi que les domaines de données. Vous pouvez afficher les résultats du profil d'objet de données d'un objet de données sélectionné dans la vue **Profil de colonne**.

1. Vérifiez que vous êtes dans la vue **Profil de colonne**.
2. Sélectionnez un objet de données dans le volet **Objets de données profilés**.
3. Dans le volet **Colonnes** à droite, sélectionnez une colonne.
4. Cliquez avec le bouton droit de la souris sur une colonne et sélectionnez **Afficher le profil d'objet de données**.

Les résultats du profil d'objet de données s'affichent dans un onglet. Les résultats du profil de colonne s'affichent par défaut.

5. Cliquez sur **Inférence de clé principale** pour afficher les résultats du profil de clé principale.
6. Cliquez sur **Inférence de dépendance fonctionnelle** pour afficher les résultats de la découverte de dépendance fonctionnelle.
7. Cliquez sur **Découverte des domaines de données** pour afficher les résultats de la découverte des domaines de données.

Affichage des résultats de profil de colonne lors de l'exécution de la découverte des données d'entreprise

La durée nécessaire pour exécuter la découverte des données d'entreprise dépend du nombre de tâches de profil, de la taille de la source de données et du type de profil. Lorsque l'outil Developer poursuit l'exécution des tâches de découverte des données, vous pouvez afficher les résultats des profils de colonne qu'il effectue dans les phases initiales de la découverte des données.

1. Après avoir exécuté le profil, cliquez sur **Profils** dans la fenêtre **Propriétés**.
2. Sélectionnez le profil de colonne pour lequel vous voulez afficher les résultats. Vérifiez que le statut de l'exécution du profil est **Succès** dans la fenêtre **Propriétés**.
3. Cliquez sur **Ouvrir** pour afficher les résultats dans un autre onglet.
4. Dans la section **Résultats**, sélectionnez **Profilage de colonne** pour afficher les résultats dans le volet de droite.

Affichage des résultats de la découverte des domaines de données lors de l'exécution de la découverte des données d'entreprise

Lorsque l'outil Developer poursuit l'exécution des tâches de découverte des données, vous pouvez afficher les résultats de la découverte des domaines de données qu'il effectue dans les phases initiales de la découverte des données d'entreprise.

1. Après avoir commencé l'exécution du profil, cliquez sur **Profils** dans la fenêtre **Propriétés**.
2. Sélectionnez le profil pour lequel vous voulez afficher les résultats des domaines de données. Vérifiez que le statut de l'exécution du profil est **Succès** dans le volet **Propriétés**.
3. Cliquez sur **Ouvrir** pour afficher les résultats dans un autre onglet.
4. Dans la section **Résultats**, sélectionnez **Profilage de découverte de domaines de données** pour afficher les résultats dans le volet de droite.

Affichage du statut d'exécution de la découverte des données d'entreprise

La vue **Progression** de l'outil Developer affiche la progression des opérations comme l'exécution d'un profil. Vous pouvez afficher le statut d'exécution des tâches de découverte des données d'entreprise depuis la vue **Progression**.

1. Après avoir exécuté un profil pour effectuer la découverte des données d'entreprise sur les sources de données, cliquez sur le bouton **Vue Progression** dans l'angle inférieur droit de l'outil Developer.
Le volet **Progression** s'affiche, s'il n'est pas déjà ouvert.

2. Cliquez sur le lien **Découverte des données d'entreprise en cours d'exécution : afficher le lien Statut des tâches** pour ouvrir la boîte de dialogue des sous-tâches.
La boîte de dialogue répertorie les tâches de profil faisant partie de la découverte des données d'entreprise. Vous pouvez afficher le nom, le type et le statut du profil.
3. Cliquez sur l'en-tête de colonne pour trier les tâches de profil. Par exemple, pour trier les tâches de profil par statut, cliquez sur l'en-tête de colonne **Statut**.
4. Si vous devez annuler une tâche de profil spécifique, sélectionnez-la et cliquez sur **Annuler**.
Le statut de la tâche annulée est remplacé par **Terminée**.

Fichiers d'exportation de la découverte des données d'entreprise

Après avoir exécuté un profil de découverte des données d'entreprise, vous pouvez exporter les informations, notamment toutes les relations de l'objet de données, les domaines de données et les résultats de chaque tâche de clé étrangère. Vous pouvez enregistrer l'image graphique des relations de l'objet de données au format `.jpg`.

Lorsque vous exportez les résultats de profil, l'outil Developer enregistre tous les résultats de la découverte des données d'entreprise dans plusieurs fichiers Microsoft Excel. Vous pouvez afficher les relations d'objet de données, les résultats de profil de colonne, les résultats de la découverte des domaines de données, les entités ainsi que les résultats individuels des tâches de clé étrangère dans des fichiers séparés.

Exportation des résultats de la découverte des données d'entreprise

Vous pouvez exporter la liste des entités et la composition de chaque entité, tous les objets de données et les relations de l'objet de données au niveau colonne d'une entité, les domaines de données ainsi que les résultats de profilage de colonne.

1. Exécutez un profil pour effectuer la découverte des données d'entreprise.
2. Dans les vues **Relations**, **Domaines de données** ou **Profil de colonne**, cliquez sur l'icône **Exporter** dans la zone supérieure droite de la fenêtre.
La boîte de dialogue **Exporter les données vers un fichier** s'ouvre.
3. Entrez le nom de fichier. Utilisez éventuellement le nom de fichier par défaut.
4. Dans **Enregistrer**, choisissez **Enregistrer sur le client** et cliquez sur **Parcourir** pour sélectionner un emplacement et enregistrer le fichier localement sur votre ordinateur. Par défaut, Informatica Developer enregistre le fichier dans un emplacement défini dans les propriétés du service d'intégration de données d'Informatica Administrator.
5. Cliquez sur **OK**.

CHAPITRE 24

Bureau de glossaire métier dans Informatica Developer

Ce chapitre comprend les rubriques suivantes :

- [Recherche dans le glossaire métier, 172](#)
- [Recherche d'un terme métier, 173](#)
- [Personnalisation des touches de raccourci pour rechercher un terme métier, 173](#)

Recherche dans le glossaire métier

Recherchez la signification d'un nom d'objet de l'outil Developer en tant que terme métier dans le Bureau de glossaire métier pour comprendre sa spécification dans l'entreprise et son implémentation actuelle.

Un glossaire métier est un ensemble de termes qui utilisent un langage métier spécifique pour définir des concepts pour les utilisateurs de l'entreprise. Un terme métier fournit la définition et l'utilisation par l'entreprise d'un concept. Le Bureau de glossaire métier est un client qui se connecte au service Metadata Manager, qui héberge le glossaire métier. Utilisez le Bureau de glossaire métier pour rechercher des termes métier dans un glossaire métier.

Si le Bureau de glossaire métier est installé sur votre machine, vous pouvez sélectionner un objet dans l'outil Developer et utiliser des touches de raccourci ou le menu Rechercher pour rechercher le nom de l'objet dans le glossaire métier. Vous pouvez rechercher le nom des objets qui se trouvent dans les vues de l'outil Developer telles que la vue **Explorateur d'objets** ou un éditeur, des noms de colonnes, des profils, et des ports de mappage dans l'éditeur.

Par exemple, un développeur souhaite rechercher dans un glossaire métier un terme métier qui correspond à l'objet de données Sales_Audit dans l'outil Developer. Le développeur souhaite afficher les détails du terme métier pour comprendre les spécifications d'entreprise et l'implémentation actuelle de l'objet Sales_Audit dans l'outil Developer. Cela peut aider le développeur à comprendre la signification de l'objet de données et les modifications devant éventuellement être implémentées sur l'objet.

Recherche d'un terme métier

Recherchez un nom d'objet Developer en tant que terme métier dans le Bureau de glossaire métier pour comprendre sa spécification d'entreprise et son implémentation actuelle.

Le Bureau de glossaire métier doit être installé sur votre machine

1. Sélectionnez un objet.
 2. Choisissez d'utiliser les touches de raccourci ou le menu Rechercher pour ouvrir le Bureau de glossaire métier.
 - Pour utiliser les touches de raccourci, utilisez la combinaison de touches de raccourci suivante :
`CTRL+Shift+F`
 - Pour utiliser le menu Rechercher, cliquez sur **Rechercher > Glossaire Métier**.
- Le **Bureau de glossaire métier** apparaît et affiche les termes métier qui correspondent au nom de l'objet.

Personnalisation des touches de raccourci pour rechercher un terme métier

Personnaliser les touches de raccourci pour changer la combinaison de touches qui ouvrent le Bureau de glossaire métier.

1. Dans le menu de l'outil Developer, cliquez sur **Fenêtre > Préférences > Général > Touches**.
2. Pour trouver ou rechercher **Rechercher dans le glossaire métier** dans la liste des commandes, sélectionnez l'un des choix suivants :
 - Pour rechercher les touches, entrez Rechercher dans le glossaire métier dans la zone de recherche.
 - Pour faire défiler les touches, parcourez la liste pour trouver la commande **Rechercher dans le glossaire métier** dans la colonne **Commande**.
3. Cliquez sur la **Commande Rechercher dans le glossaire métier**.
4. Cliquez sur **Annuler la liaison de la commande**.
5. Dans le champ **Liaison**, entrez une combinaison de touches.
6. Cliquez sur **Appliquer**, puis sur **OK**.

ANNEXE A

Glossaire

clé candidate

Colonne ou ensemble de colonnes qui identifie de façon unique chaque ligne source dans une table de base de données.

clé documentée

Clé primaire déclarée dans la base de données source.

clé inférée

Colonne ou ensemble de colonnes que l'outil Analyst ou l'outil Developer considère comme une clé candidate en fonction des données de colonne.

colonne dépendante

Dans une dépendance fonctionnelle, la colonne contenant les valeurs qui sont déterminées par une colonne déterminante.

colonne déterminante

Dans une dépendance fonctionnelle, un ensemble de colonnes qui détermine la valeur de la colonne dépendante. Si la déterminante comporte zéro colonnes, la dépendante est une constante.

conservation

Processus de validation et de gestion des métadonnées d'une source de données découvertes afin qu'elles soient exploitables et puissent figurer dans des rapports.

correspondance directe

Dans une recherche globale, une correspondance directe est une ressource qui correspond à l'ensemble de la requête de recherche. Dans la recherche de découverte, une correspondance directe est une correspondance avec tout ou partie des métadonnées de la ressource qui correspond à la requête de recherche.

correspondance directe

Dans une recherche globale, une correspondance directe est une ressource qui correspond à l'ensemble de la requête de recherche. Dans la recherche de découverte, une correspondance directe est une correspondance avec tout ou partie des métadonnées de la ressource qui correspond à la requête de recherche.

découverte de clé étrangère

Le processus qui recherche dans une source de données les colonnes correspondant aux colonnes de clé primaire dans la source de données parent.

découverte de clé primaire

Le processus d'identification d'une colonne ou d'une combinaison de colonnes qui identifie de façon unique une ligne dans la source de données.

découverte de dépendance fonctionnelle

Le processus qui recherche les relations de dépendance fonctionnelle entre les colonnes d'une source de données.

découverte de domaines de données

Processus qui identifie tous les domaines de données associés à une colonne en fonction des valeurs ou du nom de la colonne.

découverte des données d'entreprise

Processus qui recherche les statistiques de profil de colonne, les domaines de colonne, les clés principales et les clés étrangères dans un grand nombre de sources de données réparties sur plusieurs connexions ou schémas.

dépendance fonctionnelle

La relation existant entre les colonnes d'un ensemble de colonnes d'une table donnée, dans laquelle la colonne déterminante détermine la colonne dépendante de manière fonctionnelle.

domaine de données

Objet de référentiel modèle prédéfini ou défini par l'utilisateur qui représente la signification fonctionnelle d'une colonne en fonction des données ou du nom de la colonne. Par exemple, le numéro de sécurité sociale, le numéro de carte de crédit ou l'identifiant de messagerie.

entrepôt de profilage

Base de données relationnelle qui stocke les informations de profilage, telles que les résultats de profils et de fiches d'évaluation.

fiche d'évaluation

Représentation graphique de valeurs valides pour une colonne source ou sortie d'une règle dans les résultats de profil. Utilisez des fiches d'évaluation pour mesurer la progression de la qualité des données.

glossaire de domaine de données

Conteneur de tous les domaines de données et les groupes de domaine de données dans l'outil Analyst et l'outil Developer.

groupe de domaines de données

Ensemble de domaines de données d'une catégorie de domaine de données spécifique.

groupe de mesures

Un groupe de mesures définies par l'utilisateur.

lignage des fiches d'évaluation

Diagramme qui affiche l'origine des données, décrit le chemin et indique comment les données circulent pour une métrologie ou un groupe de métrologies dans une fiche d'évaluation. Dans l'analyse du lignage des fiches d'évaluation, les zones ou nœuds représentent des objets. Les flèches représentent les relations des flux de données.

mesure

Colonne d'une source de données ou de sortie d'une règle faisant partie d'une fiche d'évaluation.

Poids métrique

Nombre entier supérieur ou égal à 0 affecté à une mesure. Un poids métrique définit la contribution de la mesure au score du groupe de mesures.

profil

Un objet contenant des règles régissant la découverte de formes dans les données source. Exécutez un profil pour évaluer la structure des données et vérifier que les colonnes de données contiennent les types d'informations que vous prévoyez.

profil d'objet de données

Un objet du référentiel qui définit le type d'analyse que vous effectuez sur la source de données.

profil de colonne

Un type de profil qui détermine les caractéristiques des colonnes dans la source de données, telles que la fréquence de la valeur, les pourcentages, les formes et les types de données.

profil de découverte des données d'entreprise

Type de profil que vous utilisez pour effectuer la découverte des données d'entreprise.

profil de jointure

Un type de profil qui détermine le degré de chevauchement entre un ensemble composé d'une ou plusieurs colonnes dans une source de données et un ensemble similaire dans la même source de données ou une source de données différente.

recherche de découverte

Type de recherche dans l'outil Analyst qui identifie les ressources en fonction de correspondances directes avec la requête de recherche et de relations avec d'autres objets qui correspondent à la requête de recherche.

règle

Logique métier réutilisable qui définit les conditions appliquées aux données lorsque vous exécutez un profil. Utilisez les règles pour valider les données dans un profil et mesurer la progression de la qualité des données. Vous pouvez créer une règle dans Informatica Analyst ou Informatica Developer.

règle de données

Logique métier réutilisable qui identifie une colonne par ses valeurs comme appartenant à un domaine de données particulier.

règle de nom de colonne

Logique métier réutilisable qui identifie une colonne par son nom comme appartenant à un domaine de données particulier.

score de groupe de mesures

Moyenne pondérée calculée de tous les scores de mesures du groupe de mesures.

score de mesure

Pourcentage de valeurs valides dans une mesure.

INDEX

A

analyse de jointure
présentation [156](#)

C

conservation
concepts [32](#)
Informatica Analyst [51](#)
Informatica Developer [125](#)
processus [33](#)
tâches [33](#)
création d'un profil personnalisé
profils [39](#)
création d'une règle d'expression
règles [57](#)

D

découverte de clé étrangère
présentation [154](#)
découverte de clé primaire
présentation [114](#)
découverte de dépendance fonctionnelle
présentation [116](#)
découverte de domaines de données
présentation [29](#)
présentation d'Informatica Analyst [75](#)
présentation d'Informatica Developer [135](#)
processus [31](#)
découverte de données
processus [20](#)
découverte des données d'entreprise
affichage des relations d'objet de données [163](#)
édition [153](#)
exécution dans l'outil Analyst [89](#)
modification dans l'outil Analyst [90](#)
présentation [147](#)
processus [148](#)
statut d'exécution [170](#)
vue domaines de données [168](#)
vue Profil de colonne [169](#)
vue Profilage de clé étrangère [163](#)
vue relations [162](#)
vue tabulaire [166](#)
découverte des données d'entreprise dans l'outil Analyst
conflit de type de données [94](#)
présentation [86](#)
processus [87](#)
vue Profils [95](#)
vue Résumé [92](#)
découverte du chevauchement
exécution de [159](#)

découverte du chevauchement (*a continué*)
présentation [158](#)
résultats [158](#)
domaine de données
création à partir de résultats de profil dans Informatica Analyst [77](#)
création à partir de résultats de profil dans Informatica Developer [137](#)
création dans Informatica Analyst [76](#)
création dans Informatica Developer [137](#)
présentation [30](#)
rechercher dans Informatica Developer [138](#)
domaines de données
exportation [139](#)
importation [138](#)
rechercher dans Informatica Analyst [77](#)

E

exporter
lignage des fiches d'évaluation vers XML [130](#)

F

fiche d'évaluation
configuration des notifications [71](#)
configuration des paramètres généraux des notifications [72](#)
consultation dans des applications externes [73](#)
fiches d'évaluation
affichage [62](#)
ajout de colonnes à une fiche d'évaluation [61](#)
coût des données non valides [60](#)
coût fixe [60](#)
coût variable [60](#)
création d'un groupe de mesures [64](#)
définition des seuils [63](#)
déplacement des scores [64](#)
développement [65](#)
édition [62](#)
exécution [62](#)
graphe de tendance [66](#)
Groupes de mesures [63](#)
Informatica Analyst [58](#)
Informatica Developer [129](#)
mesures [60](#)
modification d'un groupe de mesures [64](#)
notifications [70](#)
poids métriques [60](#)
présentation [28](#)
processus de Informatica Analyst [59](#)
Suppression d'un groupe de mesures [65](#)

G

- glossaire de domaine de données
 - Informatica Analyst [75](#)
 - Informatica Developer [136](#)
 - présentation [31](#)
- graphes de tendance
 - affichage [67](#)
 - coût [67](#)
 - exportation depuis Informatica Analyst [68](#)
 - score [66](#)
- groupe de domaines de données
 - création dans Informatica Analyst [76](#)
 - création dans Informatica Developer [136](#)
 - présentation [30](#)

I

- Informatica Analyst
 - présentation des profils de colonne [36](#)
 - règles [54](#)
 - résultats de profil de colonne [44](#)
- Informatica Developer
 - présentation des profils [109](#)
 - règles [127](#)
 - vues de profils [111](#)
- intégration de fiche d'évaluation
 - Informatica Analyst [73](#)

L

- lignage des fiches d'évaluation
 - affichage dans Informatica Analyst [74](#)
 - affichage dans Informatica Developer [130](#)

O

- objet de données fichier plat
 - synchronisation [42](#)
- objet de données table
 - synchronisation [43](#)
- objet de mappage
 - exécution d'un profil [132](#)
- options de configuration
 - découverte des données d'entreprise dans l'outil Analyst [87](#)
- options de découverte des domaines de données
 - Informatica Developer [140](#)
- options de profil
 - découverte des données d'entreprise [149](#)

P

- profil
 - composants [24](#)
- profil de clé étrangère
 - découverte [155](#)
- profil de colonne
 - développement [50](#)
 - Informatica Developer [113](#)
 - options [27](#)
 - présentation [26](#)
 - processus [37](#)
- profil de découverte des données d'entreprise
 - création de scripts DDL [160](#)

- profil de découverte des données d'entreprise (*a continué*)
 - exécution [153](#)
 - Scripts DDL [160](#)
- profilage
 - architecture [19](#)
- Profilage de mapplet et de mappage
 - Présentation [132](#)
- profils
 - création d'un profil personnalisé [39](#)
 - exécution [41](#)
- profils d'objet de données
 - création d'un profil unique [118](#)
 - création de plusieurs profils [119](#)
 - découverte des données d'entreprise [152](#)
 - présentation [112](#)
- Profils de mapplet et de mappage
 - Présentation [132](#)
- projets
 - terme métier Metadata Manager [106](#)

R

- recherche de découverte
 - prérequis [97](#)
- recherche de découverte dans l'outil Analyst
 - processus [98](#)
- rechercher
 - glossaire métier [172](#)
- règles
 - application d'une règle prédéfinie [55](#)
 - application dans Informatica Developer [128](#)
 - application dans PowerCenter Express [128](#)
 - création d'une règle d'expression [57](#)
 - création dans Informatica Developer [128](#)
 - expression [56](#)
 - prédéfinies [55](#)
 - prérequis [127](#)
 - présentation [27](#)
- règles d'expression
 - processus [56](#)
- règles prédéfinies
 - processus [55](#)
- résultats de fiche d'évaluation
 - exportation [69](#)
 - exportation depuis Informatica Analyst [69](#)
 - exporter vers Excel [69](#)
- Résultats de la découverte de domaines de données
 - Informatica Analyst [82](#)
- résultats de la découverte des domaines de données
 - exportation dans Informatica Analyst [84](#)
 - exportation depuis Informatica Analyst [84](#)
 - exportation depuis Informatica Developer [146](#)
- Résultats de la découverte des domaines de données
 - Informatica Developer [143](#)
- résultats de la découverte des données d'entreprise
 - enregistrer en tant qu'image [165](#)
 - exportation [171](#)
 - présentation [161](#)
- résultats de la découverte des données d'entreprise dans l'outil Analyst
 - présentation [91](#)
- résultats de la recherche de découverte
 - interface [100](#)
- résultats de la recherche de découverte dans l'outil Analyst
 - présentation [100](#)
- résultats de profil
 - approbation des domaines de données [83](#)

résultats de profil (*a continué*)

approbation des domaines de données dans Informatica Developer [145](#)

Approbation des types de données [51](#)

approbation des types de données dans Informatica Developer [125](#)

conservation des relations de colonne dans Informatica Developer [167](#)

développement [50](#)

Excel [52](#)

exportation [52](#)

exportation dans Informatica Developer [126](#)

exportation depuis Informatica Analyst [53](#)

formes de colonne [47](#)

rejet des domaines de données [83](#)

rejet des domaines de données dans l'outil Developer [146](#)

rejet des types de données [51](#)

rejet des types de données dans l'outil Developer [125](#)

résumé [45](#)

statistiques de colonne [49](#)

types de données de colonne [48](#), [124](#)

résultats de profil (*a continué*)

valeurs de colonne [46](#)

résultats de profil de colonne

Informatica Developer [121](#)

profil de colonne [121](#)

résultats de profil de la découverte des domaines de données
Microsoft Excel [84](#)

T

terme métier

recherche d'un terme métier [107](#)

terme métier Metadata Manager

gestion des termes métier [107](#)

projets [106](#)

termes métier

personnalisation des touches de raccourci [173](#)

recherche [173](#)