



Informatica®

10.1.1 HotFix 2

Guide de la découverte de données

Ce logiciel et la documentation associée sont fournis uniquement sous un accord de licence séparé contenant des restrictions d'utilisation et de divulgation. Il est interdit de reproduire ou de transmettre sous quelle que forme et par quel que moyen que ce soit (électronique, photocopie, enregistrement ou autre) tout ou partie de ce document sans le consentement préalable d'Informatica LLC.

Informatica et le logo Informatica sont des marques ou des marques déposées d'Informatica LLC aux États-Unis et dans de nombreux autres pays. Une liste actuelle des marques déposées d'Informatica est disponible sur le site <https://www.informatica.com/trademarks.html>. Les autres noms de société ou de produit peuvent être des marques de commerce ou des marques déposées de leurs détenteurs respectifs.

U.S. GOVERNMENT RIGHTS Les programmes, les logiciels, les bases de données et les documents connexes et les données techniques fournis aux clients du gouvernement américain sont des « logiciels commerciaux » ou des « données techniques commerciales », conformément au règlement fédéral sur les acquisitions et aux règlements supplémentaires propres à l'Agence. En tant que tel, l'utilisation, la duplication, la divulgation, la modification et l'adaptation sont assujetties aux restrictions et aux conditions de licence énoncées dans le contrat gouvernemental applicable et, dans la mesure applicable par les termes du contrat gouvernemental, les droits additionnels énoncés dans la réglementation FAR 52.227-19, licence de logiciel d'ordinateur commercial.

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 © University of Toronto. Tous droits réservés. Copyright © Daniel Veillard. Tous droits réservés. Copyright © Unicode, 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. Copyright © Tableau Software, Inc. Tous droits réservés. Copyright © MaxMind, Inc. Tous droits réservés. Copyright © TMate Software s.r.o. Tous droits réservés. Copyright © MapR Technologies Inc. Tous droits réservés. Copyright © Amazon Corporate LLC. Tous droits réservés. Copyright © Highsoft. Tous droits réservés. Copyright © Python Software Foundation. Tous droits réservés. Copyright © BeOpen.com. Tous droits réservés. Copyright © CNRI. 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 LÉTAT », 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://www.cryptix.org/LICENSE.TXT>, <http://www.cryptix.org/LICENSE.TXT>.

hsqldb.org/web/hsqLicense.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/license.html> ; <http://www.jgraph.com/jgraphdownload.html> ; <http://www.jcraft.com/jsch/LICENSE.txt> ; http://jotm.objectweb.org/bsd_license.html ; <http://www.w3.org/Consortium/Legal/2002/copyright-software-20021231> ; <http://www.slf4j.org/license.html> ; <http://nanoxml.sourceforge.net/orig/copyright.html> ; <http://www.json.org/license.html> ; <http://forge.ow2.org/projects/javaservice/> , <http://www.postgresql.org/about/licence.html>, <http://www.sqlite.org/copyright.html>, <http://www.tcl.tk/software/tcltk/license.html>, <http://www.jaxen.org/faq.html>, <http://www.jdom.org/docs/faq.html>, <http://www.slf4j.org/license.html>; <http://www.iodbc.org/dataspace/iodbc/wiki/IODBC/License>; <http://www.keplerproject.org/md5/license.html>; <http://www.toedter.com/en/jcalendar/license.html>; <http://www.edankert.com/bounce/index.html>; <http://www.net-snmp.org/about/license.html>; <http://www.openmdx.org/#FAQ>; http://www.php.net/license/3_01.txt; <http://srp.stanford.edu/license.txt>; <http://www.schneier.com/blowfish.html>; <http://www.jmock.org/license.html>; <http://xsom.java.net>; <http://benalman.com/about/license/>; <https://github.com/CreateJS/EaselJS/blob/master/src/easeljs/display/Bitmap.js>; <http://www.h2database.com/html/license.html#summary>; <http://jsoncpp.sourceforge.net/LICENSE>; <http://jdbc.postgresql.org/license.html>; <http://protobuf.googlecode.com/svn/trunk/src/google/protobuf/descriptor.proto>; <https://github.com/rantav/hector/blob/master/LICENSE>; <http://web.mit.edu/Kerberos/krb5-current/doc/mitK5license.html>; <http://jibx.sourceforge.net/jibx-license.html>; <https://github.com/lyokato/libgeohash/blob/master/LICENSE>; <https://github.com/hjiang/jsonxx/blob/master/LICENSE>; <https://code.google.com/p/lz4/>; <https://github.com/jedisct1/libsodium/blob/master/LICENSE>; <http://one-jar.sourceforge.net/index.php?page=documents&file=license>; <https://github.com/EsotericSoftware/kryo/blob/master/license.txt>; <http://www.scala-lang.org/license.html>; <https://github.com/tinkerpop/blueprints/blob/master/LICENSE.txt>; <http://gee.cs.oswego.edu/dl/classes/EDU/oswego/cs/dl/util/concurrent/intro.html>; <https://aws.amazon.com/asl/>; <https://github.com/twbs/bootstrap/blob/master/LICENSE>; <https://sourceforge.net/p/xmlunit/code/HEAD/tree/trunk/LICENSE.txt>; <https://github.com/documentcloud/underscore-contrib/blob/master/LICENSE>, and <https://github.com/apache/hbase/blob/master/LICENSE.txt>.

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>), le 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.

Consultez les brevets applicables à l'adresse <https://www.informatica.com/legal/patents.html>.

EXCLUSION DE RESPONSABILITÉ : Informatica LLC 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 LLC 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.

Les renseignements contenus dans cette documentation sont sujets à modification sans préavis. Si vous constatez des problèmes liés à la documentation, merci de les signaler par courriel à l'adresse infa_documentation@Informatica.com.

Les produits Informatica sont garantis conformément aux termes et conditions des accords en vertu desquels ils sont fournis. INFORMATICA FOURNIT LES INFORMATIONS DE CE DOCUMENT « EN L'ÉTAT » SANS GARANTIE D'AUCUNE SORTE, EXPRESSE OU IMPLICITE, NOTAMMENT AUCUNE GARANTIE DE QUALITÉ MARCHANDE, D'ADAPTATION À UN USAGE PARTICULIER ET D'ABSENCE DE CONTREFAÇON

Date de publication: 2018-07-03

Sommaire

Préface.....	14
Ressources Informatica.	14
Informatica Network.	14
Base de connaissances Informatica.	14
Documentation Informatica.	15
Matrices de disponibilité de produit Informatica.	15
Informatica Velocity.	15
Informatica Marketplace.	15
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.	23
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
Verrous de profil de référentiel et gestion de profil avec versions.	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
Options d'échantillonnage.	38
Options de développement.	38
Environnement d'exécution.	39
Présentation des profils de système d'exploitation dans Informatica Analyst.	39
Sélection d'un profil de système d'exploitation.	39
Présentation des verrous des ressources du référentiel et du développement basé sur l'équipe. . . .	39
Création d'un profil de colonne dans Informatica Analyst.	40
Modification d'un profil de colonne.	42
Exécution d'un profil.	42
Option de synchronisation.	43
Synchronisation d'un objet de données fichier plat.	43
Synchronisation d'un objet de données relationnel.	44
Chapitre 7: Règles dans Informatica Analyst.	45
Présentation des règles dans Informatica Analyst.	45
Règles d'un profil de colonne.	45
Règles prédéfinies.	46
Processus de règles prédéfinies.	46
Application d'une règle prédéfinie.	46
Règles d'expression.	47
Processus de règles d'expression.	47
Création d'une règle d'expression.	48
Chapitre 8: Filtres dans Informatica Analyst.	50
Présentation des filtres dans Informatica Analyst.	50
Création d'un filtre.	50
Création d'un filtre simple.	51
Création d'un filtre avancé.	52
Création d'un filtre SQL.	53
Gestion des filtres.	53
Chapitre 9: Résultats de profil de colonne dans Informatica Analyst.	55
Présentation des résultats de profil de colonne dans Informatica Analyst.	55
Vue de résumé.	56
Propriétés de la vue de résumé.	57
Filtres par défaut de la vue de résumé.	58

Vue détaillée.	58
Volets de la vue détaillée.	59
Statistiques.	60
Aperçu des données.	61
Types de données.	61
Valeurs hors norme.	62
Modèles.	63
Valeurs.	64
Types d'exécution de profil.	67
Dernière exécution de profil.	67
Exécution de profil historique.	67
Exécution de profil consolidée.	68
Sélection d'une exécution de profil.	68
Présentation des résultats de la comparaison de plusieurs profils.	69
Résultats de la comparaison de plusieurs profils.	69
Vue de résumé des résultats de la comparaison de profils	70
Vue détaillée des résultats de la comparaison de profils	73
Développement de profil de colonne.	74
Développement des données de ligne.	74
Application de filtres aux données de développement.	74
Conservation dans l'outil Analyst.	75
Approbation de types de données et de domaines de données.	75
Rejet de types de données et de domaines de données.	76
Fichiers d'exportation de profil de colonne dans Informatica Analyst.	76
Résultats de l'exportation de profil dans un fichier CSV.	76
Résultats de l'exportation de profil dans Microsoft Excel.	77
Exportation de résultats de profil depuis Informatica Analyst.	77

Chapitre 10: Termes d'entreprise, commentaires et balises dans Informatica Analyst. 79

Présentation des termes d'entreprise, des commentaires et des balises dans Informatica Analyst.	79
Termes d'entreprise.	79
Attribution de termes d'entreprise à des colonnes.	80
Commentaires.	80
Ajout de commentaires à un profil ou à des colonnes.	80
Balises.	81
Attribution de balises à un profil ou à des colonnes.	81

Chapitre 11: Fiches d'évaluation dans Informatica Analyst. 83

Présentation des fiches d'évaluation dans Informatica Analyst.	83
Processus de fiche d'évaluation de Informatica Analyst.	84
Création d'une fiche d'évaluation dans Informatica Analyst.	85
Ajouter des colonnes à une fiche d'évaluation existante.	86

Exécution d'une fiche d'évaluation.	88
Affichage d'une fiche d'évaluation.	88
Modification d'une fiche d'évaluation.	89
Mesures.	90
Poids métriques.	90
Valeur de la qualité des données.	90
Définition des seuils.	91
Groupes de mesures.	91
Création d'un groupe de mesures.	91
Déplacement des scores vers un groupe de mesures.	92
Modification d'un groupe de mesures.	92
Suppression d'un groupe de mesures.	93
Développement dans les colonnes.	93
Graphes de tendance.	93
Graphe de tendance de score.	94
Graphe de tendance de coût.	94
Affichage des graphes de tendances.	95
Exportation des graphes de tendance.	96
Tableau de bord des fiches d'évaluation dans Informatica Analyst.	97
Fiches d'évaluation par projet.	98
Tendance d'exécution des fiches d'évaluation.	99
Objets de données avec fiche d'évaluation.	100
Tendance de scores cumulatifs.	101
Fichier d'exportation de fiche d'évaluation dans Informatica Analyst.	103
Résultats de l'exportation de fiche d'évaluation dans Microsoft Excel.	103
Exportation des résultats de fiche d'évaluation à partir d'Informatica Analyst.	103
Notifications de la fiche d'évaluation.	104
Modèle de notification de message électronique.	104
Configuration des notifications de la fiche d'évaluation.	106
Configuration des paramètres généraux des notifications de la fiche d'évaluation.	106
Lignage des fiches d'évaluation.	107
Affichage du lignage des fiches d'évaluation dans Informatica Analyst.	107

Chapitre 12: Découverte des domaines de données dans Informatica Analyst. 109

Présentation de la découverte de domaines de données dans Informatica Analyst.	109
Glossaire de domaine de données dans Informatica Analyst.	110
Création d'un groupe de domaines de données dans Informatica Analyst.	110
Création d'un domaine de données dans Informatica Analyst.	110
Création d'un domaine de données depuis les résultats de profil dans Informatica Analyst.	111
Rechercher des domaines de données et des groupes de domaines de données dans Informatica Analyst.	111
Options de découverte des domaines de données dans Informatica Analyst.	112

Sélection de colonne de domaine de données dans Informatica Analyst.	112
Sélection des domaines de données dans Informatica Analyst.	113
Options d'inférence des domaines de données dans Informatica Analyst.	113
Création d'un profil de colonne pour découvrir des domaines de données dans Informatica Analyst.	116
Modification d'un profil dans Informatica Analyst.	117
Exécution d'un profil pour effectuer la découverte de domaines de données.	117
Résultats de la découverte de domaines de données dans Informatica Analyst.	118
Approbation des domaines de données.	118
Rejet des domaines de données.	118
Fichiers d'exportation de la découverte des domaines de données dans Informatica Analyst.	119
Résultats de la découverte de domaines de données dans Microsoft Excel.	119
Exportation des résultats de la découverte de domaines de données depuis Informatica Analyst.	120

Chapitre 13: Découverte des données d'entreprise dans Informatica Analyst 121

Présentation de la découverte de données d'entreprise dans Informatica Analyst.	121
Processus de découverte des données d'entreprise dans Informatica Analyst.	122
Options de configuration pour la découverte des données d'entreprise.	122
Paramètres de découverte de domaines de données.	122
Paramètres de profil de colonne.	124
Création d'un profil de découverte des données d'entreprise dans Informatica Analyst.	124
Modification des options de découverte de données d'entreprise.	126

Chapitre 14: Résultats de la découverte des données d'entreprise dans Informatica Analyst. 127

Présentation des résultats de la découverte des données d'entreprise dans Informatica Analyst.	127
Vue Résumé.	127
Résultats de profil de la vue Résumé.	128
Affichage des résultats de la découverte des domaines de données.	129
Affichage des résultats de profil de colonne.	129
Conflit de type de données.	130
Afficher les conflits de type de données.	130
Vue Profils.	131
Affichage des propriétés de profil.	131

Chapitre 15: Recherche de découverte dans Informatica Analyst. 132

Présentation de la recherche de découverte dans Informatica Analyst.	132
Prérequis de la recherche de découverte.	133
Processus de recherche de découverte dans Informatica Analyst.	134
Options de recherche de découverte.	134
Critères de recherche de découverte.	135
Recherche d'une ressource.	135

Résultats de la recherche de découverte dans Informatica Analyst.	136
Panneau des résultats de la recherche de découverte.	136
Filtrage des résultats de la recherche de découverte.	137
Types de correspondances.	138
Correspondance directe.	138
Correspondance indirecte.	138
Affichage des informations de correspondance.	138
Ouverture des ressources à partir des résultats de la recherche de découverte.	139
Ressources liées.	139
Ressources liées pour chaque type de ressource.	140
Affichage des ressources liées.	140
Forum Aux Questions (FAQ).	141
 Chapitre 16: Bureau de glossaire métier dans Informatica Analyst.	 142
Termes métier.	142
Gestion des termes métier dans le glossaire métier du gestionnaire de métadonnées.	143
Recherche d'un terme métier dans le Bureau de glossaire métier.	143
 Partie III: Découverte de données avec Informatica Developer.	 144
 Chapitre 17: Profils de Informatica Developer.	 145
Présentation des profils d'Informatica Developer.	145
Vues des profils d'Informatica DeveloperVues des profils.	147
Verrous d'objets du référentiel et développement basé sur l'équipe contenant des objets avec version.	148
 Chapitre 18: Profils d'objet de données.	 149
Présentation de profils d'objet de données.	149
Profils de colonne dans Informatica Developer.	150
Options de filtrage.	151
Options d'échantillonnage.	152
Environnement d'exécution.	153
Découverte de clé primaire.	153
Propriétés d'inférence de clés primaires.	153
Propriétés de clés primaires inférées.	154
Propriétés de violations de clés.	154
Découverte de dépendance fonctionnelle.	155
Propriétés d'inférence de dépendances fonctionnelles.	155
Propriétés des dépendances fonctionnelles inférées.	156
Propriétés de violations de dépendances fonctionnelles.	156
Profils du système d'exploitation dans Informatica Developer.	156
Sélection d'un profil de système d'exploitation.	157
Création d'un profil d'objet de données unique dans Informatica Developer.	157

Création de plusieurs profils d'objet de données dans Informatica Developer.	158
Modification d'un profil.	158
Option de synchronisation.	159
Synchronisation d'un objet de données de fichier plat.	159
Synchronisation d'un objet de données relationnel.	159
Commentaires.	160
Ajout de commentaires dans Informatica Developer.	160

Chapitre 19: Profils de colonne dans des sources de données semi-structurées. 161

Présentation de profils de colonne dans des sources de données semi-structurées.	161
Profil de colonne dans une source de données XML ou JSON.	162
Création d'un objet de données sur une source de données XML ou JSON.	162
Profil de colonne dans une source de données semi-structurée dans HDFS.	163
Profil de colonne sur une source de données XML ou JSON dans HDFS.	163
Profil de colonne d'une source de données Avro ou Parquet dans HDFS.	163
Création d'une connexion HDFS.	164
Création d'un objet de données de fichier complexe sur un fichier JSON ou XML dans HDFS.	164
Création d'un objet de données sur une source de données Avro ou Parquet.	165
Exécution d'un profil de colonne sur des sources de données semi-structurées.	166

Chapitre 20: Résultats de profil de colonne dans Informatica Developer. . . . 168

Résultats de profil de colonne dans Informatica DeveloperRésultats de profil de colonne.	168
Propriétés des valeurs de colonne.	169
Propriétés des formes de colonne.	170
Propriétés des statistiques de colonne.	170
Propriétés des types de données de colonne.	171
Conservation dans Informatica DeveloperConservation dans Informatica Developer.	172
Approbation des types de donnéesApprobation des types de données dans Informatica Developer.	172
Rejet de types de donnéesRejet de types de données dans Informatica Developer.	172
Exportation des résultats de profil depuis Informatica Developer.	173

Chapitre 21: Règles dans Informatica Developer. 174

Présentation des règles dans Informatica DeveloperLignes directrices relatives aux règles.	174
Création d'une règle dans Informatica Developer.	175
Application d'une règle dans Informatica DeveloperApplication d'une règle.	175

Chapitre 22: Fiches d'évaluation dans Informatica Developer. 176

Présentation des fiches d'évaluation dans Informatica Developer.	176
Création d'une fiche d'évaluation.	176
Exportation d'un fichier de ressource pour le lignage des fiches d'évaluation.	177
Affichage du lignage des fiches d'évaluation dans Informatica Developer.	177

Chapitre 23: Profilage de mapplet et de mappage.	179
Présentation du profilage de mapplet et de mappage	179
Exécution d'un profil sur un mapplet ou un objet de mappage.	179
Comparaison des profils pour les objets de mappage ou de mapplet.	180
Génération d'un mappage à partir d'un profil.	180
 Chapitre 24: Découverte des domaines de données dans Informatica Developer.	 182
Présentation de la découverte de domaines de données dans Informatica Developer.	182
Glossaire de domaine de données dans Informatica Developer.	183
Création d'un groupe de domaines de données dans Informatica Developer.	183
Création d'un domaine de données dans Informatica Developer.	184
Création d'un domaine de données depuis les résultats de profil dans Informatica Developer	184
Rechercher des domaines de données dans Informatica Developer.	185
Importation des domaines de données.	185
Exportation des domaines de données.	186
Options de découverte des domaines de données dans Informatica Developer.	187
Sélection de domaines de données dans Informatica Developer.	187
Sélection de colonne de domaine de données dans Informatica Developer.	188
Options d'inférence des domaines de données dans Informatica Developer.	188
Création d'un profil pour découvrir des domaines de données dans Informatica Developer.	190
Modification d'un profil dans Informatica Developer.	190
Exécution d'un profil pour effectuer une découverte des domaines de données dans Informatica Developer.	191
Résultats de la découverte des domaines de données dans Informatica Developer.	191
Affichage par groupes de domaines de données.	192
Affichage par colonnes.	193
Vérification des résultats.	193
Approbation des domaines de données.	193
Rejet des domaines de données.	194
Exportation des résultats de la découverte des domaines de données depuis Informatica Developer.	194
 Chapitre 25: Découverte des données d'entreprise dans Informatica Developer.	 195
Présentation de la découverte des données d'entreprise dans Informatica Developer.	195
Processus de découverte des données d'entreprise.	196
Options de profil pour la découverte des données d'entreprise.	197
Sélection de domaine de données pour la découverte des données d'entreprise.	197
Options d'échantillonnage de profil de colonne pour la découverte des données d'entreprise.	198
Option d'environnement d'exécution.	198
Options d'inférence de clé primaire pour la découverte des données d'entreprise.	199
Options d'inférence de clé étrangère pour la découverte des données d'entreprise.	199

Paramètres de conservation automatique pour l'inférence de clé étrangère.	200
Création d'un profil de découverte des données d'entreprise dans Informatica Developer.	202
Modification d'un profil.	203
Exécution d'un profil de découverte des données d'entreprise.	204
Découverte de clé étrangère.	205
Définition des relations entre les objets parent et enfant.	205
Découverte de relations de clés étrangères entre les objets de données.	206
Résultats d'analyse de clé étrangère.	206
Analyse de jointure.	207
Création d'un profil de jointure.	207
Résultats d'analyse de jointure.	208
Exportation des résultats de profil de jointure dans un fichier.	208
Découverte du chevauchement.	209
Résultats de découverte du chevauchement.	209
Découverte de données de chevauchement.	210
Fichiers de script DDL.	211
Création de scripts DDL à partir d'un profil de découverte des données d'entreprise.	211
Chapitre 26: Résultats de la découverte des données d'entreprise.	212
Présentation des résultats de la découverte des données d'entreprise.	212
Vue Relations.	213
Rechercher un objet de données.	213
Navigation dans la vue Profilage de clé étrangère.	214
Vue Profilage de clé étrangère.	214
Affichage des relations d'objets de données.	214
Zoom avant et zoom arrière dans la vue.	215
Recherche d'un objet de données.	215
Affichage des relations de colonne.	216
Enregistrement du diagramme de la relation de l'entité sous la forme d'une image.	216
Affichage des résultats de profil d'objet de données depuis la vue Profilage de clé étrangère.	216
Vue tabulaire.	217
Volet Détails de la table.	217
Vérification des résultats de la découverte des données d'entreprise.	217
Conservation des relations de colonne.	218
Validation des résultats dans le référentiel modèle.	218
Vue Domaines de données.	219
Affichage des résultats de la découverte des domaines de données.	219
Vérification des résultats de la découverte des domaines de données.	219
Développement des lignes.	219
Affichage des résultats de profil d'objet de données depuis la vue Domaines de données.	220
Vue Profil de colonne.	220
Affichage des résultats de profil d'objet de données.	220

Affichage des résultats de profil de colonne lors de l'exécution de la découverte des données d'entreprise.	221
Affichage des résultats de la découverte des domaines de données lors de l'exécution de la découverte des données d'entreprise.	221
Affichage du statut d'exécution de la découverte des données d'entreprise.	221
Fichiers d'exportation de la découverte des données d'entreprise.	222
Exportation des résultats de la découverte des données d'entreprise.	222
Chapitre 27: Bureau de glossaire métier dans Informatica Developer.	223
Recherche dans Business Glossary.	223
Recherche d'un terme métier.	224
Personnalisation des touches de raccourci pour rechercher un terme métier.	224
Index.	225

Préface

Le *Guide de découverte de données* s'adresse aux utilisateurs d'Informatica Analyst et d'Informatica Developer. Il contient des informations sur l'utilisation des profils pour l'analyse du contenu, de la qualité et de la structure des sources de données, et l'utilisation de la découverte de données pour découvrir les métadonnées des systèmes source qui incluent le contenu et la structure.

Utilisez les profils pour découvrir les problèmes de qualité des données dans une source de données et pour comprendre les relations entre les colonnes dans une ou plusieurs sources de données.

Ressources Informatica

Informatica Network

Informatica Network héberge le support client international Informatica, la base de connaissances Informatica et d'autres ressources de produits. Vous pouvez accéder à Informatica Network à l'adresse <https://network.informatica.com>.

En tant que membre, vous pouvez :

- Accéder à toutes les ressources Informatica d'un emplacement.
- Rechercher des ressources de produits dans la base de connaissances, notamment la documentation, les FAQ et les meilleurs pratiques.
- Afficher les informations de disponibilité de produit.
- Vérifier votre cas de support.
- Rechercher votre réseau de groupe d'utilisateurs local Informatica et collaborer avec vos pairs.

Base de connaissances Informatica

La base de connaissances Informatica Network vous permet de rechercher les ressources de produits telles que la documentation, les articles de procédures pratiques, les meilleures pratiques et les matrices de disponibilité de produit (PAM).

Pour accéder à la base de connaissances, visitez le site <https://kb.informatica.com>. 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.

Documentation Informatica

Pour obtenir la dernière documentation relative à votre produit, parcourez la base de connaissances Informatica à l'adresse

https://kb.informatica.com/_layouts/ProductDocumentation/Page/ProductDocumentSearch.aspx.

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.

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.

Si vous êtes un membre d'Informatica Network, vous pouvez accéder aux PAM à l'adresse

<https://network.informatica.com/community/informatica-network/product-availability-matrices>.

Informatica Velocity

Informatica Velocity est un ensemble de conseils et de meilleures pratiques développé par les services professionnels d'Informatica. 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 êtes membre d'Informatica Network, vous pouvez accéder aux ressources d'Informatica Velocity à l'adresse <http://velocity.informatica.com>.

Si vous avez des questions, des commentaires et des suggestions sur Informatica Velocity, contactez le support des services professionnels d'Informatica à l'adresse ips@informatica.com.

Informatica Marketplace

Informatica Marketplace est un forum dans lequel vous pouvez trouver des solutions qui permettent d'augmenter, d'étendre ou d'améliorer vos implémentations Informatica. L'utilisation d'une des centaines de solutions créées par les développeurs et partenaires Informatica vous permettra d'améliorer votre productivité et d'accélérer le temps d'implémentation de vos projets. Vous pouvez accéder à Informatica Marketplace à l'adresse <https://marketplace.informatica.com>.

Support client international Informatica

Vous pouvez contacter un centre de support international par téléphone ou via le support en ligne sur Informatica Network.

Pour trouver le numéro de téléphone du support client international Informatica, visitez le site Web Informatica à l'adresse

<http://www.informatica.com/us/services-and-training/support-services/global-support-centers>.

Si vous êtes un membre d'Informatica Network, vous pouvez utiliser le support en ligne à l'adresse

<http://network.informatica.com>.

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 de l'entreprise permet d'appréhender les points forts et les points faibles de ses données et métadonnées.

Vous pouvez utiliser Informatica Analyst et Informatica Developer pour analyser les données et les métadonnées source. 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 tool ou de l'outil Developer tool. 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 tool et l'outil Analyst tool :

- 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.
- Effectuer la découverte de domaine 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.
- Choisissez un profil de système d'exploitation pour créer et exécuter des profils de colonne, des profils de découverte de données d'entreprise et des fiches d'évaluation en fonction des autorisations de l'utilisateur du système d'exploitation que vous définissez dans le profil du système d'exploitation.

- Utiliser des verrous des ressources du référentiel pour empêcher les autres utilisateurs de remplacer les données.
- Utiliser le système de contrôle de version pour enregistrer plusieurs versions d'un profil.
- Créer et affecter des balises aux objets de données.
- Rechercher la signification d'un nom d'objet comme un terme métier dans Business Glossary Desktop. 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 tool :

- 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 de données incluent le profil de colonne, l'inférence des relations de clé primaire à clé étrangère, la découverte de domaine de données et la génération d'un résumé graphique consolidé des relations des données.

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

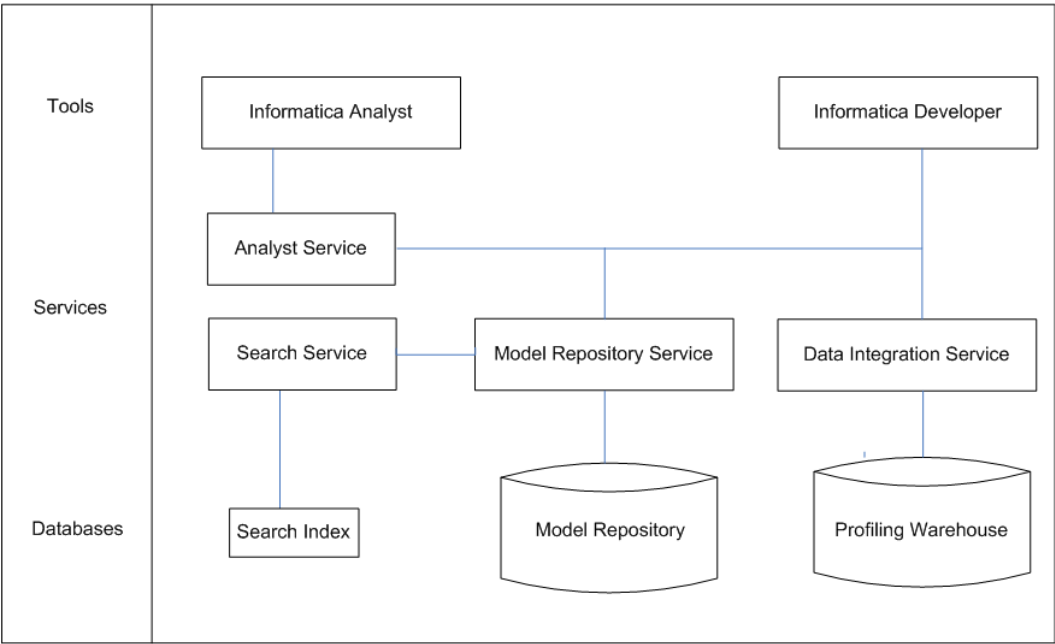
- Effectuer la découverte de données d'entreprise sur un grand nombre de sources de données sur plusieurs connexions. Vous pouvez afficher un résumé consolidé des résultats de découverte de métadonnées de colonne et de domaines de données.
- Effectuer une recherche de découverte pour trouver où les données 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.
- Afficher les résultats d'une exécution de profil historique.
- Comparer les résultats de profil pour deux exécutions de profils dans un profil de colonne.
- Afficher le lignage des fiches d'évaluation pour chaque score de fiche d'évaluation et groupe de scores.
- Afficher le tableau de bord des fiches d'évaluation.
- Ajouter des commentaires à un profil ou aux colonnes d'un profil.
- Attribuer des balises à un profil ou aux colonnes d'un profil.
- Attribuer des termes d'entreprise aux colonnes d'un profil.

Architecture de profilage

L'architecture de profilage comprend des outils, des services et des bases de données. Le composant Outils comprend les applications clientes. 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.

L'image suivante montre les composants de l'architecture du profilage :



Lors de l'exécution d'un profil, le service Analyst ou l'outil Developer tool reçoit la définition du profil depuis le service de référentiel modèle. Puis, le service Analyst ou l'outil Developer tool appelle le plug-in de profilage dans le service d'intégration de données. Ensuite, le plug-in de profilage traite la tâche de profilage et l'envoie au service d'intégration de données. Le service d'intégration de données 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	Application cliente Web que vous pouvez utiliser pour découvrir et analyser les données et les métadonnées des sources de données, ainsi que pour établir le rapport correspondant.
Informatica Developer	Application cliente que vous utilisez pour effectuer une découverte de données avancée telle que la découverte de clé primaire, la découverte de clé étrangère et la découverte de données d'entreprise.
Service Analyst	Service d'application qui exécute l'outil Analyst tool et gère les connexions entre les composants du service et les utilisateurs de l'outil Analyst tool.
Service de recherche	Service d'application qui gère la recherche dans l'outil Analyst tool. 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.
Service d'intégration de données	Service d'application qui effectue des tâches d'intégration de données pour l'outil Analyst tool, l'outil Developer tool et les clients externes.
Référentiel modèle	Base de données relationnelle qui stocke les métadonnées de projets créés par l'outil Analyst tool ou l'outil Developer tool.
Entrepôt de profilage	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, 23](#)
- [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 fait référence à des valeurs de données, des fréquences et des 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 primaire

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 tool ne s'affichent pas dans l'outil Developer tool tant que vous n'actualisez pas la connexion de l'outil Developer tool au référentiel modèle.

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 exécuter un profil sur cet échantillon de données.

Règle

Logique métier qui définit les conditions appliquées aux données lorsque vous exécutez un profil. Ajoutez une règle au profil pour 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. Attribuez des balises à un profil ou aux colonnes d'un profil dans l'outil Analyst tool.

Commentaire

Description concernant le profil. Utilisez des commentaires pour partager des informations sur les profils avec d'autres utilisateurs des outils Analyst tool et Developer tool. Ajoutez des commentaires à un profil ou aux colonnes d'un profil dans l'outil Analyst tool.

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 Null, distinctes et non distinctes dans les colonnes et types de données inférés pour les valeurs de colonne.- Modèles de fréquence et de caractère des valeurs de données dans une colonne sélectionnée et un résumé des statistiques pour la colonne.- 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 l'exécution du profil.
Profil de clé primaire	<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.
Profil de dépendance fonctionnelle	<ul style="list-style-type: none">- Dépendances fonctionnelles inférées.- Nombre de violations de dépendances fonctionnelles.

Type de profil	Résultats
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 primaires 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 de colonne et données qui correspondent aux domaines de données prédéfinis. - 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 génération de rapports tiers pour lire les résultats de profil à partir de 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](#)
- [Verrous de profil de référentiel et gestion de profil avec versions, 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 modèles.

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

- Le nombre et le pourcentage de valeurs Null, uniques et non uniques dans chaque colonne.
- 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.
- Les types de données documentées, les types de données inférées et les conflits possibles entre ces types de données.
- Les valeurs hors norme de la fréquence de valeur et du modèle.

Vous pouvez configurer les options suivantes lorsque vous créez ou modifiez un profil :

- Options de profil de colonne. Vous pouvez sélectionner les colonnes sur lesquelles vous voulez exécuter un profil, choisir une option d'échantillonnage et une option d'exploration.
- Ajoutez, modifiez ou supprimez des filtres et des règles.

Dans les résultats de profil, vous pouvez ajouter des commentaires et des balises à un profil et aux colonnes d'un profil. Vous pouvez attribuer des termes d'entreprise aux colonnes.

Le référentiel modèle verrouille les profils pour empêcher les utilisateurs de remplacer des données avec les verrous du référentiel de profil. Le système de contrôle de version enregistre plusieurs versions d'un profil et attribue un numéro à chaque version. Vous pouvez extraire un profil puis l'archiver après avoir fait des modifications. Vous pouvez annuler l'extraction d'un profil avant de l'archiver à nouveau.

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.

Utilisez le service de planificateur pour planifier l'exécution des profils et des fiches d'évaluation à un moment ou à un intervalle spécifique. Le service de planificateur gère les planifications pour les profils, les fiches d'évaluation, les mappages déployés et les flux de travail déployés. Vous pouvez créer, gérer et exécuter les planifications dans Informatica Administrator.

Options de profil de colonne

Lorsque vous créez un profil, vous pouvez utiliser l'assistant de création de profil pour définir des filtres, des règles, des options d'exploration, des options d'échantillonnage et la connexion. Ces options déterminent la manière dont le profil lit les lignes à partir des données source.

Vous pouvez définir les options suivantes dans un profil de colonne, un profil de découverte de domaines de données ou un profil de découverte des données d'entreprise :

- Filtres. Vous pouvez créer des filtres et les appliquer à un profil.
- Règles. Vous pouvez ajouter des règles après avoir créé un profil. Vous pouvez réutiliser les règles que vous créez dans l'outil Analyst tool ou Developer tool.
- Options d'exploration. Vous pouvez choisir de lire les données actuelles de la source de données ou lire les données de profil stockées dans l'entrepôt de profilage.
- Options d'échantillonnage. Vous pouvez choisir l'une des options d'échantillonnage pour déterminer le nombre de lignes sur lesquelles exécuter un profil.
- Connexion. Vous pouvez exécuter les profils dans l'environnement d'exécution natif.

Verrous de profil de référentiel et gestion de profil avec versions

Le référentiel modèle verrouille les profils pour empêcher les utilisateurs de remplacer des données. Lorsque vous commencez à modifier un profil, celui-ci est verrouillé afin d'empêcher d'autres utilisateurs d'enregistrer des modifications. Le verrouillage est désactivé lorsque vous enregistrez le profil. La gestion de profil avec version crée des versions d'un profil et vous permet d'afficher l'historique des versions.

Le référentiel modèle verrouille un profil lorsque vous le modifiez dans l'outil Developer tool ou Analyst tool. Si l'outil s'arrête inopinément, le verrouillage est conservé. Ainsi, lorsque vous vous reconnectez au référentiel modèle, vous pouvez afficher les profils que vous avez verrouillés. Vous pouvez continuer à les modifier ou à les déverrouiller.

Lorsque le référentiel modèle est intégré à un système de contrôle de version, vous pouvez gérer les versions d'un profil. Par exemple, vous pouvez extraire et archiver des profils, annuler des extractions, afficher des versions d'historiques spécifiques et afficher les profils que vous avez extraits. Pour obtenir plus d'informations sur les verrous des ressources du référentiel et la gestion des ressources avec version dans l'outil Analyst tool, consultez le *Guide de l'outil Analyst tool*. Pour obtenir plus d'informations sur les verrous d'objets du référentiel et sur la gestion d'objets avec version dans l'outil Developer tool, consultez le *Guide de l'outil Developer tool*.

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.

Vous pouvez créer un profil avec une option d'échantillonnage et des filtres pour effectuer la découverte des domaines de données. Lorsque vous exécutez le profil, vous appliquez l'option d'échantillonnage et les filtres sur la source de données et générez un ensemble de données. Le processus de découverte des domaines de données utilise l'ensemble de données pour découvrir les domaines de données.

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 été 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 choisir le pourcentage minimal de lignes source ou le nombre minimal de lignes source comme critères de conformité pour la correspondance de domaine de données. Vous pouvez également exclure les valeurs Null lorsque vous effectuez la découverte de domaines de données dans un profil de colonne.

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 :

- Types 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 par un profil dans le cadre de son exécution. Par exemple, vous pouvez approuver ou rejeter les types de données inférés dans les résultats de profil de colonne et dans les résultats 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. Vous pouvez conserver les

résultats de profil pour mettre des informations de profil précises à la disposition de la recherche de découverte et du processus postérieur de 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 l'ensemble de données complet. 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 lorsqu'une partie des données présentes dans une colonne correspondent à d'autres domaines de données. Par exemple, un numéro de téléphone à 10 chiffres auquel il manquerait un chiffre peut être assimilé au modèle du numéro de sécurité sociale (9 chiffres). Cela indique d'éventuels problèmes au niveau de la qualité des données présentes dans une colonne ou un modèle correspondant dans plusieurs domaines de données. Dans ce cas, l'outil Developer peut inférer à la fois le domaine de données du numéro de téléphone et celui du numéro de sécurité sociale. Vous pouvez conserver les résultats de profil afin de pouvoir sélectionner le domaine de données le plus approprié et l'approuver. Dans l'exemple, le domaine de numéro de téléphone est pertinent ; en effet, l'inférence du domaine de données du numéro de sécurité sociale s'est produite en raison de la qualité médiocre des données.

Lors de l'exécution de la découverte des données d'entreprise, l'outil Developer tool peut inférer plusieurs types de données, tels que Date, Chaîne et Varchar pour une colonne de date. En tant qu'architecte de données, vous pouvez choisir d'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. Certaines de ces relations peuvent être indésirables pour les clés candidates découvertes. Par exemple, il est possible que l'outil Developer infère des colonnes représentant une séquence en tant que clés candidates 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 inférés 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 de certaines préférences concernant les métadonnées, telles que les types de données ou les 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 de certaines préférences concernant les métadonnées, telles que les types de données ou les 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](#)
- [Règles dans Informatica Analyst, 45](#)
- [Filtres dans Informatica Analyst, 50](#)
- [Résultats de profil de colonne dans Informatica Analyst, 55](#)
- [Termes d'entreprise, commentaires et balises dans Informatica Analyst, 79](#)
- [Fiches d'évaluation dans Informatica Analyst, 83](#)
- [Découverte des domaines de données dans Informatica Analyst, 109](#)
- [Découverte des données d'entreprise dans Informatica Analyst, 121](#)
- [Résultats de la découverte des données d'entreprise dans Informatica Analyst, 127](#)
- [Recherche de découverte dans Informatica Analyst, 132](#)
- [Bureau de glossaire métier dans Informatica Analyst, 142](#)

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](#)
- [Environnement d'exécution, 39](#)
- [Présentation des profils de système d'exploitation dans Informatica Analyst, 39](#)
- [Présentation des verrous des ressources du référentiel et du développement basé sur l'équipe, 39](#)
- [Création d'un profil de colonne dans Informatica Analyst, 40](#)
- [Modification d'un profil de colonne, 42](#)
- [Exécution d'un profil, 42](#)
- [Option de synchronisation, 43](#)

Présentation des profils de colonne dans Informatica Analyst

Lorsque vous créez un profil, vous sélectionnez les colonnes de l'objet de données sur lesquelles vous souhaitez l'exécuter. Vous pouvez configurer les options d'échantillonnage et d'exploration pour accélérer le profilage. Vous pouvez choisir un environnement d'exécution. Lorsque vous créez un profil, vous pouvez y ajouter des règles et des filtres. 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 au maximum 1 000 colonnes. Lorsque vous créez ou exécutez un profil, vous avez la possibilité de sélectionner toutes les colonnes ou seulement certaines colonnes. Vous pouvez sélectionner toutes les colonnes pour explorer et afficher les fréquences de valeur pour ces colonnes.

Vous pouvez créer des profils de colonne en utilisant les méthodes suivantes dans Informatica Analyst :

- Pour créer un profil, cliquez avec le bouton droit de la souris sur l'objet de données dans l'espace de travail **Bibliothèque**.
- Pour créer un profil par défaut, utilisez les options par défaut.

- Pour créer un profil personnalisé, personnalisez les paramètres du profil.

Remarque: Vous pouvez afficher et exécuter le profil sur des sources de données Avro, JSON, Parquet et XML. Vous pouvez créer et modifier un profil de colonne sur des sources de données Avro, JSON, Parquet et XML dans Informatica Developer.

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 options d'échantillonnage, les options d'exploration et l'environnement d'exécution.

Les étapes suivantes décrivent le processus de profilage de colonne :

1. Choisissez un nom, une description et un emplacement pour le profil de colonne.
2. Sélectionnez une source externe ou un objet de données importé sur lequel vous souhaitez exécuter le profil.
3. Vous pouvez éventuellement prévisualiser les données source.
4. Sélectionnez les colonnes sur lesquelles vous souhaitez exécuter le profil.
5. Déterminez si vous souhaitez conserver ou modifier les options par défaut pour créer le profil. Les options que vous pouvez configurer comprennent les options d'échantillonnage, les options d'exploration et l'environnement d'exécution.
6. Vous pouvez éventuellement ajouter des règles et des filtres lors de la création du profil.
7. 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 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 tool change les types de données Date/heure, Numérique et Décimal en fonction des paramètres régionaux de votre navigateur.
- Tri des données multilingues. Vous pouvez trier des données multilingues. L'outil Analyst tool 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 les options d'échantillonnage et les options d'exploration des données. Vous pouvez configurer ces options lorsque vous créez ou modifiez un profil de colonne pour un objet de données.

Vous pouvez configurer les options de profil dans l'espace de travail **Découverte**. Lors de la création d'un profil, vous pouvez conserver les options de colonne, d'échantillonnage et d'exploration par défaut. Utilisez l'option d'exploration pour choisir les données en direct ou intermédiaires.

Options d'échantillonnage

Les options d'échantillonnage déterminent le nombre de lignes sur lesquels l'outil Analyst tool 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 de l'objet de données.
Échantillon sur les <nombre> premières lignes	Nombre de lignes pour lesquelles vous souhaitez exécuter le profil. L'outil Analyst tool choisit les lignes parmi les premières lignes de la source.
Échantillon aléatoire de <nombre> lignes	Nombre de lignes d'un échantillon aléatoire pour lequel vous souhaitez exécuter le profil. L'échantillonnage aléatoire force l'outil Analyst tool à explorer les données intermédiaires. Notez que cette option peut affecter les performances d'exploration.
Échantillon aléatoire	Taille de l'échantillon aléatoire basée sur le nombre de lignes dans l'objet de données. L'échantillonnage aléatoire force l'outil Analyst tool à explorer les données intermédiaires. Notez que cette option peut affecter les performances d'exploration.
Excluez les types et les domaines de données approuvés de l'inférence du type et du domaine de données dans les exécutions de profil suivantes.	Exclut le type de données approuvé ou le domaine de données du type de données et l'inférence de domaine de données de l'exécution de profil suivante.

Options de développement

Vous pouvez configurer les options d'exploration lors de la définition ou de la modification d'un profil.

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

Options	Description
En direct	Explore les données en direct pour lire les données actuelles de la source de données.
Stocké	Explore les données intermédiaires pour lire les données de profil intermédiaires de l'entrepôt de profilage.
Sélectionner des colonnes	Identifie les colonnes d'exploration que vous n'avez pas sélectionnées pour le profilage.

Environnement d'exécution

Choisissez l'environnement d'exécution Natif pour un profil de colonne. Informatica Analyst définit l'environnement d'exécution dans la définition de profil une fois que vous avez choisi un environnement d'exécution.

Lorsque vous exécutez un profil dans un environnement d'exécution Natif, l'outil Analyst tool soumet les tâches de profil au module de service de profilage. Celui-ci divise alors les tâches de profil en un ensemble de mappages. Le service d'intégration de données exécute ces mappages sur la même machine où il exécute et écrit les résultats de profil dans l'entrepôt de profilage. Par défaut, tous les profils s'exécutent dans un environnement d'exécution natif.

Vous pouvez utiliser des sources natives pour créer et exécuter des profils dans l'environnement natif. Une source de données native est une source non-Hadoop, comme un fichier plat, une source relationnelle ou une source principale. Vous pouvez également exécuter un profil sur une spécification de mappage ou une source de données logique avec une source de données Hive ou HDFS dans l'environnement natif.

Présentation des profils de système d'exploitation dans Informatica Analyst

Vous pouvez choisir un profil de système d'exploitation dans l'outil Analyst tool. Après avoir choisi un profil de système d'exploitation, le service d'intégration de données crée et exécute les profils de colonne, les profils de découverte de données d'entreprise et les fiches d'évaluation en fonction des autorisations de l'utilisateur du profil de système d'exploitation.

L'outil Analyst tool utilise le profil par défaut pour exécuter les profils et les fiches d'évaluation. Si vous disposez uniquement d'un profil de système d'exploitation, il est sélectionné par défaut. Si vous disposez de plusieurs profils de système d'exploitation, vous pouvez choisir l'un des profils du système d'exploitation.

Sélection d'un profil de système d'exploitation

Vous pouvez sélectionner un profil de système d'exploitation dans Informatica Analyst. Le service d'intégration de données utilise les autorisations de l'utilisateur du profil de système d'exploitation pour exécuter les tâches de profilage.

1. Dans La zone d'en-tête d'Informatica Analyst, cliquez sur **<nom d'utilisateur> > Paramètres**.
La boîte de dialogue **Paramètres** s'affiche.
2. Sélectionnez un profil de système d'exploitation. Cliquez sur **Enregistrer**.

Présentation des verrous des ressources du référentiel et du développement basé sur l'équipe

Le référentiel modèle verrouille les profils pour empêcher les utilisateurs d'écraser le travail des autres développeurs. Si le référentiel modèle est intégré à un système de contrôle de version, il enregistre plusieurs

versions des ressources et attribue un numéro à chaque version. Vous pouvez extraire et archiver des profils et annuler des extractions. Vous pouvez afficher une version spécifique d'un profil que vous avez extrait.

Lorsque vous commencez à modifier un profil dans l'outil Analyst tool, le référentiel modèle le verrouille afin d'empêcher les autres utilisateurs de le modifier. Lorsque vous enregistrez le profil, vous conservez le verrou. Lorsque vous fermez le profil, le référentiel modèle le déverrouille.

Le référentiel modèle empêche d'autres membres de l'équipe de remplacer les profils grâce à la gestion des ressources avec version. Lorsque vous tentez de modifier un profil qu'un autre membre a extrait, vous recevez un message indiquant quel utilisateur a extrait le profil. Vous pouvez ouvrir un profil extrait en mode de lecture seule ou enregistrer le profil sous un autre nom.

Vous pouvez sélectionner une version du profil dans la boîte de dialogue Propriétés de profil pour afficher la définition de profil pour cette version. Vous pouvez accéder à l'option Propriétés de profil dans le menu Actions. Pour obtenir plus d'informations sur les verrous des ressources du référentiel et sur la gestion des ressources avec version, consultez le *Guide de l'outil Analyst tool*.

Création d'un profil de colonne dans Informatica Analyst

Vous pouvez créer un profil personnalisé ou un profil par défaut. Lors de la création d'un profil personnalisé, vous pouvez configurer les options des colonnes, des lignes d'échantillonnage et d'exploration. Lors de la création d'un profil par défaut, le profil de colonne et la découverte de domaines de données s'exécutent sur l'ensemble de données complet avec tous les domaines de données.

1. Dans l'espace de travail **Découverte**, cliquez sur **Profil** ou sélectionnez **Nouveau > profil** dans la zone d'en-tête.

Remarque: Vous pouvez cliquer avec le bouton droit sur l'objet de données dans l'espace de travail **Bibliothèque** et créer un profil. Dans ce profil, le nom de profil, le nom d'emplacement et l'objet de données sont extraits des propriétés d'objet de données. Vous pouvez créer un profil par défaut ou personnaliser les paramètres pour créer un profil personnalisé.

L'assistant **Nouveau profil** s'affiche.

2. L'option **Source unique** est sélectionnée par défaut. Cliquez sur **Suivant**.
3. Dans l'écran **Spécifier les propriétés générales**, entrez un nom et une description facultative pour le profil. Dans le champ Emplacement, sélectionnez le projet ou le dossier dans lequel vous souhaitez créer le profil. Cliquez sur **Suivant**.
4. Dans l'écran **Sélectionner une source**, cliquez sur **Choisir** pour sélectionner un objet de données, ou sur **Nouveau** pour importer un objet de données. Cliquez sur **Suivant**.
 - Dans la boîte de dialogue **Choisir un objet de données**, sélectionnez un objet de données. Cliquez sur **OK**.
Le volet Propriétés affiche les propriétés de l'objet de données sélectionné. Le volet Aperçu des données affiche les colonnes de l'objet de données.
 - Dans la boîte de dialogue **Nouvel objet de données**, vous pouvez choisir une connexion, un schéma, une table ou une vue pour créer un profil, sélectionner un emplacement et créer un dossier pour importer l'objet de données. Cliquez sur **OK**.
5. Dans l'écran **Sélectionner une source**, sélectionnez les colonnes sur lesquelles vous souhaitez exécuter un profil. Sélectionnez éventuellement **Nom** pour sélectionner toutes les colonnes. Cliquez sur **Suivant**.

Toutes les colonnes sont sélectionnées par défaut. L'outil Analyst tool répertorie les propriétés de colonne, telles que le nom, le type de données, la précision, l'échelle, l'autorisation de la valeur Null et les participations à la clé primaire de chaque colonne.

6. Dans l'écran **Spécifier les paramètres**, choisissez d'exécuter un profil de colonne, une découverte de domaines de données ou un profil de colonne et une découverte de domaines de données. Par défaut, l'option de profil de colonne est sélectionnée.

- Choisissez **Exécuter le profil de colonne** pour exécuter un profil de colonne.
- Choisissez **Exécuter la découverte de domaines de données** pour effectuer la découverte de domaines de données. Dans le volet **Domaine de données**, sélectionnez les domaines de données que vous souhaitez découvrir, sélectionnez un critère de conformité, puis sélectionnez les colonnes pour la découverte de domaines de données dans la boîte de dialogue **Modifier la sélection des colonnes pour la découverte des données Domin**.
- Choisissez **Exécuter le profil de colonne** et **Exécuter la découverte de domaines de données** pour exécuter le profil de colonne et la découverte de domaines de données. Sélectionnez les options de domaine de données dans le volet **Domaine de données**.

Remarque: Par défaut, les colonnes que vous avez sélectionnées sont destinées au profil de colonne et à la découverte de domaines de données. Cliquez sur **Modifier** pour sélectionner ou désélectionner les colonnes de la découverte de domaines de données.

- Choisissez les données et/ou les colonnes sur lesquelles vous souhaitez exécuter la découverte de domaines de données.
- Choisissez une option d'échantillonnage. Vous pouvez choisir **Toutes les lignes (analyse complète)**, **Échantillon 1re ligne**, **Échantillon aléatoire** ou **Échantillon aléatoire (auto)** comme option d'échantillonnage dans le volet **Exécuter le profil sur**. Cette option s'applique sur le profil de colonne et la découverte de domaines de données.
- Choisissez une option d'exploration. Vous pouvez choisir l'option d'exploration **En direct** ou **Stocké**, ou bien choisir **Désactivé** pour désactiver l'exploration dans le volet **Exploration**. Cliquez éventuellement sur **Sélectionner les colonnes** pour sélectionner les colonnes à explorer. Vous pouvez choisir d'exclure l'inférence du type de données et du domaine de données pour les colonnes associées à un type de données ou à un domaine de données approuvé.
- Choisissez l'option d'environnement d'exécution **Natif**.

Remarque: Vous ne pouvez pas exécuter de profils sur le moteur Hive ou Blaze dans l'environnement d'exécution Hadoop, car l'environnement Hadoop n'est pas pris en charge dans la version actuelle d'Informatica.

7. Cliquez sur **Suivant**.

L'écran **Spécifiez les règles et les filtres** s'ouvre.

8. Dans l'écran **Spécifiez les règles et les filtres**, vous pouvez effectuer les tâches suivantes :

- Créer, modifier ou supprimer une règle. Vous pouvez appliquer des règles existantes au profil.
- Créer, modifier ou supprimer un filtre.

Remarque: Lorsque vous créez une fiche d'évaluation pour ce profil, vous pouvez réutiliser les filtres que vous avez créés pour le profil.

9. Cliquez sur **Enregistrer et terminer** pour créer le profil, ou sur **Enregistrer et exécuter** pour créer et exécuter le profil.

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, ou sélectionnez le profil dans le volet **Ressources**.
2. Cliquez sur le nom du profil.
La vue de résumé s'affiche dans l'espace de travail **Découverte**.
3. Si le système de contrôle de version est activé, cliquez sur **Actions** > **Extraire** pour extraire le profil.
4. Cliquez sur **Actions** > **Modifier le profil**.
L'assistant de **profil** s'affiche.
5. Sélectionnez l'une des options suivantes selon les modifications que vous souhaitez effectuer :
 - **Spécifiez les propriétés générales.** Modifiez les propriétés de base telles que le nom, la description et l'emplacement.
 - **Sélectionnez une source.** Choisissez une autre source de données correspondante et des colonnes sur lesquelles exécuter le profil.
 - **Spécifiez les paramètres.** Choisissez d'exécuter le profil de colonne uniquement ou le profil de colonne avec la découverte de domaines de données. Sélectionnez les domaines de données que vous voulez découvrir et modifiez les options de découverte de domaines de données, d'échantillonnage et d'exploration.
 - **Spécifiez les règles et les filtres.** Créez, modifiez ou supprimez des règles et des filtres.
6. Cliquez sur **Enregistrer et terminer** pour modifier le profil ou sur **Enregistrer et exécuter** pour modifier et exécuter le profil.
7. Si le système de contrôle de version est activé, vous devez exécuter les tâches suivantes :
 - Cliquez sur **Enregistrer et terminer** pour terminer la modification du profil.
 - Dans la vue de résumé, cliquez sur **Archiver** pour archiver le profil.
 - Cliquez sur **Actions** > **Exécuter le profil** pour 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 sur une colonne ou une règle uniquement, sans l'exécuter sur toutes les colonnes source après l'exécution initiale.

1. Dans l'espace de travail **Bibliothèque**, sélectionnez le projet ou le dossier contenant le profil dans le volet Projets, ou sélectionnez le profil dans le volet **Ressources**.
2. Cliquez sur **Actions** > **Ouvrir**.
La vue de résumé s'ouvre dans l'espace de travail **Découverte**.
3. Cliquez sur **Actions** > **Exécuter le profil**.
L'outil Analyst tool lance une exécution de profil et affiche les résultats dans la vue de résumé.
4. Cliquez sur une colonne de la vue de résumé pour afficher les résultats de colonne.
La vue détaillée s'affiche.

Option de synchronisation

Lorsque vous modifiez les métadonnées d'une source de données externe, les métadonnées de l'objet de données dans le référentiel de modèles ne sont pas mises à jour par défaut. Utilisez l'option de synchronisation pour synchroniser les métadonnées de l'objet de données avec les métadonnées de la source de données. Vous pouvez utiliser l'option de synchronisation pour les profils de colonnes, les profils de découverte d'entreprise et les fiches d'évaluation. La source de données externe peut être une source de données relationnelle ou une source de données de fichier plat.

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 a 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

Règles dans Informatica Analyst

Ce chapitre comprend les rubriques suivantes :

- [Présentation des règles dans Informatica Analyst, 45](#)
- [Règles prédéfinies, 46](#)
- [Règles d'expression, 47](#)

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 valider les données.

Vous pouvez utiliser une règle dans différentes situations. 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.

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 tool. 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 tool. Les règles qu'un développeur crée dans l'outil Developer tool comme mapplets peuvent s'afficher dans l'outil Analyst tool en tant que règles réutilisables.

Règles d'un profil de colonne

Vous pouvez créer des règles lors de la création d'un profil de colonne.

Après avoir ajouté une règle à un profil, vous pouvez exécuter à nouveau le profil pour la colonne de règle. L'outil Analyst tool affiche les résultats de profil de la colonne de règle dans la vue de résumé. Vous pouvez afficher les résultats de colonne d'une règle dans la vue détaillée. La sortie d'une règle peut se composer d'une ou de plusieurs colonnes virtuelles. Les résultats de profil incluent les colonnes virtuelles. L'outil Analyst tool exécute un profil sur les colonnes virtuelles. Par exemple, vous devez utiliser une règle prédéfinie qui divise une colonne contenant le prénom et le nom en colonnes virtuelles FIRST_NAME et LAST_NAME. L'outil Analyst tool exécute le profil sur 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 tool 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

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 l'espace de travail **Bibliothèque**, sélectionnez le projet qui contient le profil, ou sélectionnez le profil dans le volet **Ressources**.
2. Cliquez sur **Actions > Ouvrir** pour ouvrir le profil.
La vue de résumé s'affiche dans l'espace de travail **Découverte**.
3. Cliquez sur **Actions > Modifier le profil**.
L'**assistant de profil** s'affiche.
4. Cliquez sur **Spécifiez les règles et les filtres**.
5. Dans l'écran **Spécifiez les règles et les filtres**, cliquez sur **Actions > Appliquer une règle existante** dans le panneau **Règles**.
La boîte de dialogue **Assistant d'application de règle** s'affiche.
6. Sélectionnez une règle et cliquez sur **Suivant**.
7. Cliquez sur **Ajouter**.
La boîte de dialogue **Sélectionner les colonnes pour les ports d'entrée** s'affiche.

8. Sélectionnez un champ et une colonne d'entrée. Cliquez sur **OK**.
La boîte de dialogue **Assistant d'application de règle** affiche les colonnes d'entrée et de sortie.
9. Cliquez sur **OK** dans la boîte de dialogue **Assistant d'application de règle**.
La règle s'affiche sur l'écran **Spécifiez les règles et les filtres**.

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 de **profil** pour créer une règle d'expression et l'ajouter à un profil.

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.

Après avoir créé et validé une règle d'expression, vous pouvez modifier la valeur de précision de la colonne de règle de sortie. Par défaut, la valeur de précision de la colonne de règle de sortie est définie sur 10. La valeur de précision est tronquée lorsque la colonne de règle de sortie dépasse la valeur de précision définie.

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 tool 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 de **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. Ouvrez un profil.
2. Dans la vue de résumé, cliquez sur **Actions > Modifier le profil** pour ouvrir l'assistant de **profil**.
3. Cliquez sur **Spécifiez les règles et les filtres**.
4. Dans le volet Règles, cliquez sur **Actions > Ajouter une règle**.

La boîte de dialogue **Nouvelle règle** s'affiche.

5. Dans la boîte de dialogue **Nouvelle règle**, entrez le nom, et éventuellement la description, de la règle. Vous pouvez créer une règle dans le panneau Fonctions ou Colonnes.
 - Dans le panneau Fonctions, sélectionnez une catégorie de fonction, puis cliquez sur les flèches vers la droite (>>). Dans la boîte de dialogue, spécifiez les paramètres et cliquez sur **OK**. La fonction s'affiche avec les colonnes et les valeurs dans le panneau Expression.
 - Dans le panneau Colonnes, sélectionnez une colonne, puis cliquez sur les flèches vers la droite (>>). La colonne s'affiche dans le panneau Expression. Ajoutez des fonctions, des expressions et des valeurs pour créer une règle.
6. Pour vérifier la règle, cliquez sur **Valider**.

7. Vous pouvez éventuellement choisir de définir la règle comme règle réutilisable et de configurer l'emplacement du projet et du dossier. Si vous transformez une règle en règle réutilisable, vous-même et les autres utilisateurs pourrez l'utiliser comme règle prédéfinie dans un autre profil.
8. Cliquez sur **OK**.
L'écran **Spécifiez les règles et les filtres** s'affiche avec la règle dans le volet Règles.

CHAPITRE 8

Filtres dans Informatica Analyst

Ce chapitre comprend les rubriques suivantes :

- [Présentation des filtres dans Informatica Analyst, 50](#)
- [Création d'un filtre, 50](#)
- [Gestion des filtres, 53](#)

Présentation des filtres dans Informatica Analyst

Vous pouvez créer un filtre pour constituer un sous-ensemble de la source de données d'origine correspondant aux critères de filtre. Vous pouvez ensuite exécuter un profil sur les données filtrées.

Vous pouvez créer un filtre pour afficher les résultats de profil qui répondent aux critères du filtre. Vous pouvez afficher les résultats de profil avec les filtres par défaut disponibles dans la vue de résumé.

Création d'un filtre

Vous pouvez créer un filtre pour constituer un sous-ensemble de la source de données d'origine correspondant aux critères de filtre.

1. Ouvrez un profil.
2. Dans la vue de résumé, cliquez sur **Actions > Modifier le profil**.
L'assistant de **profil** s'affiche.
3. Cliquez sur **Spécifiez les règles et les filtres**.
4. Dans le volet **Filtres**, cliquez sur **Actions > Ajouter un filtre**.
La boîte de dialogue **Nouveau filtre** s'affiche.
5. Créez un filtre simple, avancé ou SQL.

Remarque: Pour créer un filtre simple ou avancé sur une colonne de date, indiquez la condition au format AAAA/MM/JJ HH:MM:SS.

Le volet **Aperçu des données** affiche le sous-ensemble de la source de données d'origine répondant aux critères de filtre.

6. Cliquez sur **OK**.
L'écran **Spécifiez les règles et les filtres** s'affiche le filtre dans le volet **Filtres**.

Création d'un filtre simple

Vous pouvez créer un filtre simple à l'aide d'opérateurs conditionnels, tels que =, !=, > ou <. Utilisez le filtre pour créer un sous-ensemble de la source de données d'origine.

1. Dans la boîte de dialogue **Nouveau filtre**, cliquez sur **Simple**.

L'image suivante présente les options que vous pouvez utiliser pour créer un filtre simple dans la boîte de dialogue **Nouveau filtre** :

New Filter

Create a filter. The filter is used to create a subset of the data rows before profiling.

Name*:

Description:

Choose the filter type*: ☒ Simple ☐ Advanced ☐ SQL

Columns	Operator	Values(s)
<input type="text" value="-Select-"/>	<input type="text" value="-Select-"/>	<input type="text" value=""/>

Filter Preview

2. Entrez un nom et éventuellement une description.
3. Sélectionnez une colonne.
4. Sélectionnez un opérateur conditionnel.
5. Entrez une valeur.
6. Vous pouvez éventuellement cliquer sur l'icône plus (+) pour ajouter d'autres filtres.
7. Cliquez sur **OK**.

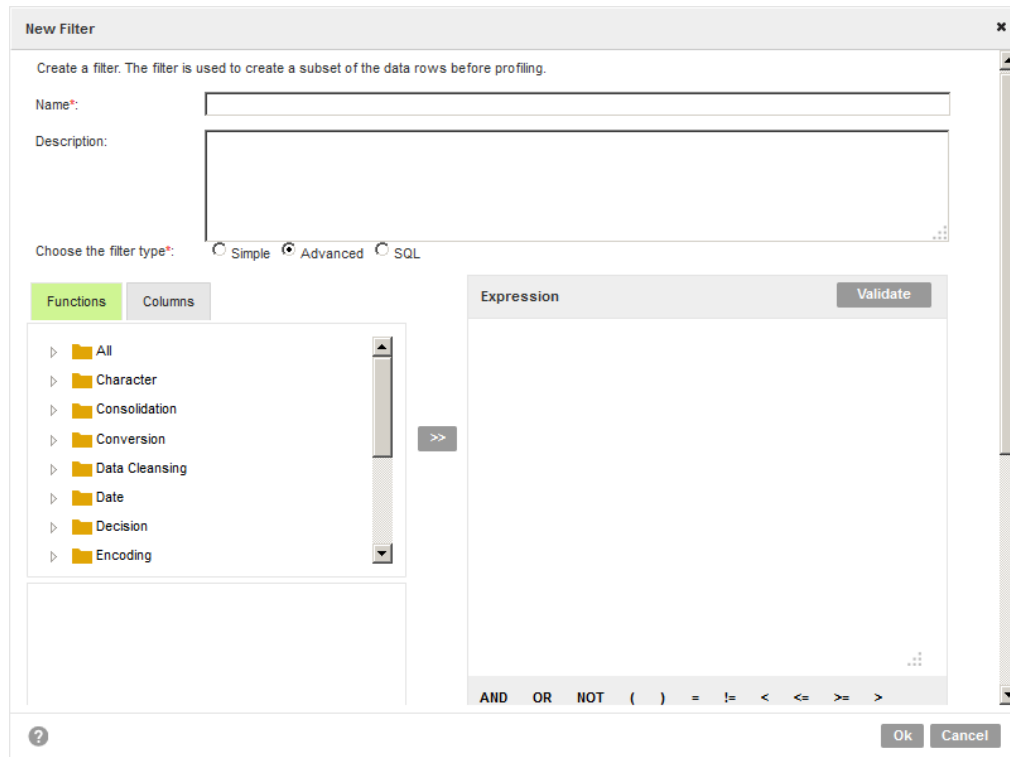
La page **Spécifiez les règles et les filtres** s'affiche avec le filtre dans le volet Filtres.

Création d'un filtre avancé

Vous pouvez créer un filtre avancé avec des expressions comme AND, OR et NOT pour créer un sous-ensemble de la source de données d'origine.

1. Dans la boîte de dialogue **Nouveau filtre**, cliquez sur **Avancé**.

L'image suivante présente les options de filtre avancé de la boîte de dialogue **Nouveau filtre**.



2. Entrez le nom, et éventuellement la description, du filtre avancé.
3. Vous pouvez créer un filtre avancé à l'aide du panneau Fonctions ou Colonnes.
 - Dans le panneau Fonctions, sélectionnez une catégorie de fonction, puis cliquez sur les flèches vers la droite (>>). Dans la boîte de dialogue, spécifiez les paramètres et cliquez sur **OK**. La fonction s'affiche avec les colonnes et les valeurs dans le panneau Expression.
 - Dans le panneau Colonnes, sélectionnez une colonne, puis cliquez sur les flèches vers la droite (>>). La colonne s'affiche dans le panneau Expression. Ajoutez des fonctions, des expressions et des valeurs pour créer un filtre avancé.
4. Pour vérifier le filtre avancé, cliquez sur **Valider**.
5. Cliquez sur **OK**.

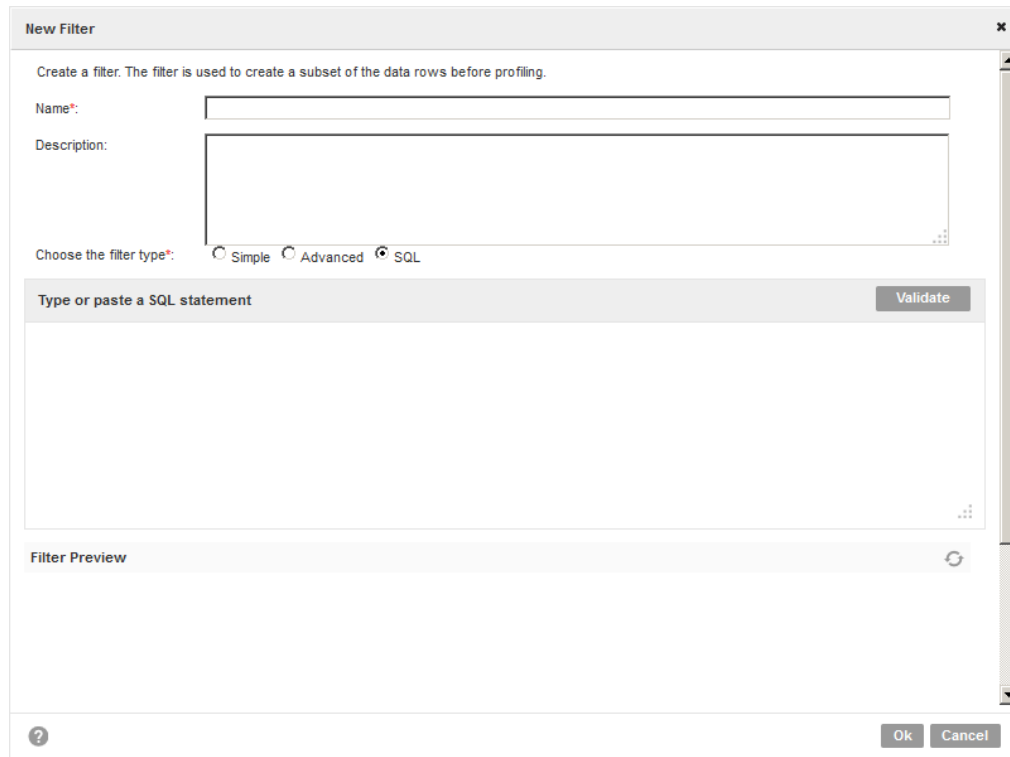
L'écran **Spécifiez les règles et les filtres** s'affiche avec le filtre dans le volet Filtres.

Création d'un filtre SQL

Vous pouvez créer un filtre SQL avec des requêtes SQL. Vous pouvez créer un filtre SQL pour les sources de données relationnelles.

1. Dans la boîte de dialogue **Nouveau filtre**, cliquez sur **SQL**.

L'image suivante présente les options de filtre SQL de la boîte de dialogue **Nouveau filtre** :



2. Entrez le nom, et éventuellement la description, du filtre SQL.
3. Dans la zone de texte, entrez ou collez une requête SQL.
4. Cliquez sur **Valider** pour vérifier la requête SQL.
5. Cliquez sur **OK**.

La page **Spécifiez les règles et les filtres** s'affiche avec le filtre SQL dans le volet Filtres.

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, ou sélectionnez dans le volet **Ressources** le profil à filtrer.
2. Ouvrez un profil.
3. Dans la vue de résumé, cliquez sur **Actions > Modifier le profil** pour ouvrir l'assistant de **profil**.
4. Cliquez sur **Spécifiez les règles et les filtres**.
5. Dans le volet Filtres, sélectionnez un filtre et cliquez sur **Actions > Modifier le filtre**.

La boîte de dialogue **Modifier le filtre** s'affiche.

6. Modifiez les paramètres de filtre et cliquez sur **OK**.
7. Pour supprimer un filtre, sélectionnez-le et cliquez sur **Actions > Supprimer un filtre**.

CHAPITRE 9

Résultats de profil de colonne dans Informatica Analyst

Ce chapitre comprend les rubriques suivantes :

- [Présentation des résultats de profil de colonne dans Informatica Analyst, 55](#)
- [Vue de résumé, 56](#)
- [Vue détaillée, 58](#)
- [Statistiques, 60](#)
- [Types d'exécution de profil, 67](#)
- [Présentation des résultats de la comparaison de plusieurs profils, 69](#)
- [Développement de profil de colonne, 74](#)
- [Conservation dans l'outil Analyst, 75](#)
- [Fichiers d'exportation de profil de colonne dans Informatica Analyst, 76](#)

Présentation des résultats de profil de colonne dans Informatica Analyst

Affichez les résultats de profil pour comprendre et analyser le contenu, la structure et la qualité des données. Vous pouvez afficher l'intégralité des colonnes et des règles d'un profil dans la vue de résumé. Vous pouvez afficher les détails des propriétés d'une colonne ou d'une règle dans la vue détaillée.

Vous pouvez afficher les résultats de profil dans l'espace de travail **Découverte**. L'en-tête de la vue affiche le type, le nombre de colonnes, le nombre de règles, les données d'échantillonnage ainsi que la date et l'heure de création du profil.

La vue de résumé vous permet d'afficher les propriétés de chaque colonne sous la forme d'une valeur, d'un graphique à barres horizontales ou d'un pourcentage. Vous pouvez afficher les propriétés de colonne, telles que les valeurs Null, uniques et non uniques, les modèles, les types de données et les domaines de données. Les filtres par défaut permettent de gérer l'affichage des résultats de profil dans la vue de résumé.

La vue détaillée vous permet d'afficher dans des volets les valeurs Null, uniques et non uniques, les types de données, les domaines de données et les modèles inférés ainsi que les valeurs, les termes d'entreprise et l'aperçu des données.

Vous pouvez afficher les résultats de profil de la dernière exécution, d'une exécution historique ou d'une exécution consolidée. Vous pouvez comparer les résultats de deux exécutions de profil et les afficher dans la

vue de résumé et la vue détaillée. Vous pouvez également afficher des statistiques de profil et organiser les données. Les statistiques de profil comprennent des valeurs, des modèles, des types de données, des valeurs hors norme et des statistiques concernant les colonnes et les règles. Vous pouvez effectuer la découverte de données et explorer les données obtenues.

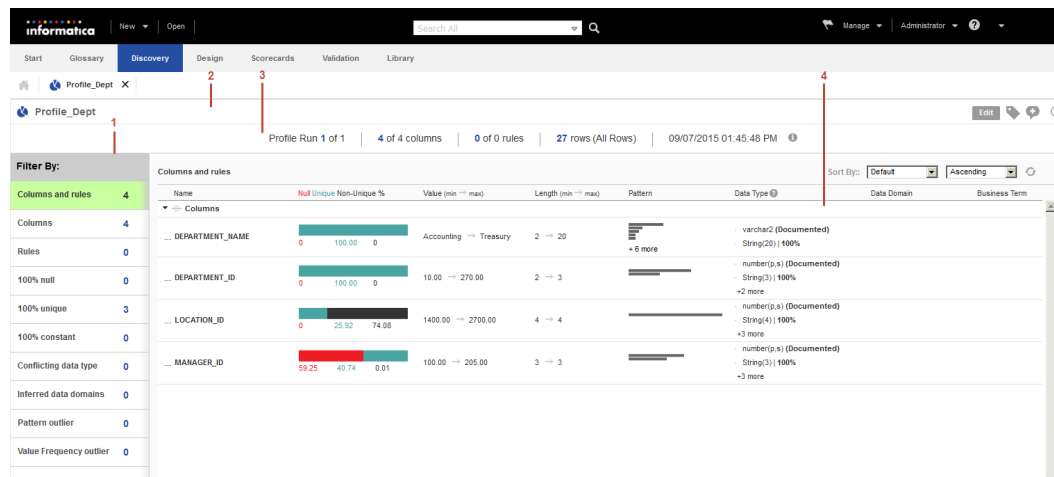
Remarque: Vous pouvez afficher et exécuter un profil sur des sources de données Avro, JSON, Parquet et XML. Vous pouvez afficher les résultats de profil de la dernière exécution, d'une exécution historique ou d'une exécution consolidée et comparer les résultats de deux exécutions de profils.

Vous pouvez exporter des fréquences de valeur, des fréquences de modèle, des données d'exploration, des commentaires, des balises et des termes d'entreprise dans un fichier CSV. Vous pouvez exporter les informations de résumé de profil dans un fichier Microsoft Excel afin d'afficher toutes les données dans un fichier pour une analyse plus poussée. Vous pouvez afficher les informations relatives aux règles dans les résultats de profil. Les résultats de profil qui s'affichent dépendent des options de configuration et d'échantillonnage du profil.

Vue de résumé

Le résumé des résultats de profil s'affiche dans un format de grille dans la vue de résumé. Vous pouvez utiliser les filtres par défaut de la vue de résumé pour afficher des statistiques spécifiques. Par exemple, lorsque vous choisissez Règles, la vue de résumé affiche toutes les règles du profil.

L'image suivante montre un exemple de représentation graphique de la vue de résumé :



1. Filtres par défaut. Les filtres par défaut permettent de voir les résultats de profil dans la vue de résumé.
2. En-tête du profil. Vous pouvez afficher le nom du profil dans l'en-tête. Vous pouvez utiliser le bouton Modifier pour modifier le profil, utiliser les icônes de balise et de commentaires pour ajouter ou modifier des balises et des commentaires et choisir les options du menu Action.
3. En-tête de la vue de résumé. Vous pouvez afficher des informations de profil spécifiques dans l'en-tête de la vue de résumé. Vous pouvez afficher le nombre d'exécutions de profil, le nombre total d'exécutions de profil, le nombre de colonnes et de règles ainsi que le nombre de lignes dans le profil.
4. Vue de résumé. Vous pouvez afficher les propriétés de toutes les colonnes et règles du profil.

Dans la vue de résumé, vous pouvez exécuter ou modifier le profil, détecter des valeurs hors norme de modèle et de fréquence de valeur, ajouter des colonnes à une fiche d'évaluation, choisir l'exécution d'un profil, comparer deux exécutions de profil, exporter les résultats de profil ou de découverte de domaine de

données vers une feuille de calcul Microsoft Excel, vérifier les résultats d'inférence de multiple colonnes, ajouter ou supprimer des commentaires et des balises ou afficher les propriétés de profil.

Propriétés de la vue de résumé

La vue de résumé affiche les propriétés de toutes les colonnes et règles d'un profil. La vue de résumé comprend une représentation visuelle des propriétés. Vous pouvez cliquer sur chaque propriété du résumé pour trier ses valeurs.

Le tableau suivant décrit les propriétés de résumé des résultats de profil :

Propriété	Description
Nom	Affiche le nom de la colonne ou de la règle dans le profil.
% Null % Distinct % Non distinct	Affiche le pourcentage de valeurs Null, distinctes et non distinctes pour une sortie de colonne ou de règle. Vous pouvez afficher les valeurs dans un graphique à barres horizontales.
Modèle	Affiche les différents modèles de la colonne sous forme de graphiques à barres horizontales. Vous pouvez survoler le graphique à barres avec votre souris pour afficher les caractères de modèle et le pourcentage de modèles similaires dans une colonne.
Valeur	Affiche les valeurs minimale et maximale de la sortie de colonne ou de règle.
Longueur	Affiche la longueur minimale et maximale des valeurs de la sortie de colonne ou de règle.
Type de données	<p>Affiche le type de données documenté de la colonne ou de la règle. Affiche les types de données inférés lorsque vous survolez le champ avec votre souris. L' découverte de Analyst tool peut inférer les types de données suivants :</p> <ul style="list-style-type: none">- Chaîne- Varchar- Décimal- Entier- Date <p>Vous pouvez également afficher le pourcentage de conformité selon les types de données inférés.</p> <p>Remarque: l' découverte de Analyst tool ne peut pas dériver le type de données des valeurs d'une colonne numérique dont la précision est supérieure à 38. Il ne peut pas non plus 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 sur laquelle vous créez un profil de colonne avec une valeur d'année antérieure à 1 800, 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.</p>
Domaine de données	Affiche les noms des domaines de données associés à la colonne ainsi que le pourcentage de conformité et le nombre de lignes conformes.
Terme d'entreprise	Affiche le terme d'entreprise attribué à la colonne.

Filtres par défaut de la vue de résumé

Les filtres par défaut permettent de gérer l'affichage des résultats de profil dans la vue de résumé.

Par défaut, la vue de résumé affiche les résultats de profil pour toutes les colonnes source, virtuelles et de règle. Le volet Filtrer par affiche le nombre de colonnes sur lesquelles vous pouvez appliquer les filtres par défaut.

La vue de résumé vous permet d'afficher les résultats de profil à l'aide des options de filtre par défaut suivantes :

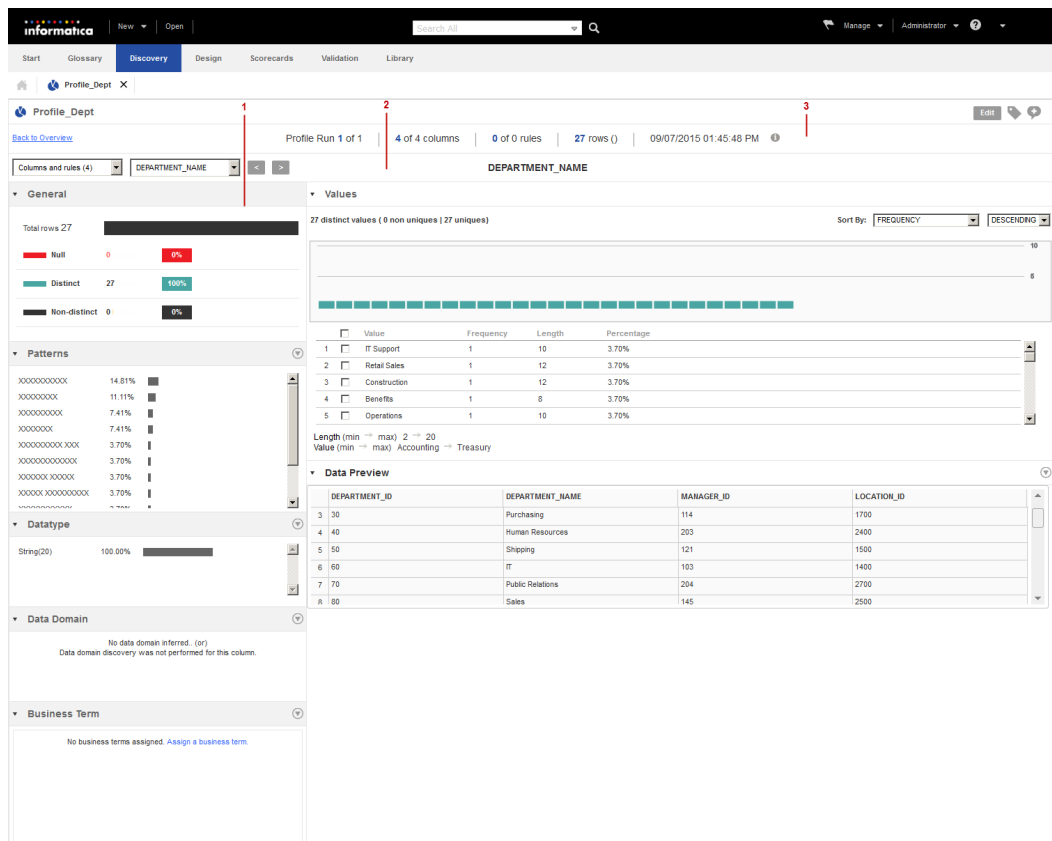
Option de filtre par défaut	Description
Colonnes et règles	Affiche les résultats de profil des colonnes source et de règle. Vous pouvez développer et réduire les colonnes source et de règle pour afficher les résultats.
Colonnes	Affiche les résultats de profil des colonnes source.
Règles	Affiche les résultats de profil des colonnes de règle.
100 % Null	Affiche les résultats de profil des colonnes qui contiennent 100 % de valeurs Null.
100 % distinct	Affiche les résultats de profil des colonnes qui contiennent 100 % de valeurs distinctes.
100 % constante	Affiche les résultats de profil des colonnes dont tous les enregistrements possèdent la même valeur. Par exemple, le filtre 100 % constante comprend les résultats de profil d'une colonne Pays si celle-ci contient uniquement la valeur « USA ».
Types de données en conflit	Affiche les résultats de profil des colonnes pour lesquelles le type de données documenté et le type de données inféré ne correspondent pas. Par exemple, le filtre affiche la colonne CustomerTier, car le type de données documenté pour la colonne est Entier (2) et le type de données inféré, Chaîne.
Domaines de données inférés	Affiche les résultats de profil des colonnes dans lesquelles le domaine de données inféré est identique au domaine de données configuré.
Valeur hors norme du modèle	Affiche les résultats de profil des colonnes qui contiennent des valeurs hors norme de modèle.
Valeur hors norme de la fréquence de la valeur	Affiche les résultats de profil des colonnes qui contiennent des valeurs hors norme de valeur ou de fréquence.

Vue détaillée

Les résultats de colonne s'affichent dans la vue détaillée. Vous pouvez afficher les propriétés de colonne en détail.

La vue détaillée d'une colonne s'affiche lorsque vous cliquez sur la colonne dans la vue de résumé.

L'image suivante montre un exemple de représentation graphique des propriétés de la colonne dans la vue détaillée :



1. Volets. Vous pouvez afficher dans des volets les propriétés générales, les valeurs de la colonne, les modèles inférés, les types de données inférés, les domaines de données inférés et les termes d'entreprise.
2. En-tête des détails de colonne. Vous pouvez afficher les résultats de la colonne en sélectionnant la colonne dans la liste déroulante ou en utilisant les boutons de navigation.
3. En-tête de la vue de résumé. Vous pouvez afficher des informations spécifiques au profil dans l'en-tête de la vue de résumé. Vous pouvez afficher l'exécution de profil, le nombre de colonnes, les règles et les lignes dans une exécution de profil ainsi que l'heure et la date de l'exécution de profil.

Dans la vue détaillée, vous pouvez exécuter ou modifier un profil, ajouter la colonne à une fiche d'évaluation, choisir une exécution de profil, comparer deux exécutions de profil, exporter les résultats de profil vers une feuille de calcul Microsoft Excel, exporter les fréquences de valeur, les fréquences de modèle, les types de données ou les données d'exploration des modèles sélectionnés vers un fichier csv, ajouter ou supprimer des commentaires et des balises dans une colonne et afficher des propriétés de profil.

Utilisez le menu Actions de chaque volet pour effectuer d'autres actions sur les propriétés de colonne. Vous pouvez réduire ou développer les volets.

Volets de la vue détaillée

La vue détaillée affiche dans des volets les propriétés de colonne, telles que le nombre et le pourcentage de valeurs distinctes, non distinctes et Null, les modèles, les types de données et domaines de données inférés, les valeurs, l'aperçu des données et les termes d'entreprise associés.

Lorsque vous cliquez sur une colonne ou une règle, la vue détaillée correspondante s'ouvre.

Le tableau suivant décrit les volets de la vue détaillée :

Volets	Description
Général	Affiche le nombre de lignes contenant des valeurs Null, distinctes et non distinctes dans différentes couleurs. Vous pouvez afficher les valeurs en pourcentage. Vous pouvez afficher un graphique sparkline pour voir l'augmentation et la diminution des valeurs générales dans des exécutions de profil consécutives. Un graphique sparkline affiche la variation du nombre de valeurs Null, distinctes ou non distinctes au cours des cinq dernières exécutions consécutives de profil dans un graphique en courbes. Vous pouvez afficher le nombre de valeurs et le pourcentage de valeurs lorsque vous placez le pointeur sur le graphique sparkline pour chaque exécution de profil. Vous pouvez ajouter des balises et des commentaires à la colonne.
Modèles	Affiche les modèles des valeurs de colonne. La fréquence d'affichage des modèles dans une colonne est représentée sous forme de graphique à barres horizontales et en pourcentage. Vous pouvez explorer un modèle, ajouter un modèle à une table de référence ou créer un domaine de données avec le modèle sélectionné.
Type de données	Affiche les types de données inférés pour la colonne. La fréquence des types de données dans une colonne s'affiche sous forme de graphique à barres horizontales et en pourcentage. Vous pouvez explorer un type de données, approuver, rejeter ou réinitialiser le type de données inféré sélectionné. L'option Afficher les éléments rejetés affiche les types de données inférés rejetés.
Domaine de données	Affiche les domaines de données inférés pour la colonne. Vous pouvez explorer un domaine de données pour les lignes conformes, les lignes non conformes ou les lignes comprenant des valeurs Null. Vous pouvez approuver, rejeter ou réinitialiser la valeur du domaine de données. L'option Afficher les éléments rejetés affiche les domaines de données rejetés. Vous pouvez vérifier la valeur du domaine de données.
Terme d'entreprise	Affiche les termes d'entreprise attribués à la colonne. Vous pouvez attribuer un terme d'entreprise à une colonne ou annuler son attribution.
Valeurs	Affiche une représentation graphique de toutes les valeurs de la colonne, ainsi que la fréquence, la longueur et le pourcentage. Vous pouvez explorer chaque valeur. Vous pouvez ajouter la valeur à une table de référence et créer une règle de fréquence de valeur et un domaine de données.
Aperçu des données	Affiche les données d'exploration pour le modèle, le type de données, le domaine de données ou la valeur sélectionnée.

Statistiques

Vous pouvez afficher des statistiques telles que des valeurs, des modèles, des types de données, des domaines de données et des valeurs hors norme pour les colonnes et les règles d'un profil.

Vous pouvez afficher les statistiques de profil dans la vue de résumé et les statistiques de colonne dans la vue de résumé et la vue détaillée. Vous pouvez afficher les statistiques de la dernière exécution de profil, l'exécution de profil historique et l'exécution de profil consolidée. Vous pouvez comparer les résultats de deux exécutions de profil et afficher les statistiques du profil et des colonnes dans la vue de résumé et la vue détaillée.

Aperçu des données

Le volet Aperçu des données vous permet d'afficher les données d'exploration pour le modèle sélectionné, le type de données, le domaine de données ou la valeur.

Vous pouvez afficher le volet Aperçu des données dans la vue détaillée. Lorsque vous cliquez sur une colonne dans la vue, la vue détaillée s'affiche et le volet Aperçu des données est réduit par défaut. Pour afficher les données de colonne, vous pouvez cliquer sur **Actions > Afficher l'aperçu**.

Le tableau suivant décrit les options du menu **Actions** dans le volet Aperçu des données :

Option	Description
Ajouter au filtre	Créez un filtre d'exploration pour filtrer les données d'exploration afin de pouvoir analyser les irrégularités présentes dans les données dans les sous-ensembles de résultats du profil.
Enregistrer le filtre	Enregistre le filtre d'exploration.
Afficher l'aperçu	Affiche les lignes source.
Exporter les données	Exporte les résultats d'exploration vers un fichier CSV ou Microsoft Excel.

Types de données

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

Vous pouvez afficher les types de données dans la vue de résumé et la vue détaillée. La vue de résumé vous permet d'afficher le type de données documenté et les types de données inférés. Le filtre **Type de données en conflit** affiche les colonnes dans lesquelles le type de données documenté et le type de données inféré sont en conflit. La vue détaillée vous permet d'afficher les types de données inférés pour la colonne. La fréquence des types de données dans une colonne s'affiche sous forme de graphique à barres horizontales et en pourcentage. Vous pouvez explorer, approuver, rejeter ou réinitialiser le type de données inféré sélectionné. L'option Afficher les éléments rejetés affiche les types de données inférés rejetés.

Le tableau suivant décrit les propriétés des types de données :

Propriété	Description
Type de données	Indique la liste des types de données documentés et inférés de la colonne dans le profil.
Fréquence	Indique le nombre d'occurrences d'un type de données dans une colonne.
Pourcentage	Indique le pourcentage d'occurrence d'un type de données dans une colonne.

Propriété	Description
Explorer	Explore des lignes source spécifiques en fonction d'un type de données de colonne. Remarque: L'exploration n'est pas disponible si plusieurs types de données inférés sont sélectionnés.
Statut	Indique le statut du type de données. Ce statut peut être Inféré, Approuvé ou Rejeté. Inféré Indique le type de données de la colonne que l'outil Analyst tool a inféré. Approuvé Indique un type de données approuvé pour la colonne. Lorsque vous approuvez un type de données, vous le validez dans le référentiel modèle. Rejeté Indique un type de données rejeté pour la colonne.

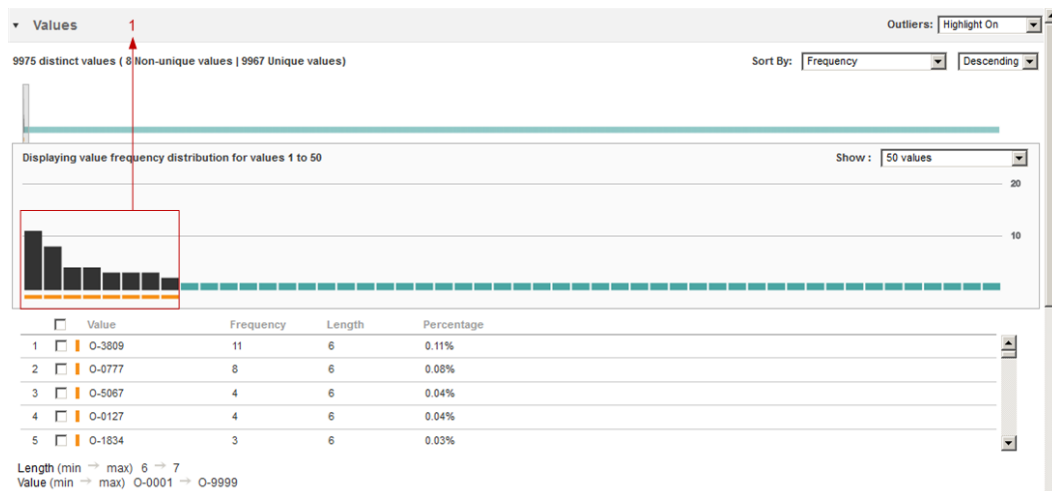
Valeurs hors norme

Une valeur hors norme est un modèle, une valeur ou une fréquence dans une colonne des résultats de profil qui ne se situe pas dans la plage de valeurs attendue.

Le plug-in de profilage du service d'intégration de données exécute un algorithme identifiant les valeurs qui ne se situent pas dans la plage de valeurs majoritaires de la colonne. Les modèles, valeurs ou fréquences qui ne se situent pas dans la plage de valeurs majoritaires attendue dans la colonne sont considérés comme des valeurs hors norme.

Par défaut, l' découverte de Analyst tool ne détermine pas les valeurs hors norme dans les résultats de profil. Vous pouvez exécuter une valeur hors norme pour afficher ses résultats dans la vue de résumé. Le filtre Valeur hors norme du modèle indique les valeurs hors norme en fonction des modèles de la colonne. Le filtre Valeur hors norme de la fréquence de la valeur affiche les valeurs hors norme en fonction des valeurs ou des fréquences de la colonne. La détection des valeurs hors norme s'effectue en arrière-plan, ce qui vous permet d'effectuer d'autres actions en parallèle dans la vue de résumé.

La vue détaillée vous permet d'afficher les valeurs hors norme dans le volet Valeurs lorsque vous sélectionnez l'option **Mise en surbrillance activée** dans la liste déroulante. La valeur hors norme s'affiche sous la forme d'une barre verticale avec un soulignement orange. Pour afficher uniquement la valeur hors norme, sélectionnez l'option **Filtre** de la liste déroulante.



1. Valeurs hors norme. Une valeur hors norme s'affiche sous la forme d'une barre verticale avec un soulignement orange.

Exécution d'une valeur hors norme

Exécutez une valeur hors norme pour identifier les modèles, les valeurs ou les fréquences d'une colonne qui ne se situent pas dans une plage de valeurs attendue.

1. Dans la vue de résumé, cliquez sur **Actions > Détecter une valeur hors norme**.
La valeur hors norme du modèle et la valeur hors norme de la fréquence de valeur du volet **Filtrer par** passe de S.O. au nombre de valeurs hors norme détectées.
2. Dans le volet **Filtrer par**, cliquez sur **Valeur hors norme du modèle**.
Les colonnes contenant des valeurs hors norme de modèle s'affichent dans la vue de résumé.
3. Dans le volet **Filtrer par**, cliquez sur **Valeur hors norme de la fréquence de la valeur**.
Les colonnes contenant des valeurs hors norme de valeur ou de fréquence s'affichent dans la vue de résumé.
4. Dans la vue détaillée, sélectionnez **Mise en surbrillance activée** dans la liste déroulante des valeurs hors norme.
Dans le volet Valeurs, les valeurs hors norme s'affichent sous la forme de barres verticales avec un soulignement orange.
5. Cliquez sur **Filtre** dans la liste déroulante des Valeurs hors norme pour afficher uniquement les valeurs hors norme.

Modèles

Vous pouvez afficher les modèles des valeurs de colonne et la fréquence à laquelle ils s'affichent dans la vue de résumé et la vue détaillée.

La vue de résumé vous permet d'afficher les différents modèles de la colonne sous forme de graphiques à barres horizontales. Vous pouvez survoler le graphique à barres avec votre souris pour afficher les caractères de modèle et le pourcentage de modèles similaires dans une colonne. La vue détaillée vous permet d'afficher sous forme de graphique à barres horizontales et en pourcentage la fréquence à laquelle les modèles s'affichent dans une colonne. Vous pouvez explorer les données, ajouter le modèle à une table de référence ou créer un domaine de données avec le modèle sélectionné.

L'entrepôt de profilage stocke un maximum de 16 000 valeurs de fréquence supérieures uniques, y compris les valeurs NULL des résultats de profil par défaut. Si les résultats de profil comprennent au moins une valeur NULL, l' découverte de Analyst tool peut afficher les valeurs NULL en tant que formes.

Remarque: L' découverte de Analyst tool ne peut pas dériver la forme d'une colonne numérique dont la précision est supérieure à 38. L' découverte de Analyst tool 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	Affiche le modèle de la colonne du profil.
Fréquence	Indique le nombre d'occurrences d'un modèle dans une colonne.
Pourcentage	Indique le pourcentage d'occurrence d'un modèle dans une 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' découverte de 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' découverte de 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.

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

Valeurs

Vous pouvez afficher les valeurs des colonnes et la fréquence à laquelle elles s'affichent dans les colonnes.

Affichez les valeurs minimale et maximale d'une colonne dans la vue de résumé. La vue détaillée vous permet d'afficher les propriétés des valeurs d'une colonne.

Valeurs de la vue de résumé

Vous pouvez afficher les valeurs minimales et maximales de toutes les colonnes et règles de la dernière exécution de profil, d'une exécution de profil historique et d'une exécution de profil consolidée dans la vue de résumé.

Exemple

La base de données d'un magasin de détail comporte une colonne nommée ID d'employé dans la table Employé. Cette colonne contient des ID employé compris entre 100 et 250 et des noms, tels que Bob et Robert. Lorsque vous exécutez un profil de colonne sur la table Employé, la colonne Valeur de l'ID d'employé dans la vue de résumé affiche 100 --> Robert.

Valeurs de la vue détaillée

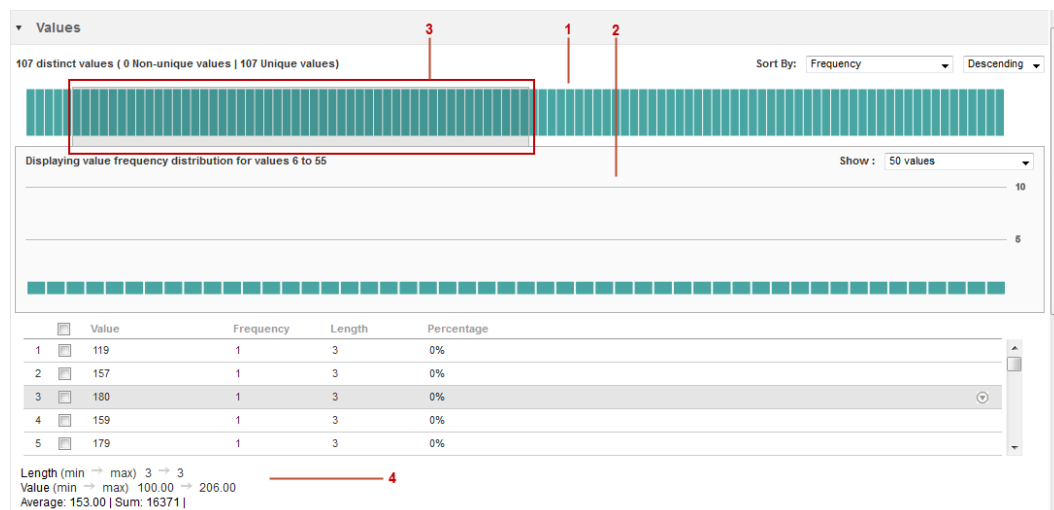
Les valeurs de colonne dans la vue détaillée comprennent les valeurs d'une colonne et la fréquence à laquelle elles s'affichent dans la colonne.

Le volet **Valeurs** affiche une représentation graphique des valeurs de colonne. Vous pouvez afficher la fréquence, la longueur et le pourcentage de chaque valeur. Vous pouvez trier les valeurs en fonction de leur valeur ou de leur fréquence. Vous pouvez explorer les données, ajouter des valeurs à une table de référence, créer une règle de fréquence de valeur ou créer un domaine de données. Vous pouvez afficher les valeurs Null sous la forme d'une barre verticale rouge, la fréquence des valeurs sous la forme d'une barre verticale noire et les valeurs hors norme sous la forme de barres verticales orange. Vous pouvez mettre en surbrillance les valeurs hors norme, désactiver leur affichage ou filtrer les résultats pour afficher uniquement les valeurs hors norme dans la colonne.

Le volet Valeurs contient les sections de présentation graphique et des valeurs.

La présentation graphique est divisée en deux panneaux.

L'image suivante présente le volet Valeurs de la vue détaillée :



1. Panneau supérieur. Vous pouvez afficher les valeurs dans un graphique à barres verticales, trier les valeurs par fréquence et valeur et trier les valeurs par ordre croissant ou décroissant. Vous pouvez afficher les valeurs hors norme sous la forme de barres verticales orange.
2. Panneau inférieur. Vous pouvez afficher les valeurs dans la réglette du panneau inférieur où chaque valeur est représentée par une barre verticale. Vous pouvez explorer la valeur, l'ajoutez à une table de référence, créez une règle de fréquence de valeur et créer un domaine de données sur la valeur. Vous pouvez afficher plus 50, 75 ou 100 valeurs simultanément.
3. Réglette. Vous pouvez faire glisser la réglette sur les valeurs du panneau supérieur. Le panneau inférieur affiche les valeurs dans la réglette.
4. Propriétés de la valeur. La section Propriétés de la valeur affiche la valeur et ses propriétés.

Le tableau suivant décrit les panneaux de la présentation graphique :

Panneau	Description
Panneau supérieur	Affiche toutes les valeurs sous forme de graphique à barres verticales. Vous pouvez afficher un maximum de 16 000 valeurs dans le panneau supérieur. Vous pouvez utiliser le curseur pour afficher un lot de valeurs.
Panneau inférieur	Affiche les valeurs du lot sélectionné dans le panneau supérieur. Par défaut, l'outil Analyst tool affiche 50 valeurs. Vous pouvez choisir d'afficher 75 ou 100 valeurs.

Le tableau suivant décrit les propriétés des valeurs de colonne dans la section des valeurs :

Propriété	Description
Valeur	Affiche la liste de valeurs du lot sélectionné dans le panneau supérieur. Remarque: L'outil Analyst tool exclut les types de données CLOB, BLOB, Brut et Binaire des valeurs de colonne.
Fréquence	Indique le nombre d'occurrences d'une valeur dans la colonne.
Longueur	Affiche la longueur de la valeur de colonne.
Pourcentage	Affiche le pourcentage d'occurrence d'une valeur dans la colonne.

Le tableau suivant décrit les statistiques de la colonne sélectionnée :

Statistiques	Description
Longueur (min. - max.)	Affiche la longueur de la valeur la plus courte et de la valeur la plus longue de la colonne.
Valeur (min. - max.)	Affiche les valeurs minimale et maximale de la colonne.
Moyenne	Affiche la moyenne des valeurs de la colonne.
Somme	Affiche la somme de toutes les valeurs de la colonne.

Valeurs dans la vue détaillée de la comparaison de résultats de profil

Le volet Valeurs de la vue détaillée pour la comparaison de résultats de profil affiche les propriétés des valeurs, telles que le nombre de valeurs distinctes, les valeurs minimale et maximale, la longueur maximale et minimale, la moyenne, l'écart-type et la somme des valeurs.

La vue détaillée d'une colonne pour la comparaison de résultats de profil affiche les propriétés des valeurs, la valeur et la fréquence de la valeur sous forme de graphique à barres horizontales.

Le tableau suivant décrit les propriétés des valeurs de colonne dans la vue détaillée lors de la comparaison des résultats de deux exécutions de profil :

Propriété	Description
Nombre de valeurs distinctes	Affiche le nombre de valeurs distinctes de la colonne.
Valeur min.	Affiche la valeur minimale de la colonne.
Valeur max.	Affiche la valeur maximale de la colonne.
Longueur (min. - max.)	Affiche la longueur de la valeur la plus courte et de la valeur la plus longue de la colonne.
Moyenne	Affiche la moyenne des valeurs de la colonne.
Écart-type	Affiche l'écart-type ou la variabilité entre toutes les valeurs de la colonne.
Somme	Affiche la somme de toutes les valeurs de la colonne.

Types d'exécution de profil

Vous pouvez afficher les résultats de la dernière exécution de profil, de l'exécution de profil historique et de l'exécution de profil consolidée. Vous pouvez afficher les résultats de l'exécution de profil dans la vue de résumé.

Dernière exécution de profil

Affichez les résultats de la dernière exécution de profil dans la vue de résumé.

Vous pouvez afficher les résultats de profil de la dernière exécution de profil dans la vue de résumé lorsque vous :

- Créez, enregistrez et exécutez un profil.
- Ouvrez un profil que vous avez exécuté préalablement à partir de l'espace de travail **Bibliothèque**.
- Cliquez sur le lien **Retour à la dernière exécution de profil** dans la vue de résumé ou la vue détaillée pour l'exécution de profil consolidée.
- Cliquez sur le lien **Retour à la dernière exécution de profil** dans la vue de résumé ou la vue détaillée pour une exécution de profil historique.
- Sélectionnez la dernière exécution de profil dans la boîte de dialogue **Sélectionnez l'exécution du profil** et cliquez sur **OK**.

Exécution de profil historique

Affichez les résultats d'une exécution de profil précédente dans la vue de résumé.

L'entrepôt de profilage enregistre les résultats de toutes les exécutions d'un profil. Pour afficher les résultats d'une version précédente de l'exécution du profil, sélectionnez l'exécution dans la boîte de dialogue **Sélectionnez l'exécution du profil**.

Exécution de profil consolidée

Affichez les derniers résultats de profil de chaque colonne du profil dans la vue de résumé.

Dans l'exécution de profil consolidée, vous pouvez afficher les derniers résultats de chaque colonne du profil. Lorsque vous choisissez Exécution de profil consolidée dans la boîte de dialogue **Sélectionnez l'exécution du profil**, l'entrepôt de profilage récupère les derniers résultats de colonne de toutes les exécutions du profil. Vous pouvez afficher les résultats dans la vue de résumé. L'en-tête de la vue de résumé affiche les exécutions de profil incrémentielles.

Exemple

En tant qu'analyste de données, vous pouvez afficher les derniers résultats de chaque colonne d'un profil. Par exemple, vous pouvez choisir les colonnes 1, 2 et 3 pour l'exécution de profil A et les colonnes 3, 4 et 5 pour l'exécution de profil B. Pour afficher les derniers résultats de toutes les colonnes, vous pouvez choisir l'option Exécution de profil consolidée dans la boîte de dialogue Sélectionnez l'exécution du profil. La vue de résumé affiche les résultats des colonnes 1 et 2 pour l'exécution A et des colonnes 3, 4 et 5 pour l'exécution B.

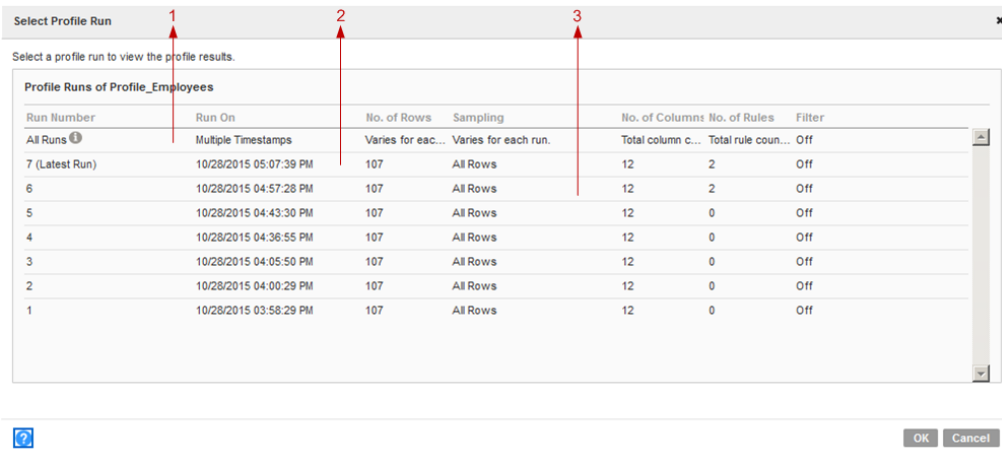
Sélection d'une exécution de profil

Vous pouvez sélectionner une exécution de profil historique, la dernière exécution de profil ou l'exécution de profil consolidée pour afficher les résultats de profil. Vous pouvez afficher les résultats de profil dans la vue de résumé et les résultats de colonne dans la vue détaillée.

1. Dans l'espace de travail **Bibliothèque**, sélectionnez le projet ou le dossier contenant le profil ou sélectionnez le profil dans le volet **Ressources**.
2. Cliquez sur **Actions > Ouvrir** pour ouvrir le profil.
La vue de résumé s'ouvre dans l'espace de travail **Découverte**.
3. Dans la vue de résumé, cliquez sur **Actions > Choisir l'exécution de profil**.

La boîte de dialogue **Sélectionnez l'exécution du profil** s'affiche.

L'image suivante présente la boîte de dialogue **Sélectionnez l'exécution du profil**.



1. Exécution de profil consolidée. Lorsque vous choisissez cette exécution de profil, vous pouvez afficher les derniers résultats de profil pour chaque colonne dans la vue de résumé.
2. Dernière exécution de profil. Lorsque vous choisissez cette exécution de profil, vous pouvez afficher les derniers résultats de profil dans la vue de résumé.

3. Exécution de profil historique. Lorsque vous choisissez cette exécution de profil, vous pouvez afficher les résultats de profil historique d'une exécution de profil précédente dans la vue de résumé.
4. Dans la boîte de dialogue **Sélectionnez l'exécution du profil**, sélectionnez l'une des exécutions de profil pour afficher ses résultats :
 - Pour afficher les résultats de profil de la dernière exécution de profil, sélectionnez la dernière exécution de profil et cliquez sur **OK**.
 - Pour afficher les résultats de profil d'une exécution de profil historique, sélectionnez une exécution de profil autre que la dernière et cliquez sur **OK**.
 - Pour afficher les résultats de profil d'une exécution de profil consolidée, sélectionnez **Toutes les exécutions** puis cliquez sur **OK**. Les derniers résultats de profil de chaque colonne s'affichent dans la vue de résumé.

L'outil Analyst tool lance une exécution de profil et affiche les résultats dans la vue de résumé.
5. Cliquez sur une colonne de la vue de résumé pour afficher les résultats de colonne.

La vue détaillée s'affiche.

Présentation des résultats de la comparaison de plusieurs profils

Vous pouvez comparer les résultats de deux exécutions de profil, puis afficher les résultats de la comparaison dans la vue de résumé et les résultats de colonne dans la vue détaillée.

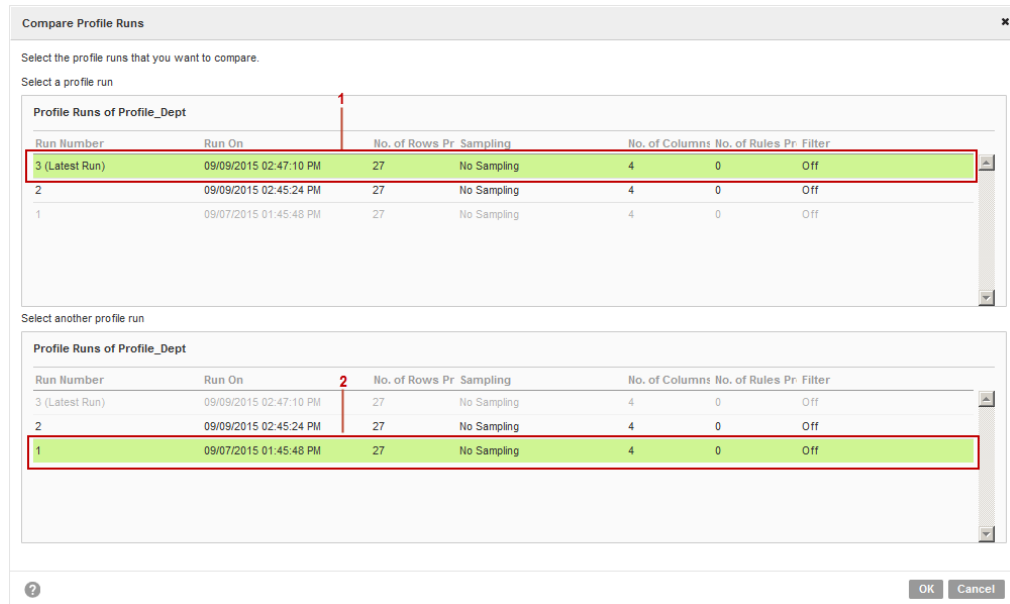
La vue de résumé vous permet d'afficher les résultats de la comparaison pour toutes les colonnes des deux exécutions de profil.

Résultats de la comparaison de plusieurs profils

Lors de la comparaison de deux exécutions de profil, vous pouvez afficher la comparaison dans la vue de résumé.

1. Dans la vue de résumé, cliquez sur **Actions > Comparer les exécutions de profil**.

L'image suivante présente la boîte de dialogue **Comparer les exécutions de profil**.



- Exécution A. Choisissez une exécution de profil comme Exécution A.
- Exécution B. Choisissez une exécution de profil comme Exécution B.

La boîte de dialogue **Comparer les exécutions de profil** s'affiche.

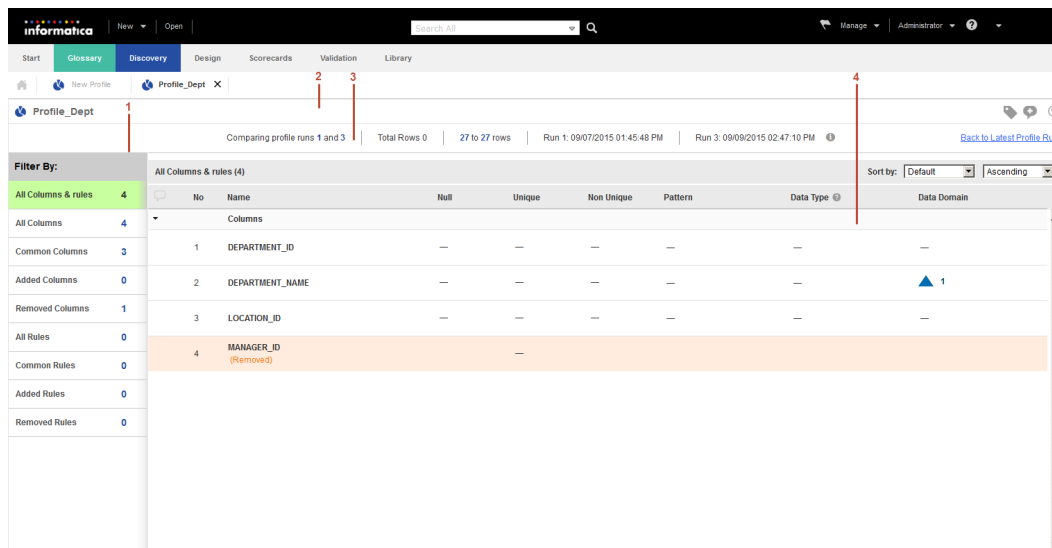
- Sélectionnez un premier profil dans le volet de l'**exécution A**, puis un deuxième dans le volet de l'**exécution B**.
- Cliquez sur **OK**.

La vue de résumé affiche une vue consolidée des résultats de profil.

Vue de résumé des résultats de la comparaison de profils

Lorsque vous comparez deux exécutions de profil, vous pouvez en afficher les résultats au format de grille dans la vue de résumé. Vous pouvez utiliser les filtres par défaut de la vue de résumé pour afficher des statistiques spécifiques.

L'image suivante présente les résultats de comparaison de profils de deux exécutions de profil dans la vue de résumé :



1. Filtre par défaut. Les filtres par défaut de la vue de résumé vous permettent de gérer l'affichage des résultats de comparaison de profil.
2. En-tête du profil. Vous pouvez afficher le nom du profil dans l'en-tête.
3. En-tête de la vue de résumé. Vous pouvez afficher des informations spécifiques au profil dans l'en-tête de la vue de résumé. Vous pouvez afficher les exécutions de profil comparées, l'augmentation ou la diminution de lignes entre les exécutions de profil, le nombre de lignes dans le profil et l'heure et la date des exécutions du profil.
4. Vue de résumé. Vous pouvez afficher la comparaison entre les colonnes dans les deux exécutions de profil.

Propriétés de la vue de résumé de la comparaison de résultats de profil

Les propriétés de la vue de résumé des résultats de la comparaison de profils comprennent le nombre et le pourcentage de valeurs distinctes, non distinctes et Null, les modèles, les types de données et domaines de données inférés ainsi que les termes d'entreprise associés. La vue de résumé comprend une représentation visuelle des propriétés. Vous pouvez cliquer sur chaque propriété du résumé pour trier les valeurs correspondantes.

Dans la vue de résumé, le service d'intégration de données attribue un numéro dans l'ordre croissant à l'ensemble des colonnes et des règles.

Remarque: Une flèche vers le haut accompagnée d'un nombre indique une augmentation des valeurs d'une propriété d'une exécution de profil à une autre. Une flèche vers le bas accompagnée d'un nombre indique une diminution des valeurs d'une propriété d'une exécution de profil à une autre.

Le tableau suivant décrit les propriétés de résumé des résultats de la comparaison de profils :

Propriété	Description
Non	Affiche le numéro de la colonne ou de la règle.
Nom	Affiche le nom de la colonne ou de la règle dans le profil.
Null	Indique une augmentation ou une diminution des valeurs Null.

Propriété	Description
Distinct	Indique une augmentation ou une diminution des valeurs distinctes.
Non distinct	Indique une augmentation ou une diminution des valeurs non distinctes.
Modèle	Affiche les variations des modèles entre les exécutions de profil.
Type de données	Affiche les variations entre les types de données inférés pour la colonne ou la règle dans les deux exécutions de profil.
Domaine de données	Affiche les variations entre les domaines de données inférés associés à la colonne ou à la règle dans les deux exécutions de profil.

Filtres par défaut pour la comparaison de résultats de profil dans la vue de résumé

Vous pouvez afficher les résultats de profil selon les filtres par défaut dans la vue de résumé.

Dans la vue de résumé, vous pouvez afficher les colonnes source et les colonnes virtuelles. La sortie pour une règle apparaît dans une colonne virtuelle dans la vue de résumé. Lorsque vous modifiez le port de sortie d'une règle et que vous comparez l'exécution du profil avec l'exécution historique, la colonne de sortie des règles historiques s'affiche dans le filtre **Règles supprimées** et la colonne de sortie des nouvelles règles s'affiche dans le filtre **Règles ajoutées**. Si vous changez la logique de règle d'une seule règle de sortie, ou si vous modifiez les entrées d'une sortie de plusieurs règles dans une exécution de profil et que vous la comparez avec une exécution historique, la sortie des filtres **Règles ajoutées** et **Règles supprimées** ne change pas. La sortie des filtres ne change pas, car les filtres ne tiennent compte que des modifications de dénomination apportées aux colonnes comme entrées valides du filtre.

Vous pouvez utiliser les options de filtre par défaut suivantes pour afficher les résultats de profil qui remplissent des conditions spécifiques :

Option de filtre par défaut	Description
Toutes les colonnes et règles	Affiche les résultats de profil des colonnes source, virtuelles et de règle. Vous pouvez développer et réduire les colonnes source et de règle pour afficher les résultats.
Toutes les colonnes	Affiche les résultats de profil des colonnes source et virtuelles.
Colonnes communes	Affiche les colonnes disponibles dans les résultats des deux exécutions de profil.
Colonnes ajoutées	Affiche les colonnes disponibles dans la dernière exécution de profil. Par exemple, lorsque vous comparez les 5e et 3e exécutions, l'option Colonnes ajoutées affiche les colonnes qui sont disponibles dans la 5e exécution, mais pas dans la 3e.
Colonnes retirées	Affiche les colonnes disponibles dans l'exécution de profil historique. Par exemple, lorsque vous comparez les 5e et 3e exécutions, l'option Colonnes retirées affiche les colonnes qui sont disponibles dans la 3e exécution, mais pas dans la 5e.
Toutes les règles	Affiche les résultats de profil pour toutes les colonnes de règle.

Option de filtre par défaut	Description
Règles ajoutées	Affiche les règles disponibles dans la dernière exécution de profil. Par exemple, lorsque vous comparez les 5e et 3e exécutions, l'option Règles ajoutées affiche les règles qui sont disponibles dans la 5e exécution, mais pas dans la 3e.
Règles supprimées	Affiche les règles disponibles dans l'historique d'exécution du profil. Par exemple, lorsque vous comparez les 5e et 3e exécutions, l'option Règles supprimées affiche les règles qui sont disponibles dans la 3e exécution, mais pas dans la 5e.

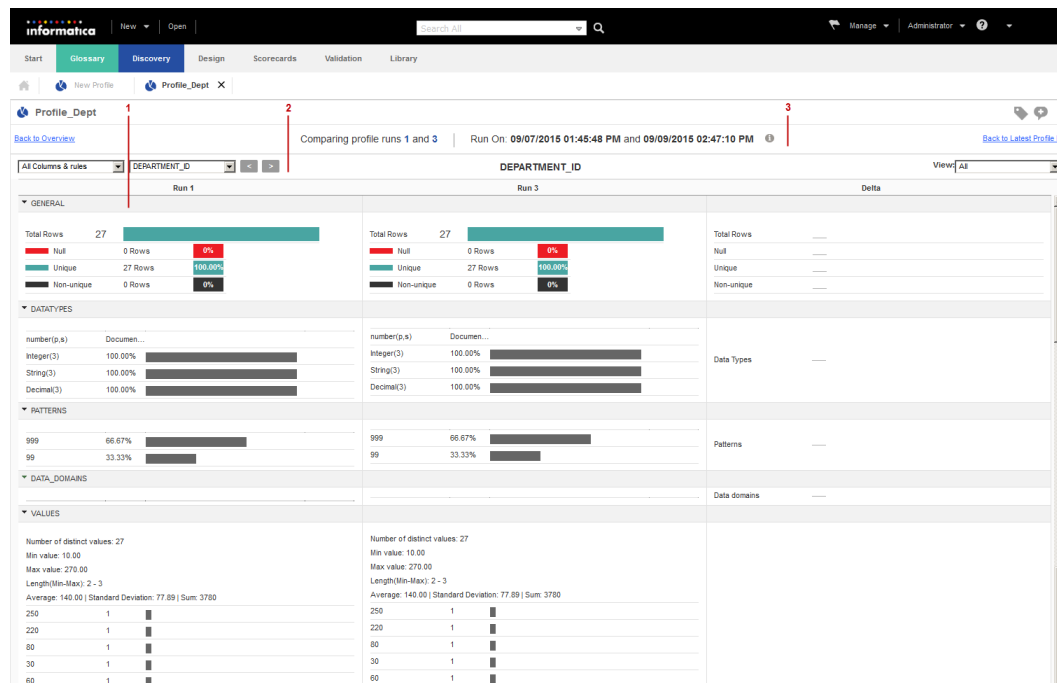
Par défaut, la vue de résumé affiche les résultats de profil pour toutes les colonnes source et virtuelles.

Vue détaillée des résultats de la comparaison de profils

Les résultats de colonne s'affichent sous la forme d'une grille dans la vue détaillée. Les détails de colonne comprennent des informations générales (valeurs distinctes, non distinctes et Null), les modèles, les types de données, les domaines de données, les termes d'entreprise, les valeurs et l'aperçu des données.

La vue détaillée d'une colonne s'affiche lorsque vous cliquez sur le nom de la colonne. Vous pouvez afficher les résultats de colonne des exécutions A et B dans des colonnes distinctes. La comparaison des données est disponible dans la colonne delta.

L'image suivante présente les résultats de la comparaison de profils pour une colonne dans la vue détaillée :



1. Volets. Vous pouvez afficher les résultats et les statistiques de profil de la colonne dans les deux exécutions de profil et afficher l'information delta de la colonne dans des volets dans les deux exécutions de profil.
2. En-tête du profil. Vous pouvez afficher les résultats de la colonne en sélectionnant la colonne dans la liste déroulante ou en utilisant les boutons de navigation. Vous pouvez afficher le nom de la colonne et des résultats spécifiques en utilisant les options de la liste déroulante de la vue.

3. En-tête de la vue de résumé. Vous pouvez afficher des informations spécifiques au profil dans l'en-tête de la vue de résumé. Vous pouvez afficher les exécutions de profil qui sont comparées ainsi que les heures et les dates d'exécution.

Volets de la vue détaillée de la comparaison de résultats de profil

La vue détaillée affiche en détail les résultats de profil et les résultats de comparaison d'une colonne dans les deux exécutions de profil.

La vue détaillée affiche les résultats de colonne pour les exécutions A et B. La comparaison des données est disponible dans la colonne delta. Pour afficher d'autres résultats de colonne, vous pouvez sélectionner un filtre ou une colonne dans la liste déroulante correspondante.

Développement de profil de colonne

Utilisez les options d'exploration d'un profil de colonne pour explorer des lignes spécifiques dans la source de données en fonction d'une valeur de colonne. Vous pouvez choisir de lire les données actuelles d'une source de données pour les explorer ou de lire les données de profil intermédiaires de l'entrepôt de profils. Lors de l'exploration d'une ligne spécifique des données de profil intermédiaires, l'outil Analyst tool crée un filtre d'exploration pour la valeur de colonne correspondante. Une fois l'exploration terminée, vous pouvez modifier, rappeler, réinitialiser et enregistrer le filtre d'exploration.

Vous pouvez sélectionner pour l'exploration des colonnes que vous n'avez pas sélectionnées pour le profilage. Vous pouvez choisir de lire les données actuelles d'une source de données pour les explorer ou de lire les données de profil intermédiaires de l'entrepôt de profilage. Après avoir exploré une valeur de colonne, vous pouvez exporter les données d'exploration des valeurs ou des modèles sélectionnés dans un fichier CSV à l'emplacement de votre choix. Bien qu'Informatica Analyst affiche uniquement les 200 premières valeurs des données d'exploration, toutes les valeurs sont exportées dans le fichier CSV.

Développement des données de ligne

Après avoir exécuté un profil, vous pouvez explorer des lignes spécifiques correspondant à la valeur, au type de données ou au modèle de la colonne.

1. Exécutez un profil.
Les résultats de profil s'affichent dans la vue de résumé.
2. Cliquez sur un nom de colonne dans la vue de résumé.
Les résultats de colonne s'affichent dans la vue détaillée.
3. Dans la vue détaillée, cliquez avec le bouton droit de la souris sur une valeur du volet **Valeurs**, puis sélectionnez **Exploration**.
Le volet **Aperçu des données** affiche les données de l'exploration.

Application de filtres aux données de développement

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

1. Sélectionnez une valeur de colonne sur l'onglet **Valeurs**.

2. Cliquez avec le bouton droit de la souris et sélectionnez **Exploration**.
Les résultats de l'exploration s'affichent dans le volet **Aperçu des données**.
3. Pour ajouter une condition de filtre, cliquez avec le bouton droit de la souris sur une valeur de colonne du volet **Aperçu des données** et sélectionnez **Ajouter au filtre**.
La boîte de dialogue **Filtre de développement** s'affiche avec la condition de filtre.
4. Ajoutez les conditions de filtre requises et cliquez sur **OK**.
Vous ne pouvez pas appliquer de filtre de développement à des types de données inférés.
5. Pour enregistrer le filtre, cliquez sur **Actions > Enregistrer le filtre**.
6. Pour effacer les filtres d'exploration, cliquez sur **Actions > Actualiser**.
7. Pour exporter des données de développement vers une feuille de calcul Microsoft Excel, cliquez sur **Actions > Exporter les données**.

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 organisez des métadonnées dans l'outil Analyst tool, vous pouvez approuver, rejeter et réinitialiser les types de données ou les domaines de données inférés dans les résultats de profil.

Vous pouvez approuver un type de données et un domaine de données pour une colonne. Vous pouvez masquer les types de données ou les domaines de données rejetés pour une colonne. Après avoir approuvé ou rejeté un type de données ou un domaine de données inféré, vous pouvez réinitialiser le type de données ou le domaine de données pour restaurer le statut inféré.

Approbation de types de données et de domaines de données

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

1. Exécutez un profil.
Les résultats de profil s'affichent dans la vue de résumé.
2. Cliquez sur un nom de colonne dans la vue de résumé.
Les résultats de colonne s'affichent dans la vue détaillée.
3. Dans la vue détaillée, sélectionnez un type de données dans le volet **Type de données** ou un domaine de données dans le volet **Domaine de données**.
4. Cliquez sur **Actions > Approuver**.
5. Pour restaurer le statut inféré du type ou du domaine de données, sélectionnez l'élément souhaité et cliquez sur **Actions > Réinitialiser**.

Rejet de types de données et de domaines de données

La vue détaillée vous permet de rejeter un type de données ou un domaine de données. Vous pouvez afficher ou masquer les types de données et les domaines de données rejetés.

1. Exécutez un profil.
Les résultats de profil s'affichent dans la vue de résumé.
2. Cliquez sur un nom de colonne dans la vue de résumé.
Les résultats de colonne s'affichent dans la vue détaillée.
3. Dans la vue détaillée, sélectionnez un type de données dans le volet **Type de données** ou un domaine de données dans le volet **Domaine de données**.
4. Cliquez sur **Actions > Rejeter**.
L'outil Analyst tool supprime le type rejeté de la liste des types de données.
5. Pour afficher les domaines de données rejetés, cliquez sur **Actions > Afficher 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 d'exploration 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 du service d'intégration de données 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 d'exploration pour afficher les données dans 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 tool enregistre le fichier au format « .xlsx ».

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

Onglet	Description
Profil de colonne	Résumé des informations exportées à partir de la vue Résumé après l'exécution du profil. Ces informations peuvent inclure les noms des colonnes, les noms des règles, le nombre de valeurs distinctes, 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 fréquence à laquelle elles s'affichent pour chaque colonne.
Modèles	Modèles de valeur des colonnes et des règles sur lesquelles vous avez exécuté le profil et fréquence à laquelle ils s'affichent.
Types de données	Tous les types de données de la colonne, fréquence de chaque type de données, valeur en pourcentage et statut du type de données (par exemple Inféré, Approuvé ou Rejeté).
Statistiques	Statistiques concernant chaque colonne et chaque règle, par exemple, moyenne, longueur, valeurs supérieures, valeurs inférieures et écart-type.
Propriétés	Informations de la vue Propriétés, notamment le nom et le type du profil, la stratégie d'échantillonnage et le nombre de lignes.

Exportation de 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 l'espace de travail **Bibliothèque**, sélectionnez le projet ou le dossier contenant le profil.
2. Cliquez sur le profil pour l'ouvrir.
Les résultats de profil s'affichent dans la vue de résumé.
3. Dans la vue de résumé, cliquez sur **Actions > Exporter les données**.
La boîte de dialogue **Exporter les données vers un fichier** s'ouvre.
4. Dans la boîte de dialogue **Exporter les données vers un fichier**, entrez un nom de fichier. Vous pouvez éventuellement utiliser le nom de fichier par défaut.
5. Sélectionnez **Tout (Résumé, Valeurs, Modèles, Statistiques, Propriétés)** ou **Résultats de la découverte de domaines de données**, et sélectionnez une **Page de code**. Cliquez sur **OK**.
Les données sont exportées vers une feuille de calcul Microsoft Excel.
6. Cliquez sur une colonne dans la vue de résumé.
Les résultats de colonne s'affichent dans la vue détaillée.
7. Dans la vue détaillée, cliquez sur **Actions > Exporter les données**.
La boîte de dialogue **Exporter les données vers un fichier** s'ouvre.
8. Dans la boîte de dialogue **Exporter les données vers un fichier**, entrez un nom de fichier. Vous pouvez éventuellement utiliser le nom de fichier par défaut.

9. Sélectionnez l'une des options suivantes :
 - Tout (résumé, valeurs, formes, statistiques, propriétés)
 - Fréquences de valeur pour la colonne sélectionnée.
 - Fréquences de forme pour la colonne sélectionnée.
 - Types de données de la colonne sélectionnée.
 - Données d'exploration des valeurs sélectionnées.
 - Données d'exploration des modèles sélectionnés.
 - Données d'exploration des types de données sélectionnés.
10. Entrez un format de fichier. Le format est **Excel** pour l'option **Tout** et **CSV** pour les autres options. Vous pouvez choisir d'exporter le nom de champ comme première ligne du fichier.
11. Sélectionnez la page de codes du fichier.
12. Cliquez sur **OK**.

Les données sont exportées vers le fichier.

CHAPITRE 10

Termes d'entreprise, commentaires et balises dans Informatica Analyst

Ce chapitre comprend les rubriques suivantes :

- [Présentation des termes d'entreprise, des commentaires et des balises dans Informatica Analyst, 79](#)
- [Termes d'entreprise, 79](#)
- [Commentaires, 80](#)
- [Balises, 81](#)

Présentation des termes d'entreprise, des commentaires et des balises dans Informatica Analyst

Vous pouvez ajouter des termes d'entreprise, des commentaires et des balises à un profil ou à des colonnes du profil. Vous pouvez attribuer, afficher et modifier des termes d'entreprise, des commentaires et des balises dans la vue de résumé et la vue détaillée.

Termes d'entreprise

Vous pouvez attribuer des termes d'entreprise aux colonnes d'un profil dans l'outil Analyst tool. Vous pouvez modifier un lien de ressource ou supprimer un terme d'entreprise d'une colonne. Un glossaire d'entreprise est un ensemble de termes qui définissent des concepts dans un langage d'entreprise spécifique pour les utilisateurs professionnels. Un terme d'entreprise fournit la définition et l'utilisation par l'entreprise d'un concept.

Vous pouvez attribuer, afficher ou supprimer les termes d'entreprise d'une colonne dans la vue de résumé et dans la vue détaillée. Pour afficher le terme d'entreprise dans l'espace de travail **Glossaire**, cliquez sur le terme d'entreprise dans la vue détaillée.

Vous pouvez également modifier les propriétés d'un lien de ressource pour un terme d'entreprise. Vous pouvez également exporter des termes d'entreprise en tant que colonne virtuelle ainsi que les résultats de profil dans un fichier CSV.

Attribution de termes d'entreprise à des colonnes

Attribuez des termes d'entreprise aux colonnes d'un profil dans la vue de résumé et la vue détaillée. Vous pouvez supprimer un terme d'entreprise d'une colonne dans le panneau **Termes d'entreprise**. Vous pouvez également modifier les propriétés d'un lien de ressource pour un terme d'entreprise dans la boîte de dialogue **Modifier le lien de ressource**.

1. Dans la vue de résumé, cliquez avec le bouton droit de la souris sur le nom de la colonne et sélectionnez **Gérer les termes d'entreprise**. Dans la vue détaillée, sélectionnez **Gérer les termes d'entreprise** dans le menu **Actions** de l'onglet **Terme d'entreprise**.

Le panneau **Termes d'entreprise** s'affiche.

2. Cliquez sur l'icône plus (+) ou sur le lien **Attribuer des termes d'entreprise** pour ajouter un terme d'entreprise.

Le panneau **Attribuer des termes d'entreprise** s'affiche.

3. Sélectionnez un terme d'entreprise dans la liste du panneau **Attribuer des termes d'entreprise**. Entrez un nom de ressource. Vous pouvez éventuellement ajouter le contexte et la description de la ressource. Cliquez sur **OK**.

Le terme d'entreprise s'affiche dans le panneau **Termes d'entreprise**.

Commentaires

Vous pouvez ajouter des commentaires à un profil et à ses colonnes pour fournir des informations supplémentaires qui faciliteront la collaboration et l'analyse.

Au niveau du profil, vous pouvez ajouter des commentaires, une définition ou des métadonnées concernant le profil. Vous pouvez afficher les commentaires relatifs à un profil dans la vue de résumé.

Vous pouvez ajouter et afficher les commentaires relatifs à une colonne dans la vue de résumé et dans la vue détaillée.

Vous pouvez effectuer les tâches suivantes sur les commentaires :

- Exporter les commentaires en tant que colonne virtuelle dans un fichier CSV avec les résultats de profil. Le fichier CSV contient tous les commentaires concernant un profil et les colonnes de ce profil.
- Rechercher des résultats de profil à l'aide de mots clés dans la colonne de commentaires.
- Ajouter des commentaires aux colonnes source et aux colonnes virtuelles d'un profil.

Remarque: si vous ne sélectionnez aucune colonne ou si vous n'ajoutez aucun commentaire de colonne, le panneau Commentaires de la vue de résumé affiche des commentaires de profil.

Ajout de commentaires à un profil ou à des colonnes

Vous pouvez ajouter ou afficher un commentaire dans le panneau **Commentaires**.

1. Vous pouvez ajouter des commentaires dans la vue de résumé ou la vue détaillée.

- Pour ajouter des commentaires sur le profil dans la vue de résumé, cliquez sur **Actions > Afficher les commentaires**.
- Pour ajouter des commentaires sur la colonne dans la vue de résumé, cliquez avec le bouton droit de la souris sur une colonne et sélectionnez **Afficher les commentaires**.
- Dans la vue détaillée, cliquez sur **Ajouter un commentaire** dans le volet **Général**.

Le panneau **Commentaires** s'affiche.

2. Cliquez sur **Ajouter un commentaire**.

Une zone de texte s'affiche dans le panneau **Commentaires**.

3. Ajoutez un commentaire descriptif, puis cliquez sur **Enregistrer**.

Le commentaire s'affiche dans le panneau **Commentaires** avec le nom de l'utilisateur actuel ainsi que la date et l'heure de création.

Balises

Vous pouvez attribuer des balises à un profil ou aux colonnes d'un profil pour grouper des objets selon leur utilisation professionnelle.

Vous pouvez afficher ou attribuer des balises de profil dans la vue de résumé. Vous pouvez afficher ou attribuer des balises de colonne dans la vue de résumé et dans la vue détaillée.

Vous pouvez effectuer les tâches suivantes sur les balises de colonne :

- Exporter des balises en tant que colonne virtuelle dans un fichier CSV avec les résultats de profil. Le fichier CSV contient toutes les balises d'un profil et de ses colonnes.
- Attribuer des balises aux colonnes source et aux colonnes virtuelles du profil.

Remarque: Lorsque vous ne sélectionnez aucune colonne ou que vous n'ajoutez aucune balise de colonne, le panneau Balises affiche les balises de profil dans la vue de résumé.

Attribution de balises à un profil ou à des colonnes

Ajoutez des balises à un profil dans la vue de résumé. Ajoutez des balises à une colonne dans la vue de résumé et dans la vue détaillée.

1. Vous pouvez ajouter des balises dans la vue de résumé ou dans la vue détaillée.
 - Pour attribuer des balises à un profil, cliquez sur **Actions > Afficher les balises** dans la vue de résumé.
 - Pour ajouter des balises à une colonne, cliquez avec le bouton droit de la souris sur une colonne et cliquez sur **Afficher les balises** dans la vue de résumé.
 - Cliquez sur **Ajouter une balise** dans le volet **Général** de la vue détaillée.

Le panneau **Balises** s'affiche.

2. Cliquez sur l'icône plus (+) ou sur le lien **Attribuer des balises** pour attribuer une balise.

La boîte de dialogue **Attribuer des balises** s'affiche.

3. Sélectionnez une ou plusieurs balises à attribuer à un profil ou à une colonne. Cliquez sur **OK** pour ouvrir le panneau **Balises**.

Remarque: Pour créer une balise, cliquez sur **Ajouter une nouvelle balise** dans le panneau **Attribuer des balises**.

CHAPITRE 11

Fiches d'évaluation dans Informatica Analyst

Ce chapitre comprend les rubriques suivantes :

- [Présentation des fiches d'évaluation dans Informatica Analyst, 83](#)
- [Processus de fiche d'évaluation de Informatica Analyst, 84](#)
- [Création d'une fiche d'évaluation dans Informatica Analyst, 85](#)
- [Ajouter des colonnes à une fiche d'évaluation existante, 86](#)
- [Exécution d'une fiche d'évaluation, 88](#)
- [Affichage d'une fiche d'évaluation, 88](#)
- [Modification d'une fiche d'évaluation, 89](#)
- [Mesures, 90](#)
- [Groupes de mesures, 91](#)
- [Développement dans les colonnes, 93](#)
- [Graphes de tendance, 93](#)
- [Tableau de bord des fiches d'évaluation dans Informatica Analyst, 97](#)
- [Fichier d'exportation de fiche d'évaluation dans Informatica Analyst, 103](#)
- [Notifications de la fiche d'évaluation, 104](#)
- [Lignage des fiches d'évaluation, 107](#)

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.

Lorsque vous créez ou modifiez une fiche d'évaluation, vous pouvez créer des filtres de fiche d'évaluation en fonction des données source. Les filtres de fiche d'évaluation vous permettent de recalculer les scores de mesure en fonction de la condition de filtre. 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 coûts des données incorrectes dans les mesures sur une certaine période de temps. Vous pouvez réutiliser les filtres de profil dans une fiche d'évaluation.

Lorsque le système de contrôle de version est activé dans l'outil Analyst tool, vous pouvez créer plusieurs versions d'une fiche d'évaluation et en afficher l'historique des versions. Par défaut, la fiche d'évaluation est extraite après sa création. Vous devez archiver la fiche d'évaluation pour que d'autres utilisateurs puissent la modifier.

Vous pouvez afficher le tableau de bord des fiches d'évaluation dans l'espace de travail **Fiches d'évaluation**. Dans le tableau de bord des fiches d'évaluation, vous pouvez afficher les objets de données qui disposent de fiches d'évaluation, les fiches d'évaluation d'un projet, la tendance d'exécution des fiches d'évaluation sur les six derniers mois, une agrégation de scores Bon, Acceptable et Inacceptable pour l'ensemble des exécutions de fiches d'évaluation en un mois.

Informatica Analyst vous permet de configurer et de gérer les notifications par courriel pour les fiches d'évaluation. Utilisez le service de messagerie électronique pour gérer les notifications par courriel. Ce service est un service système que vous pouvez configurer dans Informatica Administrator.

Processus de fiche d'évaluation de Informatica Analyst

Vous pouvez créer et modifier une fiche d'évaluation dans l'outil Developer tool et l'outil Analyst tool. Vous pouvez exécuter une fiche d'évaluation dans l'outil Analyst tool. 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 tool ou l'outil Analyst tool et ajouter des colonnes depuis un profil.
2. Ouvrir une fiche d'évaluation dans l'outil Analyst tool.
3. Après avoir exécuté un profil, ajoutez des colonnes de profil comme mesures à la fiche d'évaluation.

4. Vous pouvez éventuellement créer des filtres de fiche d'évaluation en fonction des données source.
5. Vous pouvez également configurer le coût des données non valides pour chaque mesure.
6. Exécuter la fiche d'évaluation pour générer les scores pour les colonnes.
7. Afficher une fiche d'évaluation pour consulter les scores de chaque colonne dans un enregistrement.
8. Développer un score dans les colonnes.
9. Éditer une fiche d'évaluation.
10. Définissez les seuils de chaque mesure dans une fiche d'évaluation.
11. Créez un groupe pour ajouter ou déplacer les mesures associées de la fiche d'évaluation.
12. Éditer ou supprimer un groupe, selon les besoins.
13. Afficher le graphe de tendance de score pour chaque score afin de surveiller la manière dont le score change au cours du temps.
14. Vous pouvez également afficher le graphe de tendance de coût pour chaque mesure pour surveiller la valeur de la qualité des données.
15. Affichez le lignage des fiches d'évaluation pour chaque métrologie ou groupe de métrologies.
16. Affichez les informations de consolidation concernant les fiches d'évaluation pour lesquelles vous disposez d'autorisations de lecture.

Création d'une fiche d'évaluation dans Informatica Analyst

Créez un fiche d'évaluation et ajoutez-y des colonnes de profil. Vous devez exécuter un profil avant d'ajouter des colonnes à la fiche d'évaluation.

1. Dans l'espace de travail **Bibliothèque**, sélectionnez le projet ou le dossier contenant le profil.

2. Cliquez sur le profil pour l'ouvrir.

Les résultats de profil s'affichent dans la vue de résumé de l'espace de travail **Découverte**.

3. Cliquez sur **Actions > Ajouter à la fiche d'évaluation**.

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

4. Dans l'écran **Ajouter à la fiche d'évaluation**, vous pouvez choisir de créer une fiche d'évaluation ou de modifier une fiche d'évaluation existante pour y ajouter des colonnes. L'option **Nouvelle fiche d'évaluation** est sélectionnée par défaut. Cliquez sur **Suivant**.

5. Dans l'écran **Étape 2 sur 8**, entrez le nom de la fiche d'évaluation. Vous pouvez éventuellement entrer une description de la fiche d'évaluation. Sélectionnez le projet et le dossier dans lequel vous voulez enregistrer la fiche d'évaluation. Cliquez sur **Suivant**.

Par défaut, l'assistant Fiche d'évaluation sélectionne les colonnes et les règles définies dans le profil. Vous ne pouvez pas ajouter de colonnes non comprises dans le profil.

6. Dans l'écran **Étape 8 sur 3**, 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 l'option **Nom de colonne** pour trier les noms de colonnes. Cliquez sur **Suivant**.

7. Dans l'écran **Étape 4 sur 8**, vous pouvez ajouter un filtre au score.

Vous pouvez appliquer aux scores le filtre que vous avez créé pour le profil ou en créer un. Sélectionnez un score dans le volet **Filtres de score** , puis cliquez sur l'icône **Gérer les filtres** pour ouvrir la boîte de dialogue **Modifier le filtre : nom de colonne** . Dans la boîte de dialogue **Modifier le filtre : nom de colonne** , vous pouvez choisir d'exécuter l'une des tâches suivantes :

- Vous pouvez choisir un filtre que vous avez créé pour le profil. Cliquez sur **Suivant**.
- Sélectionnez un filtre existant. Cliquez sur l'icône Modifier pour modifier le filtre dans la boîte de dialogue **Modifier le filtre** . Cliquez sur **Suivant**.
- Cliquez sur l'icône plus (+) pour créer des filtres dans la boîte de dialogue **Nouveau filtre** . Cliquez sur **Suivant**.

Le filtre apparaît dans le volet **Filtres de score** . Vous pouvez appliquer le même filtre à tous les scores de la fiche d'évaluation.

8. Dans l'écran **Étape 4 sur 8** , cliquez sur **Suivant**.
9. Dans l'écran **Étape 5 sur 8** , sélectionnez chaque score dans le volet **Scores** , puis configurez les valeurs valides dans la liste de toutes les valeurs du volet **Score avec : Valeurs** . Vous pouvez effectuer les tâches suivantes dans l'écran **Étape 5 sur 7** :
 - Vous pouvez sélectionner plusieurs valeurs dans le volet **Valeurs disponibles** et cliquer sur le bouton flèche droite pour les déplacer vers le volet **Valeurs valides** . Le nombre total de valeurs valides pour un score s'affiche en haut du volet **Valeurs disponibles** .
 - Dans le volet **Seuils du score** , configurez les seuils du score.
Vous pouvez définir des seuils de scores **Bon**, **Acceptable** et **Inacceptable** .
 - Sélectionnez chaque score et configurez le coût des données non valides. Pour attribuer une valeur constante au coût de la mesure, sélectionnez **Coût fixe**. Cliquez éventuellement sur **Changer l'unité de coût** pour modifier l'unité de coût ou choisissez **Aucun**. Pour attacher une colonne numérique en tant que coût variable au score, sélectionnez **Coût variable**, puis cliquez sur **Sélectionner une colonne** pour sélectionner une colonne numérique.
10. Dans l'écran **Étape 6 sur 8** , vous pouvez sélectionner un groupe de scores auquel vous pouvez ajouter des scores ou créer un groupe. Pour créer un groupe de scores, cliquez sur l'icône de groupe. Cliquez sur **Suivant**.
11. Dans l'écran **Étape 7 sur 8** , spécifiez le poids des scores et les seuils du groupe.
12. Dans l'écran **Étape 8 sur 8** , sélectionnez l'option d'environnement d'exécution **Natif** pour exécuter la fiche d'évaluation.
13. Cliquez sur **Enregistrer** pour enregistrer la fiche d'évaluation, ou sur **Enregistrer et exécuter** pour enregistrer et exécuter la fiche d'évaluation.

La fiche d'évaluation apparaît dans l'espace de travail **Fiche d'évaluation** .

Ajouter des colonnes à une fiche d'évaluation existante

Après avoir exécuté un profil, vous pouvez ajouter des colonnes à une fiche d'évaluation existante, 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.

Les résultats de profil s'affichent dans la vue de résumé.

2. Sélectionnez une colonne. Cliquez sur **Actions > Ajouter à la 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.
3. Sélectionnez **Fiche d'évaluation existante** pour ajouter les colonnes à une fiche d'évaluation prédéfinie. Cliquez sur **Suivant**.
 4. Dans l'écran **Étape 2 sur 7**, sélectionnez la fiche d'évaluation, les scores et le groupe de scores auxquels vous voulez ajouter les colonnes. Cliquez sur **Suivant**.
 5. Dans l'écran **Étape 3 sur 7**, sélectionnez les colonnes et les règles que vous voulez ajouter à la fiche d'évaluation comme mesures. Vous pouvez éventuellement cocher la case dans l'en-tête de la colonne de gauche pour sélectionner toutes les colonnes. Cliquez sur **Nom de colonne** pour trier les noms de colonne. Cliquez sur **Suivant**.
 6. Dans l'écran **Étape 4 sur 7**, vous pouvez créer des filtres pour les scores. Vous pouvez également appliquer aux mesures le filtre que vous avez créé pour le profil.
 7. Dans l'écran **Étape 5 sur 7**, vous pouvez effectuer les tâches suivantes :
 - Dans le volet **Mesures**, sélectionnez chaque mesure et configurez les valeurs de mesure dans les autres volets.
 - Dans le volet **Score avec : des valeurs**, sélectionnez des valeurs multiples dans le volet **Valeurs disponibles** et cliquez sur la flèche de droite pour les déplacer vers le volet **Valeurs valides**. Le nombre total de valeurs valides d'une mesure s'affiche en haut du volet **Valeurs disponibles**.
 - Dans le volet **Seuils des mesures** : vous pouvez configurer des seuils pour un score **Satisfaisant**, **Acceptable** et **Inacceptable**.
 - Dans le volet **Coût des données non valides**, vous pouvez :
 - Sélectionner chaque mesure et configurer le coût des données non valides de la mesure.
 - Sélectionner l'option **Coût fixe** pour attribuer une valeur constante au coût de la mesure. Vous pouvez cliquer sur **Changer l'unité de coût** pour modifier l'unité de coût.
 - Sélectionner l'option **Coût variable** pour joindre une colonne numérique en tant que coût variable à la mesure. Vous pouvez cliquer sur **Sélectionner une colonne** pour sélectionner une colonne numérique.
 8. Cliquez sur **Suivant**.
 9. Dans l'écran **Étape 7 sur 6**, vous pouvez effectuer les tâches suivantes :
 - Sélectionner le groupe de mesures auquel vous souhaitez ajouter les mesures.
 - Dans le volet **Par défaut - Mesures**, vous pouvez double-cliquer sur le poids métrique par défaut, qui est de 0, pour en modifier la valeur.
 - Dans le volet **Seuils des mesures** : vous pouvez configurer des seuils pour un score **Satisfaisant**, **Acceptable** et **Inacceptable**.

10. Cliquez sur **Suivant**.
11. Dans l'écran **Étape 7 sur 7**, sélectionnez un environnement d'exécution.
12. 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**.

Modification 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 l'espace de travail **Bibliothèque** , cliquez sur la fiche d'évaluation à modifier dans le volet **Ressources** .
La fiche d'évaluation apparaît dans l'espace de travail **Fiches d'évaluation** .
2. Si le système de contrôle de version est activé, cliquez sur **Actions** > **Extraire**.
3. Cliquez sur **Actions** > **Modifier** > **Général**.
La boîte de dialogue **Modifier la fiche d'évaluation** s'affiche.
4. Sous l'onglet **Général**, vous pouvez modifier le nom et la description de la fiche d'évaluation si nécessaire.
5. Cliquez sur l'onglet **Scores** .
6. Sélectionnez un score dans le volet **Scores** , puis configurez les valeurs valides dans la liste de toutes les valeurs du volet **Score avec : Valeurs** .
7. Dans le volet **Seuils du score** , vous pouvez modifier les seuils de score si nécessaire.
8. Vérifiez le coût des données non valides pour chaque score et effectuez des modifications si nécessaire.
9. Cliquez sur l'onglet **Filtres de fiche d'évaluation** .
10. Vous pouvez ajouter, modifier ou supprimer des filtres.
11. Cliquez sur l'onglet **Groupes de scores** .
12. Vous pouvez créer, modifier ou supprimer les groupes de scores.
Vous pouvez également modifier les poids et les seuils du score sous l'onglet **Groupes de scores** .
13. Cliquez sur l'onglet **Notifications** .
14. Vous pouvez modifier les paramètres de notification de la fiche d'évaluation si nécessaire.
Vous pouvez configurer les paramètres généraux et personnalisés des scores et groupes de scores.
15. Cliquez sur **Enregistrer** pour enregistrer les modifications dans la fiche d'évaluation, ou sur **Enregistrer et exécuter** pour enregistrer les modifications et exécuter la fiche d'évaluation.
16. Cliquez sur **Archiver**.

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.

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 l'espace de travail **Bibliothèque**, cliquez sur la fiche d'évaluation que vous voulez modifier dans le volet **Ressources**.
La fiche d'évaluation s'affiche dans l'espace de travail **Fiches d'évaluation**.
2. Cliquez sur **Actions > Éditer**.
La fenêtre **Éditer la fiche d'évaluation** s'affiche.
3. 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**.
4. Cliquez sur l'icône **Nouveau groupe** pour créer un groupe de mesures.
La boîte de dialogue **Groupes de mesures** s'ouvre.
5. Entrez un nom et une description facultative.

6. Cliquez sur **OK**.
7. 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 l'espace de travail **Bibliothèque**, cliquez sur la fiche d'évaluation que vous voulez modifier dans le volet **Ressources**.
La fiche d'évaluation s'affiche dans l'espace de travail **Fiches d'évaluation**.
2. Cliquez sur **Actions > Éditer**.
La fenêtre **Éditer la fiche d'évaluation** s'affiche.
3. 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**.
4. 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.
5. Sélectionnez le groupe de mesures vers lequel déplacer les scores.
6. 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 l'espace de travail **Bibliothèque**, cliquez sur la fiche d'évaluation que vous voulez modifier dans le volet **Ressources**.
La fiche d'évaluation s'affiche dans l'espace de travail **Fiches d'évaluation**.
2. Cliquez sur **Actions > Éditer**.
La fenêtre **Éditer la fiche d'évaluation** s'affiche.
3. 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**.
4. Sur le panneau **Groupes de mesures**, cliquez sur l'icône **Éditer le groupe**.
La boîte de dialogue **Éditer** s'ouvre.
5. Entrez un nom et une description facultative.
6. 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 l'espace de travail **Bibliothèque**, cliquez sur la fiche d'évaluation que vous voulez modifier dans le volet **Ressources**.
La fiche d'évaluation s'affiche dans l'espace de travail **Fiches d'évaluation**.
2. Cliquez sur **Actions > Éditer**.
La fenêtre **Éditer la fiche d'évaluation** s'affiche.
3. 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**.
4. 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.
5. 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.
6. 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

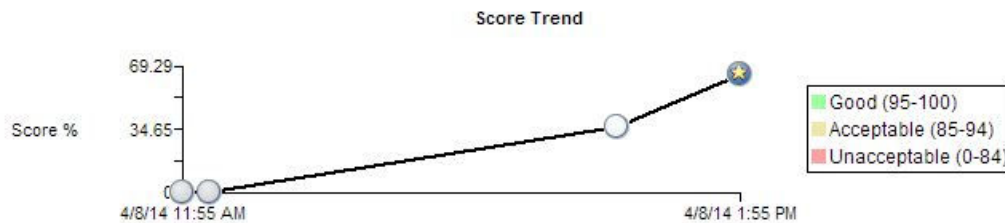
défaut, le graphe de tendance affiche les données à partir des 10 dernières exécutions de la fiche d'évaluation. Vous pouvez afficher le nombre de lignes totales et de lignes non valides pour la mesure dans le graphe de tendance. Le graphe de tendance affiche également si les tendances de score et de 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 graphique de tendance de score est une représentation graphique de la manière dont évoluent les scores de mesure sur plusieurs exécutions de profils. Le graphique 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 graphique 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 graphique visuel permettant une analyse rapide.

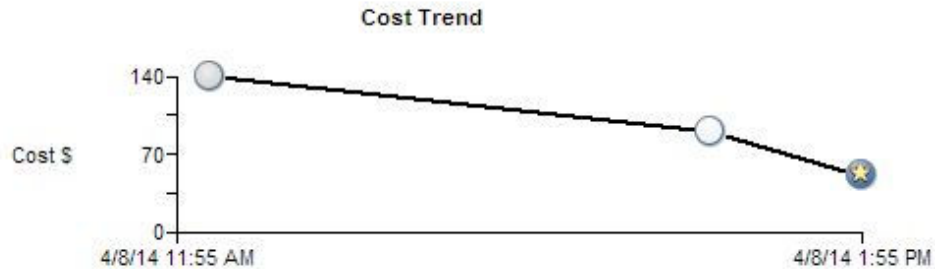
Graphe de tendance de coût

Un graphique 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 graphique de tendance de coût peut mesurer l'incidence de la qualité des données dans une entreprise. Le graphique 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 graphique de tendance de coût.

Un graphique 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 graphique 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 graphique 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

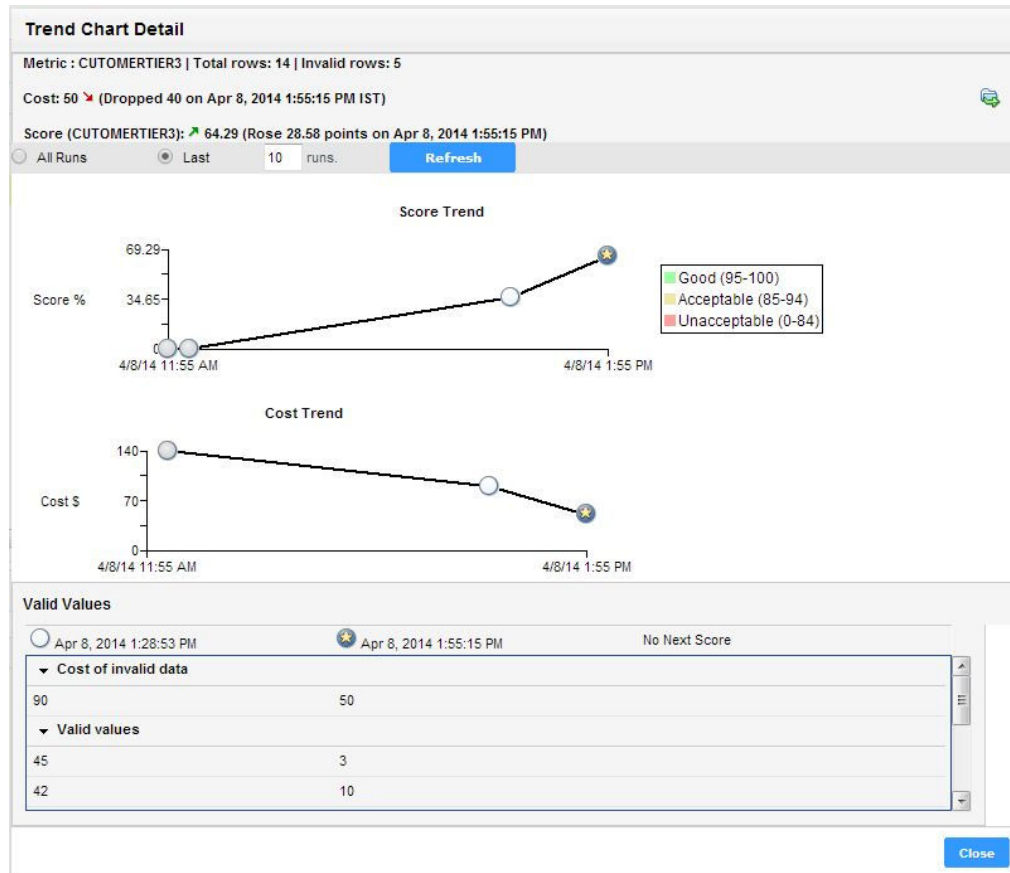
Vous pouvez afficher des graphes de tendance pour chaque mesure afin de surveiller la manière selon laquelle évolue le score ou le coût des données non valides dans le temps.

1. Dans l'espace de travail **Bibliothèque**, sélectionnez le projet ou le dossier contenant 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 graphique de tendance**.

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

L'image suivante présente la boîte de dialogue **Détail du 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**.

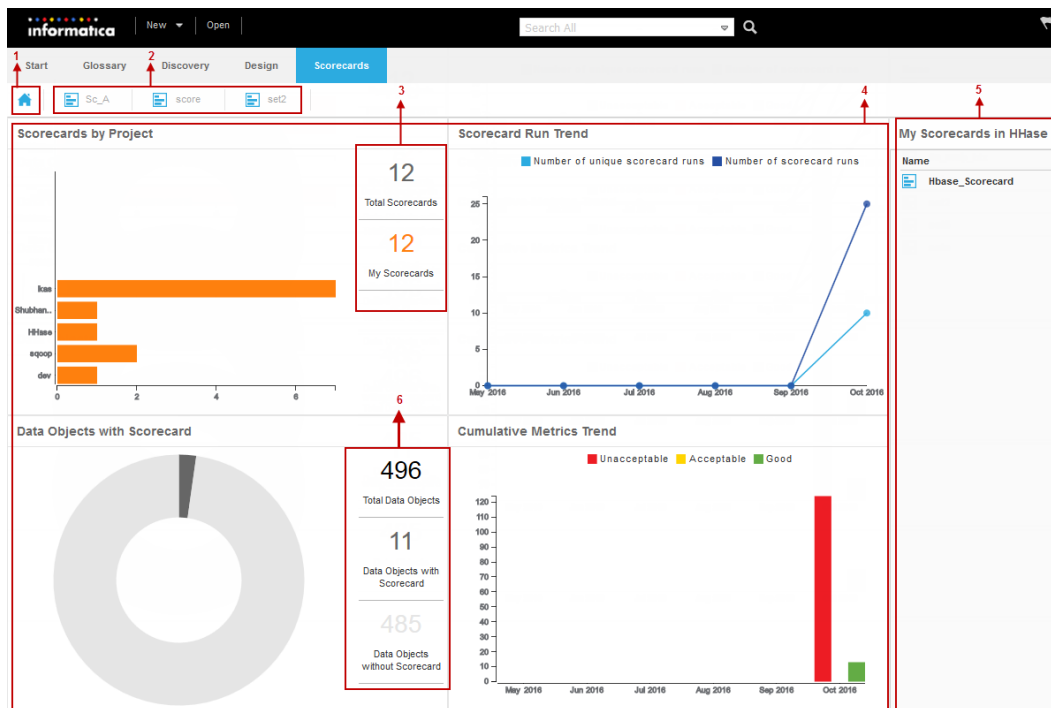
Tableau de bord des fiches d'évaluation dans Informatica Analyst

L'espace de travail **Fiches d'évaluation** dans Informatica Analyst affiche le tableau de bord des fiches d'évaluation. Dans le tableau de bord des fiches d'évaluation, vous pouvez afficher les objets de données avec fiches d'évaluation, la tendance d'exécution des fiches d'évaluation sur les six derniers mois, les fiches d'évaluation d'un projet, une agrégation de scores Bon, Acceptable et Inacceptable pour l'ensemble des exécutions de fiches d'évaluation en un mois, et le volet Liste des ressources.

La tableau de bord des fiches d'évaluation sur votre machine ne s'actualise pas automatiquement lorsque les fiches d'évaluation sont modifiées par d'autres utilisateurs. Utilisez la clé de fonction F5 ou basculez entre les espaces de travail ou les onglets de résultat des fiches d'évaluation pour actualiser le tableau de bord des fiches d'évaluation.

Vous pouvez afficher les données sous la forme de séries de données ou de points de données dans les volets. Les points de données s'affichent sous le forme de petits cercles opaques et les séries de données comme des barres horizontales, des barres verticales ou des tranches dans les graphes.

L'image suivante montre le tableau de bord des fiches d'évaluation et le volet Ressources dans l'espace de travail **Fiches d'évaluation** :



1. Icône du tableau de bord des fiches d'évaluation. Affiche le tableau de bord des fiches d'évaluation.
2. Onglets des résultats des fiches d'évaluation. Affichent les résultats des fiches d'évaluation ouvertes.
3. Légende dans le volet Fiches d'évaluation par projet. Affiche le nombre total de fiches d'évaluation dans l'ensemble des projets et le nombre total de fiches d'évaluation pour lesquelles vous disposez d'autorisations de lecture dans tous les projets.
4. Tableau de bord des fiches d'évaluation. Affiche les volets Fiches d'évaluation par projet, Tendence d'exécution des fiches d'évaluation, Objets de données avec fiche d'évaluation, et Tendence des scores cumulatifs dans le tableau de bord.
5. Volet de la liste des ressources Affiche la liste des fiches d'évaluation ou des objets de données associés à une légende, à une série de données ou à un point de données dans le graphe.
6. Légende dans le volet Objets de données avec fiche d'évaluation. Affiche le nombre total d'objets de données, le nombre d'objets de données avec fiche d'évaluation et le nombre d'objets de données sans fiche d'évaluation

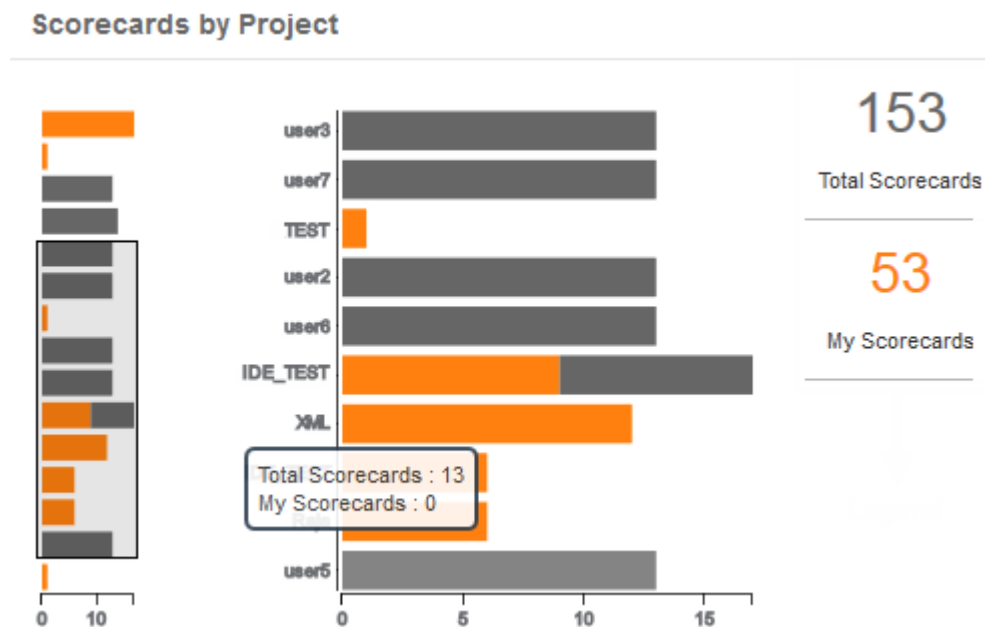
Dès que vous cliquez sur un point de données ou une série de données dans le tableau de bord des fiches d'évaluation, les fiches d'évaluation qui mappent vers le point de données ou la série de données s'affichent dans le volet Liste des ressources. Dès que vous cliquez sur une fiche d'évaluation dans le volet Liste des ressources, les résultats des fiches d'évaluation s'affichent dans l'onglet de l'espace de travail **Fiches d'évaluation**. Le volet Liste des ressources affiche les fiches d'évaluation pour lesquelles vous disposez d'autorisations de lecture.

Fiches d'évaluation par projet

Le volet **Fiches d'évaluation par projet** affiche les projets du référentiel modèle ainsi que le nombre de fiches d'évaluation dans chaque projet sous la forme de graphique à barres. Une barre du graphique à barres représente un projet. L'axe des X du graphique affiche le nombre de fiches d'évaluation et l'axe des Y les projets qui disposent de fiches d'évaluation.

Les fiches d'évaluation d'un projet s'affichent en gris dans le graphique à barres et les fiches d'évaluation pour lesquelles vous disposez d'autorisations de lecture s'affichent en orange dans le graphique à barres. La section **Total des fiches d'évaluation** dans la légende affiche le nombre total des fiches d'évaluation dans le référentiel modèle. La section **Mes fiches d'évaluation** dans la légende affiche le nombre de fiches d'évaluation pour lesquelles vous disposez d'autorisations de lecture dans le référentiel modèle.

L'image suivante présente le volet **Fiches d'évaluation par projet** dans le tableau de bord des fiches d'évaluation :



Vous pouvez afficher les graphiques suivants dans le volet :

- Graphique détaillé. Affiche tous les projets dans le référentiel modèle avec les fiches d'évaluation et le nombre de fiches d'évaluation de chaque projet. Si le nombre de projets est supérieur à 10, le volet **Fiches d'évaluation par projet** affiche une réglette.
- Graphique miniature. Affiche tous les projets et le nombre de fiches d'évaluation de chaque projet dans la réglette du graphique détaillé.

Lorsque vous placez votre curseur sur le graphique miniature, le nombre total de fiches d'évaluation d'un projet et le nombre de fiches d'évaluation pour lesquelles vous disposez d'autorisations de lecture s'affichent dans une étiquette de données.

Pour afficher les fiches d'évaluation d'un projet pour lesquelles vous disposez d'autorisations de lecture, cliquez sur la partie orange de la barre horizontale. Pour afficher les fiches d'évaluation pour lesquelles vous disposez d'autorisations de lecture dans le référentiel modèle, cliquez sur **Mes fiches d'évaluation** dans le graphique à barres. Les fiches d'évaluation s'affichent dans le volet Liste des ressources. Cliquez sur une fiche d'évaluation dans le volet Ressources pour afficher les résultats de la fiche d'évaluation.

Tendance d'exécution des fiches d'évaluation

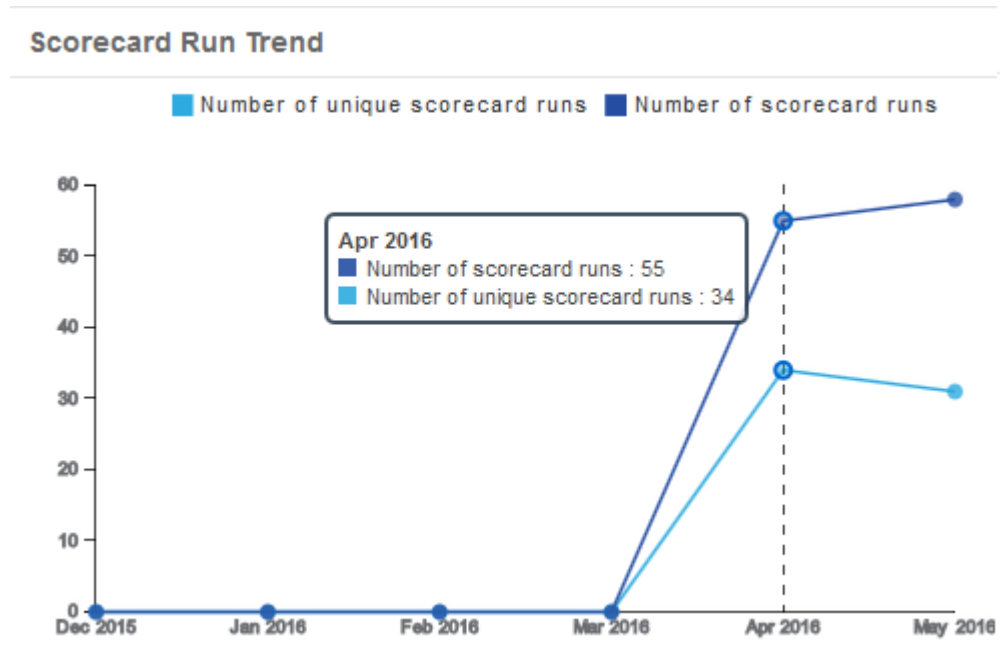
Le volet **Tendance d'exécution des fiches d'évaluation** affiche la tendance d'exécution des fiches d'évaluation du mois en cours et des cinq derniers mois sous la forme de graphiques en courbes avec des marqueurs. L'axe des X du graphique affiche le mois en cours et les cinq derniers mois et l'axe des Y affiche le nombre de fiches d'évaluation. Un marqueur est un point de données sur le graphique en courbes. Si vous déplacez votre curseur sur un marqueur d'un graphique, le résumé d'exécution de la fiche d'évaluation pour le mois s'affiche dans une étiquette de données.

Vous pouvez afficher les marqueurs dans le volet :

- **Nombre d'exécutions de fiches d'évaluation.** Le marqueur affiche le nombre total d'exécutions de fiches d'évaluation au cours du mois.

- **Nombre d'exécutions de fiches d'évaluation uniques.** Le marqueur affiche le nombre total d'exécutions de fiches d'évaluation uniques au cours du mois.

L'image suivante présente le volet **Tendance d'exécution des fiches d'évaluation** dans le tableau de bord des fiches d'évaluation :



Lorsque vous cliquez sur un marqueur du volet, les fiches d'évaluation qui mappent vers le marqueur s'affichent dans le volet Liste des ressources. Vous pouvez afficher les fiches d'évaluation pour lesquelles vous disposez d'autorisations de lecture. Cliquez sur une fiche d'évaluation dans le volet Liste des ressources pour afficher les résultats de la fiche d'évaluation dans l'espace de travail **Fiches d'évaluation**.

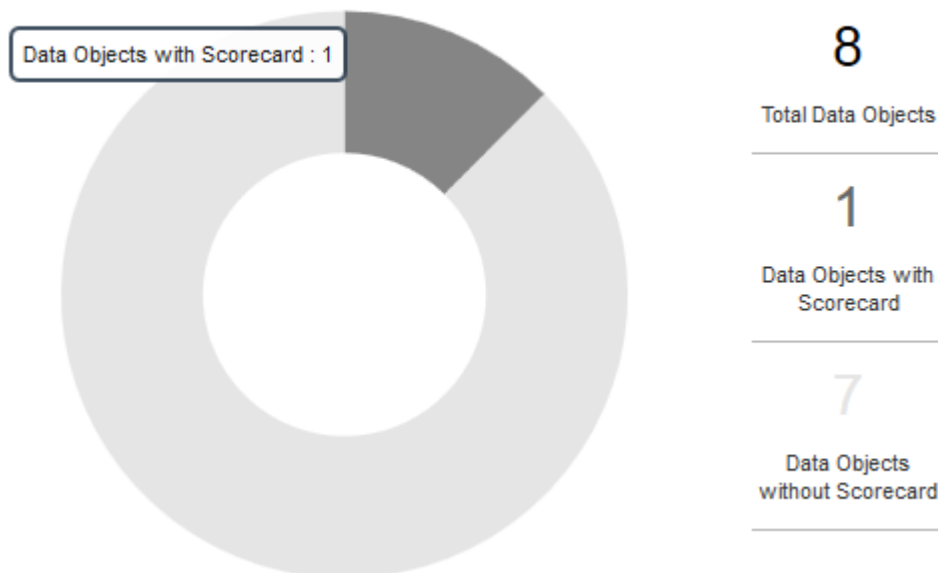
Objets de données avec fiche d'évaluation

Le volet **Objets de données avec fiche d'évaluation** affiche un graphique en anneau. Dans ce graphique, vous pouvez afficher le nombre d'objets de données avec des fiches d'évaluation et le nombre d'objets de données sans fiche d'évaluation sous la forme de tranches.

Lorsque vous placez le curseur sur le graphique en anneau, les données qui mappent vers la tranche apparaissent dans une étiquette de données.

L'image suivante représente le volet **Objets de données avec fiche d'évaluation** du tableau de bord des fiches d'évaluation :

Data Objects with Scorecard



La légende du volet **Objets de données avec fiche d'évaluation** affiche les statistiques de données suivantes :

- Total des objets de données. Affiche le nombre total d'objets de données dans le dossier **Objet de données** du volet **Ressources** dans l'espace de travail **Bibliothèque**. Les objets de données comprennent les objets de données logiques et les objets de données personnalisés.
- Objets de données avec fiche d'évaluation. Affiche le nombre d'objets de données avec fiche d'évaluation.
- Objets de données sans fiche d'évaluation. Affiche le nombre d'objets de données sans fiche d'évaluation.

Après que vous cliquez sur les tranches du graphique en anneau ou sur la légende **Objets de données avec fiche d'évaluation** et **Objets de données sans fiche d'évaluation**, les fiches d'évaluation qui mappent vers la tranche du graphique en anneau ou la légende apparaissent dans le volet Liste des ressources.

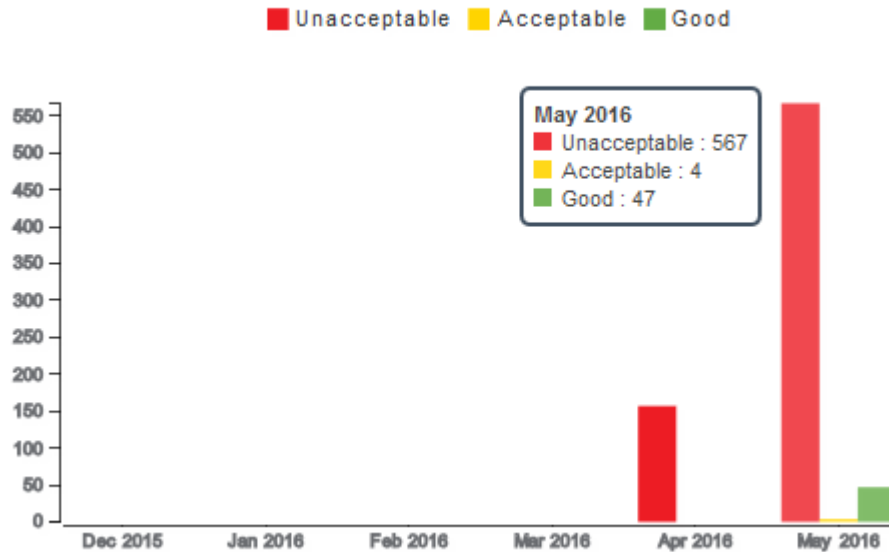
Tendance de scores cumulatifs

Le volet **Tendance de scores cumulatifs** affiche des histogrammes. Vous pouvez afficher l'agrégation des scores Bon, Acceptable et Inacceptable pour toutes les exécutions de fiche d'évaluation sur un mois dans l'histogramme sous la forme de barres verticales. Vous pouvez utiliser le volet **Tendance de scores cumulatifs** pour afficher et analyser la tendance des scores pour le mois en cours et les cinq derniers mois.

Lorsque vous déplacez le curseur sur une barre verticale, le résumé des scores pour le mois s'affiche dans une étiquette de données. Lorsque vous cliquez sur une barre verticale du volet, les fiches d'évaluation pertinentes s'affichent dans le volet Liste des ressources. Vous pouvez afficher les fiches d'évaluation pour lesquelles vous disposez d'autorisations de lecture. Cliquez sur une fiche d'évaluation dans le volet Liste des ressources pour afficher les résultats des fiches d'évaluation.

L'image suivante présente le volet **Tendance de scores cumulatifs** dans le tableau de bord des fiches d'évaluation :

Cumulative Metrics Trend



Le volet Liste des ressources affiche quelques fiches d'évaluation dans la liste des scores inacceptables et dans la liste des bons scores lorsque la tendance d'une fiche d'évaluation évolue avec le temps en un mois. Pour analyser les scores, ouvrez la fiche d'évaluation pour afficher les résultats de la fiche d'évaluation.

Exemple

Vous êtes le directeur régional d'une chaîne de points de vente dans l'État de Californie. Vous créez la fiche d'évaluation Sales_SC dans la table Sales.

Vous définissez les seuils suivants pour le score Sales_amt dans la fiche d'évaluation Sales_SC :

- Inacceptable = 0 % à 40 %
- Acceptable = 41 % à 89 %
- Bon = 90 % à 100 %

Pour saisir les données relatives aux ventes quotidiennement, vous utilisez le service de planificateur pour exécuter les fiches d'évaluation toutes les nuits sur la table Sales. Vous préparez un rapport mensuel pour la gestion du mois de mai et vous utilisez le tableau de bord des fiches d'évaluation pour vérifier votre rapport. Pour le mois de mai, lorsque vous cliquez sur les barres verticales pour consulter les scores inacceptables et les bons scores dans le graphe, la fiche d'évaluation Sales_SC s'affiche dans le volet Liste des ressources.

Lorsque vous analysez les ventes du mois, vous constatez la tendance suivante :

1. Du 1er mai au 25 mai, le score Sales_amt est calculé en dessous des 40 % et est marqué comme score inacceptable.
2. Pour la dernière semaine de mai, grâce à une hausse des ventes, le score Sales_amt se situe au-dessus des 98 % et est marqué comme bon score.

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 tool enregistre le fichier au format « .xlsx ».

Lorsque vous exportez une fiche d'évaluation, vous pouvez choisir d'exporter le résumé de la fiche d'évaluation, les graphiques de tendance, les lignes qui ne sont pas valides ainsi que les propriétés de la fiche d'évaluation dans le fichier Microsoft Excel. Vous pouvez également exporter uniquement le résumé de la fiche d'évaluation, les graphiques de tendance et les propriétés de la fiche d'évaluation dans le fichier Microsoft Excel.

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

Lorsque vous exportez les résultats d'une fiche d'évaluation, l'outil Analyst tool enregistre les informations dans plusieurs feuilles de calcul dans un fichier Microsoft Excel. Le résumé de la fiche d'évaluation, les graphiques de tendance, les lignes non valides et les propriétés de la fiche d'évaluation s'affichent en tant que feuilles de calcul dans le fichier. L'outil Analyst tool enregistre le fichier au format « .xlsx ».

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

Onglet	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 non valides, le score et le poids métrique.
Graphe de tendance	Graphes de tendance des scores.
Lignes non valides	Détails des lignes non valides pour chaque colonne. L'outil Analyst tool exporte un maximum de 100 lignes vers la feuille de calcul. La feuille de calcul Lignes non valides s'affiche lorsque vous choisissez l'option Données > Tout dans la boîte de dialogue Exporter les données vers un fichier .
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. Dans la liste **Données**, vous pouvez choisir les statistiques à exporter vers le fichier :

- Choisissez l'option **Tout** pour exporter le résumé de la fiche d'évaluation, les graphiques de tendance, les lignes non valides et les propriétés de la fiche d'évaluation dans le fichier.
 - Sélectionnez l'option **Vue Résumé** pour exporter le résumé de la fiche d'évaluation, les graphiques de tendance, les lignes non valides et les propriétés de la fiche d'évaluation dans le fichier.
5. Sélectionnez la page de codes du fichier.
 6. 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.

Avant de configurer des fiches d'évaluation de façon à envoyer des notifications par courriel, un administrateur doit configurer service de messagerie électronique dans l'outil Administrator tool.

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.
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.

L'ID de courriel de l'expéditeur par défaut est **Le courriel de l'expéditeur** configuré dans les propriétés du domaine SMTP.

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**.

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 l'espace de travail **Bibliothèque**, cliquez sur la fiche d'évaluation à afficher dans le volet **Ressources**.

La fiche d'évaluation s'affiche dans l'espace de travail **Fiches d'évaluation**.

2. Dans la vue **Fiche d'évaluation**, sélectionnez une métrologie ou un groupe de métrologies.

3. 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 177](#).

CHAPITRE 12

Découverte des domaines de données dans Informatica Analyst

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la découverte de domaines de données dans Informatica Analyst, 109](#)
- [Glossaire de domaine de données dans Informatica Analyst, 110](#)
- [Options de découverte des domaines de données dans Informatica Analyst, 112](#)
- [Création d'un profil de colonne pour découvrir des domaines de données dans Informatica Analyst, 116](#)
- [Modification d'un profil dans Informatica Analyst, 117](#)
- [Exécution d'un profil pour effectuer la découverte de domaines de données, 117](#)
- [Résultats de la découverte de domaines de données dans Informatica Analyst, 118](#)
- [Fichiers d'exportation de la découverte des domaines de données dans Informatica Analyst, 119](#)

Présentation de la découverte de domaines de données dans Informatica Analyst

Lorsque vous créez un profil pour effectuer la découverte des domaines de données, vous pouvez sélectionner 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 choisir un critère de conformité pour la découverte de domaines de données et exclure des valeurs Null lors de la découverte des domaines de données.

Vous pouvez créer un profil avec une option d'échantillonnage et des filtres pour effectuer la découverte des domaines de données. Lorsque vous exécutez le profil, vous appliquez l'option d'échantillonnage et les filtres sur la source de données et générez un ensemble de données. Le processus de découverte des domaines de données utilise l'ensemble de données pour découvrir les domaines de données.

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'affichent dans la vue de résumé.
2. Dans la vue de résumé, cliquez sur une colonne pour en afficher les résultats dans la vue détaillée.
3. Dans le volet **Valeurs** ou **Modèles**, sélectionnez la valeur ou le modèle sur lequel vous voulez baser la création d'un domaine de données.
4. Cliquez avec le bouton droit de la souris sur la valeur ou le modèle, puis sélectionnez **Nouveau domaine de données**.
La boîte de dialogue **Créer un domaine de données** s'affiche.
5. Entrez le nom du domaine de données et éventuellement une description.
6. Cliquez sur **Créer**.
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

Vous pouvez cliquer sur **Modifier** dans l'écran **Spécifier les paramètres** pour choisir les colonnes que vous voulez exécuter comme faisant partie de la découverte de domaines de données. Vous pouvez afficher toutes les colonnes de la source de données dans l'écran **Sélectionner une source** de l'assistant de profil.

Vous pouvez choisir différentes colonnes pour le profil de colonne et de découverte de domaines de données.

Le tableau suivant décrit les propriétés de la boîte de dialogue **Modifier** de la découverte de domaines de données :

Option	Description
Nom	Affiche le nom de la colonne.
Type	Affiche le type de données documenté de la colonne.
Précision	Affiche la précision maximale de la colonne.
Échelle	Affiche l'échelle de la colonne.
Valeur Null autorisée	Indique une colonne pouvant contenir des valeurs Null.
Clé	Indique si la colonne est documentée comme une clé primaire ou étrangère.

Sélection des domaines de données dans Informatica Analyst

Le volet **Domaine de données** de l'écran **Spécifier les paramètres** énumère tous les domaines de données du glossaire de domaines 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 propriétés **Domaines de données** de la découverte de domaines de données :

Option	Description
Nom	Affiche le nom du domaine de données. Vous pouvez choisir un ou plusieurs domaines de données ou un groupe de domaines de données.
Description	Description du domaine de données.
DomainGroups	Affiche le nom du groupe de domaines de données auquel le domaine de données appartient.

Options d'inférence des domaines de données dans Informatica Analyst

Les options d'inférence déterminent si la découverte de domaines de données doit s'exécuter sur les données de colonne, les noms de colonne ou les deux. Vous pouvez spécifier le nombre maximal de lignes source que le profil peut analyser. Vous pouvez choisir un critère de conformité pour la découverte de domaines de données. Vous pouvez exclure les valeurs Null de la découverte de domaines de données. Vous pouvez définir les options d'inférence du domaine de données dans l'écran **Spécifier les paramètres** de l'assistant de profil.

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

Option	Description
Données	Permet d'exécuter le profil sur les données de colonne.
Colonnes	Permet d'exécuter le profil sur les titres de colonne.
Données et colonnes	Permet d'exécuter le profil sur les données de colonne et les titres de colonne.
Pourcentage minimal de lignes	Pourcentage minimal de conformité des lignes dans l'ensemble de données requis pour une correspondance de domaine de données.
Nombre minimal de lignes	Nombre minimal de lignes dans l'ensemble de données requis pour une correspondance de domaine de données.
Exclure les valeurs Null de la découverte de domaines de données	Exclut les valeurs Null de l'ensemble de données pour la découverte de domaines de données.
Modifier	Permet de sélectionner les colonnes de la découverte de domaines de données..
Toutes les lignes	Permet d'exécuter le profil sur toutes les lignes de la source.
Échantillon 1re ligne	Permet de choisir le nombre maximal de lignes sur lesquelles le profil peut s'exécuter. L'outil Analyst tool choisit les lignes à partir de la première ligne dans la source.
Échantillon aléatoire	Permet de choisir un échantillon aléatoire de lignes depuis la source de données.
Échantillon aléatoire (auto)	L'outil Analyst tool choisit un échantillon aléatoire de lignes en fonction de la taille de la source de données.
Excluez les types et les domaines de données approuvés de l'inférence du type et du domaine de données dans les exécutions de profil suivantes.	Exclut le type de données approuvé ou le domaine de données du type de données et l'inférence de domaine de données de l'exécution de profil suivante.

Pourcentage de conformité minimal

Vous pouvez choisir un pourcentage minimal de lignes dans l'ensemble de données comme critère de conformité pour la découverte de domaines de données.

Le pourcentage de conformité correspond au 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.

Exemple

Vous disposez d'une source de données de 10 000 lignes parmi lesquelles 2 500 lignes contiennent des numéros de sécurité sociale dans la colonne Commentaires. Vous créez un profil de colonne à l'aide de la découverte de domaines de données et définissez un pourcentage minimal de lignes de 30 % comme critères de conformité. Lorsque vous exécutez le profil, les résultats du profil n'affichent pas les numéros de sécurité sociale comme un domaine de données inféré, car le critère de conformité minimum est de 30 % des lignes de la source de données, c'est-à-dire 3 000 lignes.

Nombre minimal de lignes conformes

Vous pouvez choisir un nombre minimal de lignes dans l'ensemble de données comme critère de conformité pour la découverte de domaines de données.

Exemple

Vous disposez d'une source de données de 10 000 lignes parmi lesquelles 3 lignes contiennent une adresse électronique dans la colonne Commentaires. Vous créez un profil de colonne à l'aide d'un profil de découverte de domaines de données et définissez le nombre minimal de lignes sur 1 comme critère de conformité. Lorsque vous exécutez le profil, les résultats affichent l'adresse électronique comme un domaine de données inféré comprenant trois lignes conformes ainsi que les autres domaines de données inférés.

Exclure les valeurs Null de la découverte de domaines de données

Vous pouvez exclure les valeurs Null lorsque vous effectuez la découverte de domaines de données sur une source de données. Lorsque vous sélectionnez le pourcentage minimal de lignes avec l'option d'exclusion des valeurs Null, le pourcentage de conformité est le rapport entre le nombre de lignes correspondantes et le nombre total de lignes moins les valeurs Null de la colonne.

Le processus de découverte des domaines de données est différent lorsque vous sélectionnez l'option **Exclure les valeurs Null de la découverte de domaines de données** et plusieurs options d'échantillonnage ou des filtres.

Les scénarios suivants expliquent les résultats de la découverte de domaine de données lorsque vous sélectionnez l'option Exclure les valeurs Null avec une option d'échantillonnage et des filtres :

- Option d'échantillonnage **Toutes les lignes** et aucun filtre. La découverte de domaines de données ignore toutes les valeurs Null de la colonne.
- Une option d'échantillonnage et aucun filtre. La découverte de domaines de données ignore toutes les valeurs Null dans les données de l'échantillon et s'exécute sur le reste de ces données.
- Option d'échantillonnage **Toutes les lignes** et des filtres. La découverte de domaines de données ignore toutes les valeurs Null dans les données filtrées et s'exécute sur le reste de ces données.
- Une option d'échantillonnage et des filtres. La découverte de domaines de données ignore les valeurs Null dans les données filtrées de l'échantillon et s'exécute sur le reste de ces données.

Exemple

Vous disposez d'une source de données de 10 000 lignes parmi lesquelles 3 000 lignes comprennent des numéros de sécurité sociale dans la colonne Commentaires. Vous créez un profil de colonne avec la découverte de domaines de données et choisissez les options suivantes :

- Sélectionnez l'option **Exclure les valeurs Null de la découverte de domaines de données**.
- Sélectionnez l'option d'échantillonnage **Toutes les lignes**.
- Sélectionnez l'option **Pourcentage minimal de lignes** et configurez-la sur 12 %.

Lorsque vous exécutez le profil, il s'exécute sur l'ensemble des données et ignore les valeurs Null lors de la découverte de domaines de données.

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

Avant de créer un profil pour effectuer une découverte de domaines de données dans l'outil Analyst tool, vous devez créer au moins un domaine de données. 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 l'espace de travail **Découverte**, cliquez sur **Profil** ou sélectionnez **Nouveau > profil** de n'importe où dans l'outil Analyst tool.
L'assistant **Nouveau profil** s'affiche.
2. L'option **Source unique** est sélectionnée par défaut. Cliquez sur **Suivant**.
3. Dans l'écran **Spécifier les propriétés générales**, entrez un nom et une description facultative pour le profil. Dans le champ Emplacement, sélectionnez le projet ou le dossier dans lequel vous souhaitez créer le profil. Cliquez sur **Suivant**.
4. Dans l'écran **Sélectionner une source**, cliquez sur **Choisir** pour sélectionner un objet de données, ou sur **Nouveau** pour importer un objet de données. Cliquez sur **Suivant**.
5. Dans l'écran **Spécifier les paramètres**, choisissez d'exécuter un profil de colonne, une découverte de domaines de données ou un profil de colonne et une découverte de domaines de données. Par défaut, l'option de profil de colonne est sélectionnée.
 - Choisissez **Exécuter la découverte de domaines de données** pour effectuer la découverte de domaines de données. Sélectionnez les options de domaine de données dans le volet **Domaine de données**.
 - Choisissez **Exécuter le profil de colonne** et **Exécuter la découverte de domaines de données** pour exécuter le profil de colonne et la découverte de domaines de données. Sélectionnez les options de domaine de données dans le volet **Domaine de données**.
Remarque: Par défaut, les colonnes sélectionnées pour le profil de colonne s'appliquent également à la découverte de domaines de données. Cliquez sur **Modifier** pour sélectionner ou désélectionner des colonnes pour la découverte de domaines de données indépendamment des colonnes sélectionnées pour le profil de colonne.
 - Choisissez les données et/ou les colonnes sur lesquelles vous souhaitez exécuter la découverte de domaines de données.
 - Choisissez une option d'échantillonnage dans le volet **Exécuter le profil sur**.
 - Choisissez une option d'exploration dans le volet **Exploration**. Cliquez éventuellement sur **Sélectionner les colonnes** pour sélectionner les colonnes à explorer. Vous pouvez choisir d'exclure l'inférence du type de données et du domaine de données pour les colonnes associées à un type de données ou un domaine de données approuvé.
 - Choisissez un critère de conformité, puis sélectionnez l'option **Exclure les valeurs Null de la découverte de domaines de données**.
 - Choisissez l'option **d'environnement d'exécution natif**.
Remarque: Vous ne pouvez pas exécuter de profils sur le moteur Hive ou Blaze dans l'environnement d'exécution Hadoop, car l'environnement Hadoop n'est pas pris en charge dans la version actuelle d'Informatica.
6. Dans l'écran **Spécifiez les règles et les filtres**, vous pouvez ajouter, modifier ou supprimer des règles et des filtres pour le profil.
7. Cliquez sur **Enregistrer et terminer** pour créer le profil, ou sur **Enregistrer et exécuter** pour créer et exécuter le profil.

Modification d'un profil dans Informatica Analyst

Vous pouvez modifier les propriétés d'un profil après l'avoir exécuté. Si vous avez exécuté un profil de colonne dans le cadre d'une découverte de domaines de données, vous pouvez modifier les paramètres de profil de colonne.

1. Dans l'espace de travail **Bibliothèque**, sélectionnez le projet qui contient le profil, ou sélectionnez le profil dans le volet **Ressources**.
2. Cliquez sur le nom du profil.
La vue de résumé s'affiche dans l'espace de travail **Découverte**.
3. Si le système de contrôle de version est activé, cliquez sur **Actions** > **Extraire** pour extraire le profil.
4. Cliquez sur **Actions** > **Modifier le profil**.
L'assistant de **profil** s'affiche.
5. Sélectionnez l'une des options suivantes selon les modifications que vous souhaitez effectuer :
 - **Spécifiez les propriétés générales.** Modifiez les propriétés de base telles que le nom, la description et l'emplacement.
 - **Sélectionnez une source.** Choisissez une autre source de données correspondante et des colonnes sur lesquelles exécuter le profil.
 - **Spécifiez les paramètres.** Choisissez d'exécuter le profil de colonne uniquement ou le profil de colonne avec la découverte de domaines de données. Modifiez les options de domaine de données, d'échantillonnage et d'exploration.
 - **Spécifiez les règles et les filtres.** Créez, modifiez ou supprimez des règles et des filtres.
6. Cliquez sur **Enregistrer et terminer** pour modifier le profil ou sur **Enregistrer et exécuter** pour modifier et exécuter le profil.
7. Si le système de contrôle de version est activé, vous devez exécuter les tâches suivantes :
 - Cliquez sur **Enregistrer et terminer** pour terminer la modification du profil.
 - Dans la vue de résumé, cliquez sur **Archiver** pour archiver le profil.
 - Cliquez sur **Actions** > **Exécuter le profil** pour exécuter le profil.

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

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

1. Dans le **navigateur de la bibliothèque**, sélectionnez le projet ou le dossier contenant le profil dans le volet Projets, ou le profil dans le volet Ressources.
2. Cliquez sur **Actions** > **Ouvrir**.
La vue de résumé s'ouvre dans l'espace de travail **Découverte**.
3. Cliquez sur **Actions** > **Exécuter le profil**.
L'outil Analyst tool lance une exécution de profil et affiche les résultats dans la vue de résumé.

4. Cliquez sur une colonne de la vue de résumé pour afficher les résultats de colonne.
La vue détaillée s'affiche.

Résultats de la découverte de domaines de données dans Informatica Analyst

Vous pouvez afficher les résultats de la découverte de domaines de données dans la vue de résumé et dans la vue détaillée.

Le champ de domaine de données affiche des statistiques concernant les colonnes qui correspondent aux domaines de données. Dans la vue de résumé, vous pouvez afficher les domaines de données inférés ainsi que le pourcentage de lignes conformes et le nombre de lignes conformes.

La vue détaillée vous permet d'effectuer les tâches suivantes :

- Affichez les domaines de données inférés avec le pourcentage de lignes conformes et le nombre de lignes conformes dans un graphique à barres horizontales.
- Explorer les résultats de lignes conformes, non conformes et de valeurs Null.
- Approuver, rejeter ou réinitialiser le domaine de données.
- Afficher ou masquer les domaines de données rejetés.
- Exécuter la découverte de domaines de données sur toutes les lignes de la source de données pour découvrir les domaines de données inférés.

Approbation des domaines de données

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

1. Dans l'espace de travail **Bibliothèque**, sélectionnez le projet ou le dossier contenant le profil.
2. Cliquez sur le profil pour l'ouvrir.
Les résultats de profil s'affichent dans la vue de résumé.
3. Cliquez sur la colonne pour laquelle vous voulez approuver le domaine de données.
Les résultats de colonne s'affichent dans la vue détaillée.
4. Dans la vue détaillée, sélectionnez le domaine de données dans le volet **Domaine de données**. 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 le domaine de données, 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 de bibliothèque, sélectionnez le projet ou le dossier qui contient le profil.
2. Cliquez sur le profil pour l'ouvrir.

Les résultats de profil s'affichent dans la vue de résumé.

3. Cliquez sur la colonne pour laquelle vous voulez rejeter le domaine de données.

Les résultats de colonne s'affichent dans la vue détaillée.

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 de domaines de données depuis l'outil Analyst tool, vous pouvez spécifier les valeurs de nom 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 de domaines de données dans Microsoft Excel

Lorsque vous exportez les résultats de la découverte de domaines de données vers Microsoft Excel, l'outil Analyst tool enregistre les noms de colonne, les noms de domaines de données correspondants, les critères de conformité et les valeurs Null. Le fichier Excel contient également les noms des groupes de domaines 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 :

Onglet	Description
Vue par colonnes	Résultats de la découverte de domaines de données triés par colonne de source de données.
Vue par domaines de données	Résultats de la découverte de domaines de données triés par domaine de données.
Vue par groupes de domaines de données	Résultats de la découverte de domaines de données triés par groupes de domaines de données.
Propriétés	Propriétés de base 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.

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

Vous pouvez exporter les résultats de la découverte de 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 de domaines de données.
2. Dans la vue de résumé ou dans la vue détaillée, cliquez sur **Actions > Exporter les données**.
La boîte de dialogue **Exporter les données vers un fichier** s'ouvre.
3. Entrez le nom de fichier. Vous pouvez éventuellement utiliser le nom de fichier par défaut.
4. Sélectionnez la page de codes du fichier.
5. Cliquez sur **OK**.

CHAPITRE 13

Découverte des données d'entreprise dans Informatica Analyst

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la découverte de données d'entreprise dans Informatica Analyst, 121](#)
- [Processus de découverte des données d'entreprise dans Informatica Analyst, 122](#)
- [Options de configuration pour la découverte des données d'entreprise, 122](#)
- [Création d'un profil de découverte des données d'entreprise dans Informatica Analyst, 124](#)
- [Modification des options de découverte de données d'entreprise, 126](#)

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

La découverte de 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 de 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 de 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 de 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 de données d'entreprise génère un résumé consolidé des résultats de profil.

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

Vous pouvez sélectionner un profil de système d'exploitation dans Informatica Analyst. Après avoir choisi un profil de système d'exploitation, le service d'intégration de données crée et exécute les profils de découverte

de données d'entreprise selon les autorisations de l'utilisateur du système d'exploitation défini dans le profil du système d'exploitation.

Processus de découverte des données d'entreprise dans Informatica 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 choisir un critère de conformité pour la découverte de domaines de données. Vous pouvez exclure les valeurs Null de la découverte 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.
Pourcentage de conformité minimal	Pourcentage minimal de conformité des lignes dans l'ensemble de données requis pour une correspondance de domaine de données. Le pourcentage de conformité correspond au 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.
Nombre minimal de lignes conformes	Nombre minimal de lignes dans l'ensemble de données requis pour une correspondance de domaine de données.
Exclure les valeurs Null de la découverte de domaines de données	Exclut les valeurs Null de l'ensemble de données pour la découverte de domaines de données.
Exclure les colonnes avec des domaines de données approuvés	Exclut les colonnes qui contiennent des domaines de données approuvés issus de l'inférence de domaine de données de l'exécution de profil.
Toutes les lignes	Exécute la découverte de domaines de données sur toutes les lignes source.
Premier	Nombre maximal de lignes sur lesquelles le profil peut s'exécuter. L'outil Analyst choisit les lignes à partir de la première ligne dans la source.

Paramètres de profil de colonne

Les options d'échantillonnage déterminent si l'outil Analyst tool 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 paramètres de profil de colonne que vous pouvez configurer pour un profil de 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.
Excluez les types et les domaines de données approuvés de l'inférence du type et du domaine de données dans les exécutions de profil suivantes.	Exclut le type de données approuvé ou le domaine de données du type de données et l'inférence de domaine de données de l'exécution de profil suivante.

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

Option	Description
Toutes les lignes	Exécute un profil de colonne sur toutes les lignes de la source de données.
<nombre> premières lignes	Nombre de lignes sur lesquelles vous voulez exécuter le profil de colonne. L'outil Analyst tool choisit les lignes à partir de la première ligne dans la source de données.

Le tableau suivant décrit l'option d'environnement d'exécution que vous pouvez configurer pour un profil de découverte des données d'entreprise :

Option	Description
Natif	L'outil Analyst tool envoie les tâches de profil au module de service de profilage. Celui-ci divise alors les tâches de profil en un ensemble de mappages. Le service d'intégration de données exécute les mappages et enregistre les résultats de profil dans l'entrepôt de profils.

Création d'un profil de découverte des données d'entreprise dans Informatica 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 Informatica Analyst.

1. Dans l'espace de travail **Découverte** , sélectionnez **Nouveau > profil**.
L'assistant **Nouveau profil** s'affiche.
2. Sélectionnez **Découverte des données d'entreprise** . Cliquez sur **Suivant**.
L'onglet **Spécifier les propriétés générales** s'affiche.

3. Sous l'onglet **Spécifier les propriétés générales** , entrez un nom pour le profil de découverte des données d'entreprise et une description facultative. Dans le champ Emplacement, sélectionnez le projet ou le dossier dans lequel vous souhaitez créer le profil. Cliquez sur **Suivant**.
L'onglet **Sélectionner les objets de données** s'affiche.
 4. Sous l'onglet **Sélectionner les objets de données** , cliquez sur **Choisir**.
La boîte de dialogue **Choisir les objets de données** s'affiche.
 5. Dans la boîte de dialogue **Choisir les objets de données** , choisissez un ou plusieurs objets de données à ajouter au profil. Cliquez sur **Enregistrer**.
Les objets de données s'affichent dans le volet **Objets de données** .
 6. Cliquez sur **Suivant**.
L'onglet **Sélectionner les ressources** s'affiche.
 7. Sous l'onglet **Sélectionner les ressources** , cliquez sur **Choisir** pour ouvrir l'onglet **Sélectionner les ressources** .
Vous pouvez importer des données provenant de plusieurs sources de données relationnelles.
 8. Sous l'onglet **Sélectionner les ressources** , sélectionnez les connexions, schémas, tables et vues que vous souhaitez inclure dans le profil. Cliquez sur **Enregistrer**.
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.
Les ressources s'affichent dans le volet **Ressource** .
 9. Cliquez sur **Suivant**.
L'onglet **Spécifier les paramètres** s'affiche.
 10. Sous l'onglet **Spécifier les paramètres** , vous pouvez configurer les options de profil de colonne et les options de découverte de domaines de données. Cliquez sur **Enregistrer et terminer** pour enregistrer le profil de découverte des données d'entreprise, ou sur **Enregistrer et exécuter** pour exécuter le profil.
Vous pouvez effectuer les tâches suivantes sous l'onglet **Spécifier les paramètres** .
 - Activez la découverte de domaines de données. Cliquez sur **Choisir** pour sélectionner les domaines de données que vous souhaitez découvrir dans la boîte de dialogue **Choisir les domaines de données** . Les domaines de données sélectionnés s'affichent dans le volet **Domaines de données de la découverte de domaines de données** .
 - Exécutez le domaine de données sur les données, sur le nom de colonne ou sur les données et le nom de colonne.
 - Sélectionnez toutes les lignes de la source de données ou choisissez un nombre maximal de lignes sur lesquelles exécuter une découverte de domaines.
 - Spécifiez un pourcentage de conformité minimal ou le nombre minimal de lignes conformes pour la découverte de domaines de données.
 - Activez les paramètres de profil de colonne et sélectionnez toutes les lignes ou les premières lignes de la source de données pour le profil de colonne. Vous pouvez exclure l'inférence de type de données pour les colonnes contenant des types de données approuvés dans le profil de colonne.
 - Choisissez l'option **d'environnement d'exécution natif**.
Remarque: Vous ne pouvez pas exécuter de profils sur le moteur Hive ou Blaze dans l'environnement d'exécution Hadoop, car l'environnement Hadoop n'est pas pris en charge dans la version actuelle d'Informatica.
- Vous pouvez afficher les résultats de la découverte des données d'entreprise sous les onglets **Résumé** et **Profils** .

Modification des options de découverte de données d'entreprise

Après avoir effectué une découverte de 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 de données d'entreprise.
Les résultats de profil s'affichent dans l'espace de travail **Découverte**.
2. Si le système de contrôle de version est activé, cliquez sur **Actions** > **Extraire** pour extraire le profil.
3. Cliquez sur **Éditer le profil**.
4. Dans l'onglet **Spécifiez les propriétés générales**, mettez à jour les propriétés de profil si nécessaire.
5. Pour modifier la sélection des objets de données, cliquez sur l'onglet **Sélectionner des objets de données**.
6. Pour modifier les sources de données externes de la découverte de données d'entreprise, cliquez sur l'onglet **Sélectionner des ressources**.
7. Pour modifier les options d'inférence du domaine de données et les paramètres de profil de colonne, cliquez sur l'onglet **Spécifier les paramètres**.
8. 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 de 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 de données d'entreprise.
9. Cliquez sur **Annuler** pour annuler les modifications.
10. Cliquez sur **Enregistrer et exécuter** pour enregistrer les modifications et réexécuter le profil.
11. Si le système de contrôle de version est activé, vous devez exécuter les tâches suivantes :
 - Cliquez sur **Enregistrer et terminer** pour terminer la modification du profil.
 - Dans la vue de résumé, cliquez sur **Archiver** pour archiver le profil.
 - Cliquez sur **Actions** > **Exécuter le profil** pour exécuter le profil.

CHAPITRE 14

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 Informatica Analyst, 127](#)
- [Vue Résumé, 127](#)
- [Conflit de type de données, 130](#)
- [Vue Profils, 131](#)

Présentation des résultats de la découverte des données d'entreprise dans Informatica 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 les documentés.

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 correspondante.
Nom de colonne	Nom de la colonne de correspondance.
% de conformité des données	Pourcentage minimal de conformité des lignes requis pour une correspondance de domaine de données.
Nom de la connexion	Nom de la connexion de la base de données relationnelle.
Nom de la source	Nom de la source de données.
Statut de l'inférence	Statut de l'inférence du domaine de données. Les statuts sont Accepté , Rejeté et Inféré .
% Null	Pourcentage de valeurs Null de la colonne.
Lignes totales	Nombre total de lignes.
Lignes conformes	Nombre minimal de lignes requis pour une correspondance de domaine de données.
Correspondance de nom de colonne	Détermine si le nom de colonne correspond à un nom de domaine de données.
Type de données documenté	Type de données déclaré pour la colonne de l'objet de profil.
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.

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% Null 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 correspondante.
Connexion	Nom de la 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 des domaines de données

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 de 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 retourner à la vue **Résumé**, cliquez sur **Retourner à la découverte de données d'entreprise**.

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% Null**, **100% de valeurs uniques** 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 retourner à la vue **Résumé**, cliquez sur **Retourner à la découverte des données d'entreprise**.

Conflit de type de données

La découverte des données d'entreprise identifie les conflits de type de données dans des colonnes. Un conflit de type de données est une différence entre les types de données inférés et consignés dans une colonne après l'exécution de la découverte des données d'entreprise. Un type de données inféré est le type de données provenant de l'outil Analyst tool 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 Chaîne 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.

Afficher les conflits de type de données

Lorsque vous ouvrez un profil présentant des conflits de type de données dans la vue **Résumé**, l'outil Analyst tool met en surbrillance les conflits 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 répertoriant les colonnes contenant des conflits de type 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 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 contenant les types de données en conflit, puis cliquez sur **Types de données**.
7. Cliquez sur **Actions**, puis sélectionnez **Approuver** ou **Rejeter**.

8. Pour retourner à la vue **Résumé**, cliquez sur **Retourner à la découverte des données d'entreprise**.

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 15

Recherche de découverte dans Informatica Analyst

Ce chapitre comprend les rubriques suivantes :

- [Présentation de la recherche de découverte dans Informatica Analyst, 132](#)
- [Prérequis de la recherche de découverte, 133](#)
- [Processus de recherche de découverte dans Informatica Analyst, 134](#)
- [Options de recherche de découverte, 134](#)
- [Résultats de la recherche de découverte dans Informatica Analyst, 136](#)
- [Types de correspondances, 138](#)
- [Ressources liées, 139](#)
- [Forum Aux Questions \(FAQ\), 141](#)

Présentation de la recherche de découverte dans Informatica Analyst

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. Les utilisateurs d'entreprise peuvent recourir à la recherche de découverte pour savoir où se trouvent les données et 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 la recherche de découverte, outre les correspondances de texte, les résultats de recherche incluent des objets ayant des relations avec les objets qui correspondent 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 rechercher les objets qui contiennent un modèle de données spécifique et ont des noms comportant un mot-clé particulier.

La recherche de découverte comprend les types d'informations suivants dans les résultats :

Objets dans le référentiel modèle

Recherche les objets primaires associés aux objets qui correspondent aux critères de recherche de découverte. Par exemple, lorsque vous recherchez un profil, les résultats comprennent l'objet de données du profil.

Résultats de l'entrepôt de profil

Inclut les résultats de l'inférence provenant des profils, par exemple un domaine de données ou un modèle de données.

Termes du 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 dans l'entreprise chargé de s'assurer que les données d'entreprise sensibles sont 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 créé des domaines de données pour identifier des données importantes qui n'ont pas été découvertes dans les sources de données. Vous effectuez la recherche sur la chaîne « SSN ». L'outil Analyst affiche le domaine de données de sécurité sociale et toutes les colonnes correspondantes à partir des sources de données. En outre, la chaîne de recherche de découverte peut trouver des colonnes ou des tables supplémentaires qui ont « SSN » dans leur description ou leur nom. Pour réduire votre recherche, vous pouvez filtrer les spécifications de mappage pour afficher celles qui font référence aux objets de données correspondants. Vous pouvez appliquer des filtres supplémentaires pour filtrer les spécifications de mappage en fonction de projets ou d'utilisateurs. Vous pouvez ensuite ouvrir les spécifications de mappage dans les résultats pour vérifier qu'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 tels que 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 des modèles, des 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 textuelle à rechercher.
Effacer	Efface la chaîne de recherche et tous les autres critères de recherche sélectionnés précédemment.
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 que vous voulez inclure dans la recherche.
Valeurs Null	Pourcentage de valeurs Null dans les colonnes que vous voulez 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 comprennent 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.

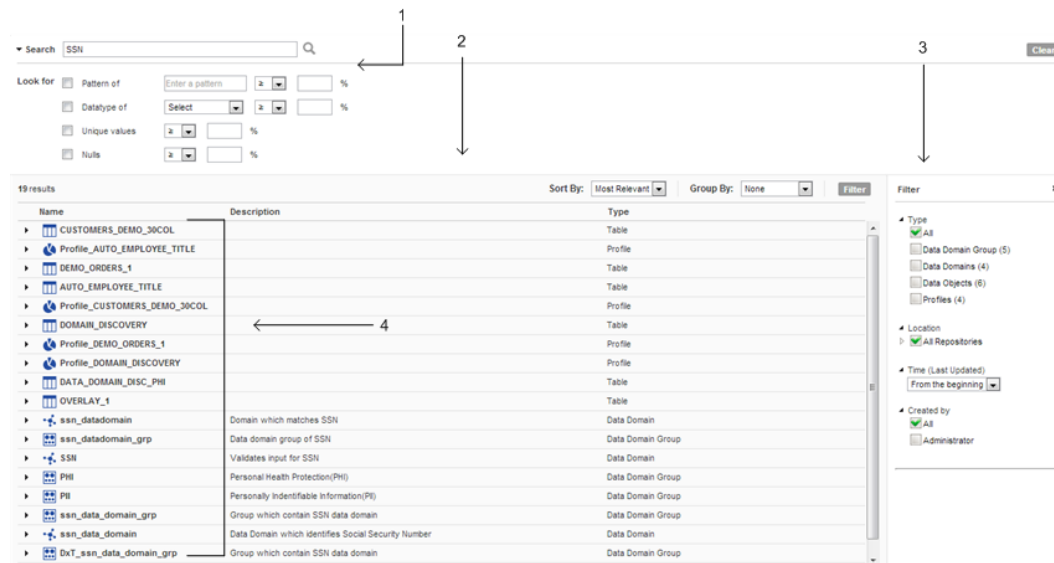
Pour afficher les résultats de la recherche, vous devez disposer des autorisations appropriées sur les projets qui contiennent les correspondances directes et indirectes.

Panneau des résultats de la recherche de découverte

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

Les résultats de la recherche s'affichent dans la grille de résultats de l'espace de travail **Bibliothèque**. Vous pouvez trier les résultats en fonction de leur pertinence. Vous pouvez également grouper les résultats selon le type de ressource, l'emplacement du référentiel, l'heure et l'utilisateur ayant créé les ressources.

L'image suivante montre 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 des résultats de la recherche de découverte affiche les sections suivantes par défaut :

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 en haut de l'outil Analyst.

Grille de résultats

Affiche le nombre total de correspondances et la liste des correspondances en fonction des critères que vous sélectionnez dans les champs de critères de recherche. La grille de résultats contient également la description de l'objet, son type et des menus déroulants permettant de trier et de grouper les résultats de la recherche.

Filtre

Affiche les filtres que vous pouvez définir pour filtrer les résultats de la recherche. La section **Filtre** s'affiche dans la zone de droite de l'outil Analyst.

Résultats de la recherche

Affiche les résultats de la recherche correspondants en fonction de la chaîne de recherche, notamment les objets qui répondent aux critères de recherche. La section des résultats de la recherche contient les propriétés de correspondance, la section **Correspondances directes** et la section **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 16

Bureau de glossaire métier dans Informatica Analyst

Ce chapitre comprend les rubriques suivantes :

- [Termes métier, 142](#)
- [Gestion des termes métier dans le glossaire métier du gestionnaire de métadonnées, 143](#)
- [Recherche d'un terme métier dans le Bureau de glossaire métier, 143](#)

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, 145](#)
- [Profils d'objet de données, 149](#)
- [Profils de colonne dans des sources de données semi-structurées, 161](#)
- [Résultats de profil de colonne dans Informatica Developer, 168](#)
- [Règles dans Informatica Developer, 174](#)
- [Fiches d'évaluation dans Informatica Developer, 176](#)
- [Profilage de mapplet et de mappage, 179](#)
- [Découverte des domaines de données dans Informatica Developer, 182](#)
- [Découverte des données d'entreprise dans Informatica Developer, 195](#)
- [Résultats de la découverte des données d'entreprise, 212](#)
- [Bureau de glossaire métier dans Informatica Developer, 223](#)

CHAPITRE 17

Profils de Informatica Developer

Ce chapitre comprend les rubriques suivantes :

- [Présentation des profils d'Informatica Developer, 145](#)
- [Vues des profils d'Informatica DeveloperVues des profils, 147](#)
- [Verrous d'objets du référentiel et développement basé sur l'équipe contenant des objets avec version, 148](#)

Présentation des profils d'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 colonnes
- Profilage de colonnes sur des sources de données semi-structurées
- Découverte de clé primaire
- 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 de 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 primaire, un profil de dépendance fonctionnelle, et pour la découverte de 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. Vous pouvez créer un profil de colonne pour un objet de données de fichier plat, un objet de données relationnelles et des objets de données semi-structurées.

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 de 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 de 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 de 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 de 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 de 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 :

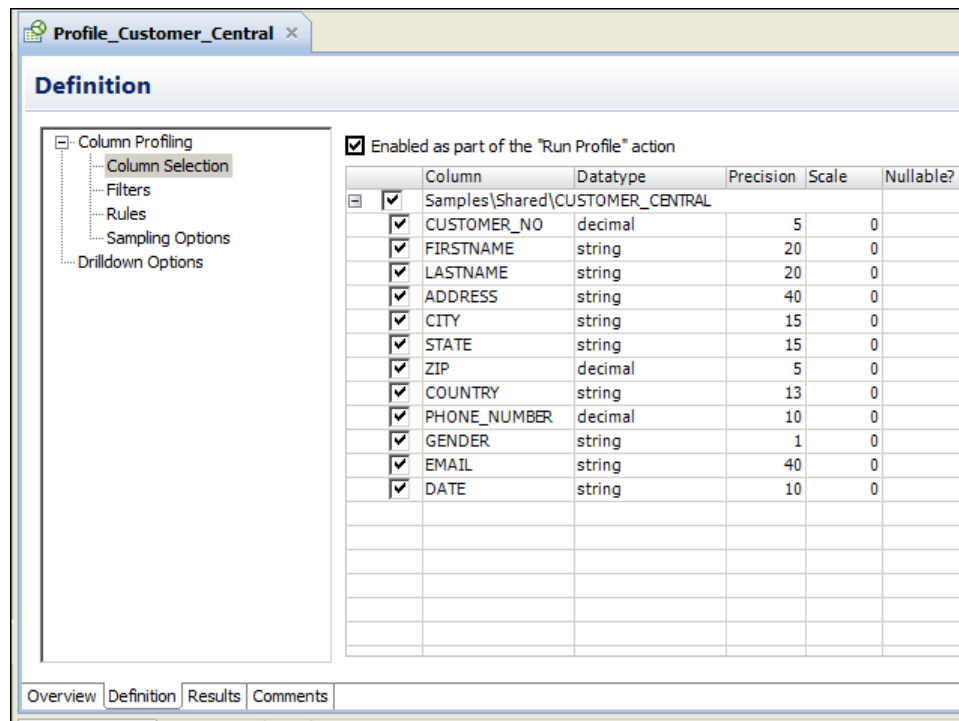
Option de profil	Opérations de profil
Profil	<ul style="list-style-type: none">- Exécuter un profil de colonne.- Rechercher des clés principales.- Rechercher des dépendances fonctionnelles.- Identifier des domaines de données
Profils multiples	Créer et exécuter des profils de colonne sur plusieurs objets en même temps.
Profil de découverte de 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 de 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.

Verrous d'objets du référentiel et développement basé sur l'équipe contenant des objets avec version

Le référentiel modèle verrouille les profils pour empêcher les utilisateurs d'écraser le travail des autres développeurs. Si le référentiel modèle est intégré à un système de contrôle de version, il enregistre plusieurs versions des ressources et attribue un numéro à chaque version. Vous pouvez extraire et archiver des profils, annuler des extractions et afficher les profils que vous avez extraits.

Le référentiel modèle conserve les verrous d'objet en cas d'arrêt inopiné de l'outil Developer tool. Lorsque vous vous reconnectez au référentiel modèle, vous pouvez voir les objets que vous avez verrouillés et continuer à les modifier ou à les déverrouiller. Vous pouvez afficher et déverrouiller des objets verrouillés dans la boîte de dialogue **Objets verrouillés**. Pour afficher la boîte de dialogue **Objets verrouillés**, cliquez sur **Afficher > Objets verrouillés**.

Si le référentiel modèle est intégré à un système de contrôle de version, vous pouvez gérer les versions d'objet dans l'outil Developer tool grâce à la gestion d'objet avec version. Vous pouvez effectuer des actions telles que l'archivage et l'extraction d'objets, l'affichage et la récupération de versions d'historiques des objets ou encore l'annulation d'une extraction.

Le référentiel modèle empêche d'autres membres de l'équipe de développement de remplacer les objets. Si vous ouvrez un objet qu'un autre utilisateur a extrait, vous recevez une notification qui identifie l'utilisateur à l'origine de l'extraction. Vous pouvez ouvrir un objet extrait en lecture seule ou l'enregistrer sous un autre nom.

Le référentiel modèle crée des nouvelles versions d'un objet lorsque celui-ci est modifié.

Le référentiel modèle incrémente le numéro de version une fois que vous avez conservé le type de données, la clé primaire, la clé étrangère ou le domaine de données.

Lorsque vous restaurez une version, les derniers résultats de profil s'affichent dans la vue Résultats plutôt que les résultats de profil de la version restaurée. Ceci est dû au fait que le système de contrôle de version conserve les définitions de profil de toutes les versions dans le référentiel modèle, et parce que les résultats de profil sont extraits de l'entrepôt de profilage. Pour obtenir plus d'informations sur les verrous d'objets du référentiel et sur la gestion d'objet avec version, consultez le *Guide de l'outil Developer tool*.

CHAPITRE 18

Profils d'objet de données

Ce chapitre comprend les rubriques suivantes :

- [Présentation de profils d'objet de données, 149](#)
- [Profils de colonne dans Informatica Developer, 150](#)
- [Environnement d'exécution, 153](#)
- [Découverte de clé primaire, 153](#)
- [Découverte de dépendance fonctionnelle, 155](#)
- [Profils du système d'exploitation dans Informatica Developer, 156](#)
- [Création d'un profil d'objet de données unique dans Informatica Developer, 157](#)
- [Création de plusieurs profils d'objet de données dans Informatica Developer, 158](#)
- [Modification d'un profil, 158](#)
- [Option de synchronisation, 159](#)
- [Commentaires, 160](#)

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 sur un seul objet de données ou sur plusieurs objets de données dans Informatica 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 colonnes	Découvre les caractéristiques de données, telles que les fréquences, les pourcentages et les modèles. 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 primaire	Découvre les colonnes avec des valeurs qui peuvent identifier de façon unique les lignes dans une source de données.

Tâche	Description
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 primaire 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 primaires 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 une source 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'occurrence de chaque valeur dans une colonne, exprimée en pourcentage ou en nombre de lignes.
- 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, les critères de conformité pour la découverte des domaines de données 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 définir des options d'échantillonnage, des options d'exploration et l'environnement d'exécution pour un profil de colonne. Vous pouvez ajouter des règles et des filtres pour un profil de colonne.

Options de filtrage

Vous pouvez ajouter des filtres avancés ou des filtres SQL pour déterminer les lignes qu'un profil de colonne utilise lorsque vous exécutez le profil. Le profil ne traite pas les lignes qui ne répondent pas aux critères de filtrage.

Création d'un filtre avancé

Vous pouvez créer un filtre avancé avec des expressions comme AND, OR et NOT pour créer un sous-ensemble de la source de données d'origine.

1. Créez ou ouvrez un profil d'objet de données unique.
2. Sélectionnez la vue **Filtre**.
3. Cliquez sur **Ajouter**.
La boîte de dialogue **Sélectionner un assistant** s'affiche.
4. Dans la boîte de dialogue **Sélectionner un assistant**, cliquez sur **Filtre avancé**.
La boîte de dialogue **Filtre** s'affiche.
5. Entrez le nom, et éventuellement la description, du filtre avancé.
6. Sélectionnez **Définir comme actif** pour appliquer le filtre au profil. Cliquez sur **Suivant**.
7. Sélectionnez **Définition de filtre** pour définir un filtre.
8. Vous pouvez créer un filtre avancé à l'aide du panneau **Fonctions** ou **Colonnes**.
 - Dans le panneau **Fonctions**, sélectionnez une catégorie de fonction, puis cliquez sur les flèches vers la droite (>>).
Dans la boîte de dialogue, spécifiez les paramètres et cliquez sur **OK**. La fonction s'affiche avec les colonnes et les valeurs dans le panneau **Expression**.
 - Dans le panneau **Colonnes**, sélectionnez une colonne, puis cliquez sur les flèches vers la droite (>>).
La colonne s'affiche dans le panneau **Expression**.
Ajoutez des fonctions, des expressions et des valeurs pour créer un filtre avancé.
9. Pour vérifier le filtre avancé, cliquez sur **Valider**.
10. Après avoir créé ou modifié le filtre, sélectionnez **Prévisualisation des données** pour afficher les données filtrées. Vous pouvez définir l'option **Nombre maximal de lignes à prévisualiser**.
11. Cliquez sur **Terminer**.
L'assistant **Nouveau profil** s'affiche avec le filtre dans la vue **Filtres**.

Création d'un filtre SQL

Vous pouvez créer un filtre SQL avec des requêtes SQL. Vous pouvez créer un filtre SQL pour les sources de données relationnelles.

1. Créez ou ouvrez un profil d'objet de données unique.
2. Sélectionnez la vue **Filtre**.
3. Cliquez sur **Ajouter**.
La boîte de dialogue **Sélectionner un assistant** s'affiche.
4. Dans la boîte de dialogue **Sélectionner un assistant**, cliquez sur **Filtre SQL**.

La boîte de dialogue **Filtre** s'affiche.

5. Entrez le nom, et éventuellement la description, du filtre avancé.
6. Sélectionnez **Définir comme actif** pour appliquer le filtre au profil. Cliquez sur **Suivant**.
7. Sélectionnez **Définition de filtre** pour définir un filtre.
8. Utilisez les colonnes du panneau **Colonnes** pour créer un filtre SQL.
9. Pour vérifier le filtre, cliquez sur **Valider**.
10. Après avoir créé ou modifié le filtre, sélectionnez **Prévisualisation des données** pour afficher les données filtrées. Vous pouvez définir l'option **Nombre maximal de lignes à prévisualiser**.
11. Cliquez sur **Terminer**.

L'assistant **Nouveau profil** s'affiche avec le filtre dans la vue **Filtres**.

Options d'échantillonnage

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

Le tableau suivant décrit les options d'échantillonnage :

Propriété	Description
Toutes les lignes	Choisit toutes les lignes de l'objet de données.
Premier	Nombre de lignes pour lesquelles vous souhaitez exécuter le profil. L'outil Developer tool choisit les lignes parmi les premières lignes de la source.
Échantillon aléatoire de	Nombre de lignes d'un échantillon aléatoire pour lequel vous souhaitez exécuter le profil.
Échantillon aléatoire (Auto)	Taille de l'échantillon aléatoire basée sur le nombre de lignes dans l'objet de données.
Excluez les types et les domaines de données approuvés de l'inférence du type et du domaine de données dans les exécutions de profil suivantes.	Exclut le type de données approuvé ou le domaine de données du type de données et l'inférence de domaine de données de l'exécution de profil suivante.

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.

Environnement d'exécution

Choisissez l'environnement d'exécution natif pour un profil de colonne. Informatica Developer définit l'environnement d'exécution dans la définition de profil une fois que vous avez choisi un environnement d'exécution.

Lorsque vous exécutez un profil dans un environnement d'exécution natif, l'outil Developer tool soumet les tâches de profil au module de service de profilage. Celui-ci divise alors les tâches de profil en un ensemble de mappages. Le service d'intégration de données exécute ces mappages sur la même machine où il exécute et écrit les résultats de profil dans l'entrepôt de profilage. Par défaut, tous les profils s'exécutent dans un environnement d'exécution natif.

Vous pouvez utiliser des sources natives pour créer et exécuter des profils dans l'environnement natif. Une source de données native est une source non-Hadoop, comme un fichier plat, une source relationnelle ou une source principale. Vous pouvez également exécuter un profil sur une spécification de mappage ou une source de données logique avec une source de données Hive ou HDFS dans l'environnement natif.

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.

Propriété	Description
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.
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.

Profils du système d'exploitation dans Informatica Developer

Vous pouvez choisir un profil de système d'exploitation dans l'outil Developer tool. Après avoir choisi un profil de système d'exploitation, le service d'intégration de données crée et exécute les profils de colonne, les profils de découverte de données d'entreprise et les fiches d'évaluation en fonction des autorisations de l'utilisateur du profil de système d'exploitation.

Sélection d'un profil de système d'exploitation

Vous pouvez sélectionner un profil de système d'exploitation dans Informatica Developer. Le service d'intégration de données utilise les autorisations de l'utilisateur du profil de système d'exploitation pour exécuter les tâches de profilage.

1. Dans Informatica Developer, cliquez sur **Windows > Préférences**.
La boîte de dialogue **Préférences** s'affiche.
2. Cliquez sur **Informatica > Exécuter les configurations > Mappage**.
La boîte de dialogue **Mappage** s'affiche.
3. Dans la boîte de dialogue **Mappage**, effacez l'option **Utiliser le service d'intégration de données par défaut**.
4. Cliquez sur **Parcourir** pour sélectionner un profil de système d'exploitation dans la liste.
5. Cliquez sur **OK**.

Création d'un profil d'objet de données unique dans Informatica Developer

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.

1. Dans la vue **Explorateur d'objets**, sélectionnez l'objet de données à profiler.
2. Cliquez sur **Fichier > Nouveau > profil** pour ouvrir l'assistant 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 colonnes
 - Découverte de clé primaire
 - Découverte de dépendance fonctionnelle
 - Découverte de domaines de données

Remarque: Pour activer une opération de profil, sélectionnez **Activé dans le cadre de l'action "Exécuter le profil"** pour cette opération. Le profilage de colonnes est activé par défaut.
9. Vérifiez les options pour votre profil.

Vous pouvez modifier la sélection de colonne pour tous les types de profil. Vérifiez les options de filtre et d'échantillonnage pour les profils de colonne. Vous pouvez vérifier les options d'inférence pour clé primaire, 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.

10. Vérifiez les options d'exploration et modifiez-les si nécessaire. Par défaut, l'option **Activer l'exploration des lignes** est sélectionnée. 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.
11. Dans la section **Paramètres d'exécution**, choisissez l'option d'environnement d'exécution **Natif**.
Remarque: Vous ne pouvez pas exécuter de profils sur le moteur Hive ou Blaze dans l'environnement d'exécution Hadoop, car l'environnement Hadoop n'est pas pris en charge dans la version actuelle d'Informatica.
12. Cliquez sur **Terminer**.

Création de plusieurs profils d'objet de données dans Informatica Developer

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**.

Modification d'un profil

Vous pouvez modifier un ou plusieurs profils d'objet de données. Si le système de contrôle de version est activé, le profil est extrait par défaut.

1. Dans la vue **Explorateur d'objets**, cliquez avec le bouton droit de la souris sur le profil, puis cliquez sur **Ouvrir**.
La vue **Résultats** s'affiche.
2. Dans la vue **Définition**, mettez à jour les propriétés si nécessaire.

3. Cliquez sur **Équipe** > **Archiver** pour archiver le profil.
 4. Cliquez avec le bouton droit de la souris sur le profil, puis cliquez sur **Exécuter le profil** pour exécuter le profil.
- Les résultats de profil s'affichent dans la vue **Résultats**.

Option de synchronisation

Lorsque vous modifiez les métadonnées d'une source de données externe, les métadonnées de l'objet de données dans le référentiel de modèles ne sont pas mises à jour par défaut. Utilisez l'option de synchronisation pour synchroniser les métadonnées de l'objet de données avec les métadonnées de la source de données.

Dans l'outil Developer tool, après avoir utilisé l'option de synchronisation, lorsque vous ouvrez un profil ou une fiche d'évaluation qui utilise l'objet de données, un astérisque apparaît en regard du nom du profil ou de la fiche d'évaluation dans l'éditeur. L'astérisque indique que les métadonnées de l'objet de données ont été modifiées pour le profil ou la fiche d'évaluation. Ouvrez et enregistrez le profil ou la fiche d'évaluation pour mettre à jour la définition du profil dans le référentiel de modèles. Notez qu'aucune modification visible n'apparaît dans l'outil Analyst tool lorsque vous ouvrez le profil ou la fiche d'évaluation après avoir synchronisé l'objet de données de celui-ci ou celle-ci. Vous pouvez utiliser l'option de synchronisation pour les profils de colonnes, les profils de découverte d'entreprise et les fiches d'évaluation. La source de données externe peut être une source de données relationnelle ou une source de données de fichier plat.

Synchronisation d'un objet de données de 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.

Commentaires

Vous pouvez ajouter une description en tant que commentaire à un profil. Vous pouvez également ajouter des commentaires aux colonnes des résultats de profil de colonne.

Vous pouvez ajouter plusieurs commentaires aux colonnes. La vue **Commentaires** de l'outil Developer tool permet d'ajouter et d'afficher des commentaires.

Ajout de commentaires dans Informatica Developer

Ajoutez des commentaires à des colonnes dans les résultats de profil de colonne. L'outil Developer tool inclut les commentaires lors de l'exportation des résultats de profil.

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 **Commentaires**.
4. Cliquez sur **Ajouter** pour ouvrir la boîte de dialogue **Ajouter des commentaires**.
5. Sélectionnez le nom du profil ou une colonne dans la liste.
Si vous avez déjà ajouté des commentaires, ils s'affichent dans la boîte de dialogue.
6. Entrez une description dans le champ **Commentaires**.
7. Cliquez sur **OK**.
L'outil Developer tool affiche les commentaires dans la vue **Commentaires**.
8. Pour supprimer un commentaire, sélectionnez-le dans la vue **Commentaires** et cliquez sur **Supprimer**.

CHAPITRE 19

Profils de colonne dans des sources de données semi-structurées

Ce chapitre comprend les rubriques suivantes :

- [Présentation de profils de colonne dans des sources de données semi-structurées, 161](#)
- [Profil de colonne dans une source de données XML ou JSON, 162](#)
- [Profil de colonne dans une source de données semi-structurée dans HDFS, 163](#)
- [Création d'une connexion HDFS, 164](#)
- [Création d'un objet de données de fichier complexe sur un fichier JSON ou XML dans HDFS, 164](#)
- [Création d'un objet de données sur une source de données Avro ou Parquet, 165](#)
- [Exécution d'un profil de colonne sur des sources de données semi-structurées, 166](#)

Présentation de profils de colonne dans des sources de données semi-structurées

Vous pouvez créer des objets de données à partir de sources de données Avro, JSON, Parquet et XML, et créer un profil de colonne sur les objets de données.

Les formats Avro, JSON, Parquet et XML sont des sources de données semi-structurées. Pour utiliser les sources de données semi-structurées afin de créer un profil de colonne, procédez comme suit :

1. Créez un objet de données physique dans la source de données semi-structurée.
2. Créez et exécutez un profil de colonne sur l'objet de données physique.

Vous pouvez créer des objets de données de fichier plat pour les sources de données XML ou JSON. Vous pouvez créer des objets de données de fichier complexe pour des sources de données Avro, JSON, Parquet et XML dans le système HDFS (Hadoop Distributed file System).

Profil de colonne dans une source de données XML ou JSON

Vous pouvez créer un objet de données de fichier plat ou complexe à partir de sources de données JSON ou XML. Vous pouvez créer et exécuter un profil de colonne dans l'objet de données.

Créez un fichier texte qui contient le chemin de la source de données XML ou JSON et utilisez-le comme source de données pour créer un objet de données de fichier plat. Vous pouvez également ajouter le chemin du fichier pour plusieurs sources de données JSON ou XML au fichier texte.

Vous pouvez créer un objet de données de fichier complexe à partir d'une source de données XML ou JSON avec un lecteur de fichier complexe. Le lecteur de fichier complexe fournit une entrée à une transformation Processeur de données qui analyse le fichier et convertit les données source en enregistrements de valeurs linéaires séparés par des virgules.

Remarque: L'outil Developer tool ne prend pas en charge les sources de données JSON avec le codage UTF-8.

Création d'un objet de données sur une source de données XML ou JSON

Vous pouvez créer un objet de données de fichier plat ou complexe sur une source de données JSON ou XML.

1. Dans la vue **Explorateur d'objets** de l'outil Developer tool, sélectionnez le projet dans lequel vous souhaitez créer l'objet de données et le profil de colonne.

2. Cliquez sur **Fichier > Nouveau > Objet de données**.

La boîte de dialogue **Nouveau** s'affiche.

3. Vous pouvez créer un objet de données de fichier plat ou de fichier complexe.

- Pour créer un objet de données de fichier plat, effectuez les tâches suivantes :

1. Sélectionnez **Objets de données physiques > Objet de données de fichier plat**, puis cliquez sur **Suivant**.

La boîte de dialogue **Nouvel objet de données de fichier plat** s'ouvre.

2. Sélectionnez **Créer à partir d'un fichier plat existant**, puis cliquez sur **Parcourir** pour choisir le fichier texte. Cliquez sur **Suivant**.

3. Vérifiez que la page de code est **MS Windows Latin 1 (ANSI), sur-ensemble de Latin 1** et que le format est délimité. Cliquez sur **Suivant**.

4. Vérifiez que le délimiteur est défini sur la **virgule**. Cliquez sur **Terminer**.

- Pour créer un objet de données de fichier complexe, effectuez les tâches suivantes :

1. Sélectionnez **Objets de données physiques > Objet de données de fichier complexe** et cliquez sur **Suivant**.

La boîte de dialogue **Nouvel objet de données de fichier complexe** s'affiche.

2. Entrez le nom de l'objet de données. Sélectionnez le type d'accès **Fichier**.

3. Cliquez sur **Parcourir** pour choisir un fichier JSON ou XML. Cliquez sur **Terminer**.

Lorsque le serveur de Developer est exécuté sous Linux, vous devez mettre à jour le chemin d'accès au fichier de la source de données sur l'emplacement de ce serveur. Pour mettre à jour le chemin d'accès au fichier, sélectionnez l'objet de données de fichier complexe, cliquez sur **Lire** dans l'onglet **Opérations de l'objet de données** et ajoutez le chemin d'accès au fichier dans l'onglet **Avancé** du volet **Détails de l'opération de l'objet de données**.

L'objet de données s'affiche dans le dossier du projet.

Profil de colonne dans une source de données semi-structurée dans HDFS

Vous pouvez créer et exécuter un profil de colonne sur un fichier Avro, JSON, Parquet ou XML qui utilise HDFS. Pour lire le fichier XML ou JSON dans HDFS, utilisez un lecteur de fichier complexe pour transmettre l'entrée XML ou JSON la transformation Processeur de données.

Profil de colonne sur une source de données XML ou JSON dans HDFS

Vous pouvez créer un objet de données de fichier complexe à partir d'un fichier XML ou JSON. Vous pouvez créer et exécuter un profil de colonne sur un objet de données.

Créer une connexion à HDFS avant de créer les objets de données pour les fichiers JSON ou XML dans HDFS.

Vous pouvez utiliser l'une des méthodes suivantes pour créer un objet de données à partir d'un fichier XML ou JSON dans HDFS :

- Créez un objet de données de fichier complexe sur un fichier XML ou JSON.
- Créez un objet de données de fichier complexe sur un dossier qui contient plusieurs fichiers JSON ou XML.

Après avoir créé l'objet de données, vous pouvez créer et exécuter un profil de colonne sur celui-ci.

Profil de colonne d'une source de données Avro ou Parquet dans HDFS

Vous pouvez créer un objet de données de fichier complexe à partir d'une source de données Avro ou Parquet dans HDFS. Vous pouvez utiliser l'objet de données pour créer et exécuter un profil de colonne.

Vous pouvez créer un objet de données de fichier complexe pour un fichier Avro ou Parquet ou pour un dossier qui contient plusieurs fichiers Avro ou Parquet. Vous pouvez créer un objet de données de fichier complexe pour une source de données Avro et Parquet avec un type d'accès Fichier ou Connexion et un format de ressource Binaire, Avro ou Parquet. Vous devez créer une connexion HDFS avant de créer un objet de données de fichier complexe sur les sources de données Avro et Parquet.

Remarque: Vous pouvez uniquement choisir le format de ressource **Avro** ou **Parquet** pour les sources de données Avro et Parquet linéaires.

Vous pouvez choisir l'une des options suivantes lorsque vous créez un objet de données à partir de fichiers Avro et Parquet dans HDFS :

- Sélectionnez le type d'accès Fichier et le format de ressource Binaire.
- Sélectionnez le type d'accès Fichier et le format de la ressource Avro ou Parquet.
- Sélectionnez le type d'accès Connexion et le format de la ressource Avro ou Parquet.

Création d'une connexion HDFS

Configurez la connexion HDFS dans Informatica Developer pour créer un profil de colonne sur des sources de données Avro, JSON, Parquet et XML dans HDFS. Vous pouvez créer un objet de données de fichier complexe après avoir créé une connexion HDFS.

1. Cliquez sur **Fenêtre > Préférences**.
2. Sélectionnez **Informatica > Connexions**.
3. Développez le domaine.
4. Sélectionnez le type de connexion **Systèmes de fichiers > Système de fichiers Hadoop**, puis cliquez sur **Ajouter**.
5. Spécifiez un nom pour la connexion.
6. Entrez éventuellement une description de la connexion.
7. Cliquez sur **Suivant**.
8. Configurez les propriétés de la connexion.
9. Cliquez sur **Tester la connexion** pour vérifier la connexion à HDFS.
10. Cliquez sur **Terminer**.

Création d'un objet de données de fichier complexe sur un fichier JSON ou XML dans HDFS

Vous pouvez créer un objet de données physiques complexe pour le fichier source JSON ou XML qui utilise HDFS, puis un profil de colonne sur l'objet physique.

1. Dans la vue **Explorateur d'objets**, sélectionnez le projet dans lequel vous souhaitez créer un objet de données physiques et un profil de colonne.
2. Cliquez sur **Fichier > Nouveau > Objet de données**.
La boîte de dialogue **Nouveau** s'affiche.
3. Sélectionnez **Objets de données physiques > Objet de données de fichier complexe** et cliquez sur **Suivant**.
La boîte de dialogue **Nouvel objet de données de fichier complexe** s'affiche.
4. Entrez le nom de l'objet de données. Sélectionnez le type d'accès **Connexion**.
5. Vous pouvez créer un objet de données sur un fichier XML ou JSON, ou sur un dossier qui contient plusieurs fichiers JSON ou XML.
 - Pour poursuivre la création d'un objet de données de fichier complexe sur un fichier XML ou JSON, effectuez les tâches suivantes :
 1. Cliquez sur **Parcourir** pour sélectionner une connexion. Dans la boîte de dialogue **Ajouter une ressource**, cliquez sur **Ajouter** pour choisir un fichier JSON ou XML. Cliquez sur **Terminer**.
 2. L'objet de données s'affiche dans le dossier du projet.

- Pour poursuivre la création d'un objet de données de fichier complexe sur un dossier avec plusieurs fichiers JSON ou XML, effectuez les tâches suivantes :
 1. Cliquez sur **Parcourir** pour sélectionner une connexion. Dans la boîte de dialogue **Ajouter une ressource**, cliquez sur **Ajouter** et choisissez un fichier JSON ou XML dans le dossier. Cliquez sur **Terminer**.
 2. L'objet de données s'affiche dans le dossier du projet.
 3. Sélectionnez l'objet de données dans le dossier de projet, puis cliquez sur **Avancé > Exécution : Lire > Répertoire du fichier source**.
 4. Supprimez le nom du fichier et conservez le nom du dossier dans le chemin.

Création d'un objet de données sur une source de données Avro ou Parquet

Vous pouvez créer un objet de données physiques complexe pour une source de données Avro ou Parquet avec un type d'accès **Fichier** ou **Connexion**. Vous pouvez créer un profil de colonne dans l'objet de données physiques.

1. Dans la vue **Explorateur d'objets**, sélectionnez un projet.
2. Cliquez sur **Fichier > Nouveau > Objet de données**.
La boîte de dialogue **Nouveau** s'affiche.
3. Sélectionnez **Objets de données physiques > Objet de données de fichier complexe** et cliquez sur **Suivant**.
La boîte de dialogue **Nouvel objet de données de fichier complexe** s'affiche.
4. Entrez un nom pour l'objet de données.
5. Vous pouvez choisir le type d'accès **Connexion** ou **Fichier**.
 - Si vous choisissez le type d'accès **Connexion**, effectuez les tâches suivantes :
 1. Cliquez sur **Parcourir** pour sélectionner une connexion HDFS.
La boîte de dialogue **Choisir une connexion** s'affiche.
 2. Dans la boîte de dialogue **Choisir une connexion**, sélectionnez une source de données, puis cliquez sur **OK**.
 3. Dans la boîte de dialogue **Nouvel objet de données de fichier complexe**, cliquez sur **Terminer**.
L'objet de données s'affiche dans le dossier du projet.
 - Si vous choisissez le type d'accès **Fichier** et le format de ressource **Binaire**, effectuez les tâches suivantes :
 1. Cliquez sur **Parcourir** pour choisir un fichier Avro ou Parquet sur la machine locale.
 2. Dans la boîte de dialogue **Nouvel objet de données de fichier complexe**, cliquez sur **Terminer**.
L'objet de données s'affiche dans le dossier du projet.
 3. Sélectionnez l'objet de données dans le dossier du projet, puis cliquez sur la vue **Opérations de l'objet de données**.
 4. Dans la vue **Opérations de l'objet de données**, cliquez sur l'onglet **Lire > Avancé**.
 5. Sous l'onglet **Avancé**, entrez le chemin de fichier de la source de données sur la machine Linux ou Windows dans le champ **Chemin du fichier**.

6. Sélectionnez le format de fichier **Entrée personnalisée**.
 7. Entrez **com.informatica.avro.AvroToXML** dans le champ **Format d'entrée** pour la source de données Avro, puis entrez **com.informatica.parquet.ParquetToXML** dans le champ **Format d'entrée** pour les sources de données Parquet. Lorsque vous ajoutez le format d'entrée, la transformation Processeur de données traite et transforme les données sources Avro ou Parquet en une source de données XML lors de l'exécution.
- Si vous choisissez le type d'accès **Fichier** et le format de ressource **Avro** ou **Parquet**, effectuez les tâches suivantes :
 1. Cliquez sur **Parcourir** pour choisir un fichier Avro ou Parquet sur la machine locale.
 2. Dans la boîte de dialogue **Nouvel objet de données de fichier complexe**, cliquez sur **Terminer**. L'objet de données s'affiche dans le dossier du projet.
 3. Après avoir créé l'objet de données, accédez à l'onglet **Opérations de l'objet de données > Lire > Avancé** et vérifiez si le chemin dans le champ **Chemin du fichier** correspond à la source de données sur la machine Linux ou Windows.

Remarque: Vous pouvez uniquement choisir le format de ressource **Avro** ou **Parquet** pour les sources de données Avro et Parquet linéaires.

Vous pouvez choisir un dossier avec plusieurs fichiers Avro ou Parquet pour créer un objet de données. Après avoir créé l'objet de données, accédez à l'onglet **Opérations de l'objet de données > Lire > Avancé** et vérifiez si le chemin dans le champ **Chemin du fichier** correspond au dossier des sources de données sur la machine Linux ou Windows.

Exécution d'un profil de colonne sur des sources de données semi-structurées

Après avoir créé un objet de données de fichier plat ou un objet de données de fichier complexe à partir de sources de données Avro, JSON Parquet ou XML, vous pouvez créer un profil de colonne dans l'objet de données et l'exécuter.

1. Dans la vue **Explorateur d'objets**, sélectionnez l'objet de données physiques du fichier Avro, JSON, Parquet ou XML.
2. Cliquez sur **Fichier > Nouveau > Profil**.
La boîte de dialogue **Nouveau** s'affiche.
3. Sélectionnez **Profil**. Cliquez sur **Suivant**.
La boîte de dialogue **Nouveau profil** s'affiche.
4. Dans la boîte de dialogue **Nouveau profil**, ajoutez le nom du profil, et éventuellement une description.
5. Sélectionnez l'option **Traiter les formats de fichier étendus**. Cliquez sur **Suivant**.

L'image suivante présente l'assistant **Nouveau profil** avec l'option **Traiter les formats de fichier étendus** :

1. Traiter les formats de fichier étendus Sélectionnez cette option pour traiter les sources de données semi-structurées.

Remarque: L'option **Traiter les formats de fichier étendus** ne s'affiche pas pour les sources de données Avro et Parquet lorsque vous choisissez le format de ressource **Avro** ou **Parquet**.

6. Sur la page **Profil d'objet de données unique**, sélectionnez les colonnes et les options appropriées sous **Sélection des colonnes** et **Découverte de domaines de données**. Cliquez sur **Terminer**.

Remarque: Si l'outil Developer tool est installé sur une machine Linux et l'objet de données physiques XML ou JSON est un objet de données de fichier plat avec un fichier texte, effectuez les tâches suivantes :

1. Sous l'onglet **Présentation**, mettez à jour la valeur **Précision** pour inclure le nombre de caractères du chemin d'accès au fichier de la source de données sur le serveur.
2. Mettre à jour le chemin d'accès au fichier de la source de données sur le serveur après avoir créé un profil sur l'objet de données de fichier plat. Pour mettre à jour le chemin d'accès au fichier, cliquez sur **Exécution : Lire > Répertoire de fichier source** dans l'onglet **Avancé** et ajoutez le chemin d'accès au fichier.
7. Cliquez avec le bouton droit de la souris sur le profil et sélectionnez **Exécuter le profil**.
Les résultats de profil s'affichent.

CHAPITRE 20

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, 168](#)
- [Propriétés des valeurs de colonne, 169](#)
- [Propriétés des formes de colonne, 170](#)
- [Propriétés des statistiques de colonne, 170](#)
- [Propriétés des types de données de colonne, 171](#)
- [Conservation dans Informatica Developer](#)[Conservation dans Informatica Developer, 172](#)
- [Exportation des résultats de profil depuis Informatica Developer, 173](#)

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 déclaré par la source de 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 déclaré par la source de 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

L'image suivante présente 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	Une liste de toutes les valeurs de la colonne dans le profil.
Fréquence	Nombre de fois qu'une valeur s'affiche dans une colonne.
Pourcentage	Nombre de fois qu'une valeur s'affiche dans une colonne, exprimé en pourcentage de toutes les valeurs de la colonne.
Graphique	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	Longueur de la valeur la plus longue de la colonne.
Longueur minimale	Longueur de la valeur la plus courte de la colonne.
Inférieur	Cinq dernières valeurs de la colonne.
Supérieur	Cinq premières valeurs de la colonne.
Somme	Somme de toutes les valeurs de la colonne associées au 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 comprennent tous les types de données inférés pour chaque colonne dans les résultats de profil.

Pour afficher les informations sur le type de données, sélectionnez **Types de données** dans la liste **Afficher**. Double-cliquez sur un type de données pour explorer les lignes contenant 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 pour la colonne du profil.
Fréquence	Nombre d'occurrences d'un type de données dans une colonne.
% de conformité	Pourcentage d'occurrence d'un type de données dans une colonne.
Statut	<p>Indique le statut du type de données. Ce statut peut être Inféré, Approuvé ou Rejeté.</p> <p>Inféré</p> <p>Indique le type de données de la colonne qui a été inféré par l'outil Developer tool.</p> <p>Approuvé</p> <p>Indique un type de données approuvé pour la colonne. Lorsque vous approuvez un type de données, vous le validez dans le référentiel modèle.</p> <p>Rejeté</p> <p>Indique un type de données rejeté pour la colonne.</p>

Conservation dans Informatica 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 organisez des métadonnées dans Informatica Developer, vous pouvez approuver, rejeter et réinitialiser les types de données ou les domaines de données inférés dans les résultats de profil.

Vous pouvez approuver un type de données ou un domaine de données pour une colonne. Vous pouvez masquer les types de données ou les domaines de données rejetés pour une colonne. Après avoir approuvé ou rejeté un type de données ou un domaine de données inféré, vous pouvez réinitialiser le type de données ou le domaine de données pour restaurer le statut inféré.

Approbation des types de données

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 d'inférence de chaque colonne de la source de données. Vous pouvez choisir et approuver un seul type de données pour chaque colonne.

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 pour la colonne s'affichent.
5. Cliquez avec le bouton droit de la souris sur la colonne à approuver, puis cliquez sur **Approuver**.
Le statut du type de données est remplacé par **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 de types de données

Rejet de types de données dans Informatica Developer

Par défaut, Informatica Developer affiche les types de données inférés dans les résultats de profil. Vous pouvez rejeter des types de données inférés ou approuvés et 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. Pour rejeter des types de données de colonne inférés, sélectionnez la vue **Types de données** dans le panneau de droite. Sélectionnez le type de données inféré à rejeter, cliquez avec le bouton droit sur la ligne et choisissez **Rejeter**.
Informatica Developer grise le type de données rejeté dans la liste des types de données.

5. Pour masquer les types de données rejetés, cliquez sur la ligne avec le bouton droit de la souris et sélectionnez **Masquer les éléments rejetés**.
6. Pour afficher les types de données rejetés, cliquez sur une ligne avec le bouton droit de la souris, 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 tool 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**, puis 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. Les valeurs disponibles sont **Valeurs pour la colonne sélectionnée**, **Modèles pour la colonne sélectionnée**, **Types de données pour la colonne sélectionnée** et **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 21

Règles dans Informatica Developer

Ce chapitre comprend les rubriques suivantes :

- [Présentation des règles dans Informatica Developer](#)[Lignes directrices relatives aux règles, 174](#)
- [Création d'une règle dans Informatica Developer, 175](#)
- [Application d'une règle dans Informatica Developer](#)[Application d'une règle, 175](#)

Présentation des règles dans Informatica Developer

Lignes directrices relatives aux 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 tool. Vous pouvez réutiliser ces règles dans les profils de l'outil Analyst tool pour valider les données source. Vous pouvez créer des règles réutilisables à partir de mapplets dans l'outil Developer tool. 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 tool. Vous pouvez appliquer la règle à un profil de colonne dans l'outil Developer tool ou Analyst tool.

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 les transformations Expression, Recherche et de qualité de 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 c'est une transformation active.

- Elle ne spécifie pas de cardinalité entre les groupes d'entrée.

Lorsque vous utilisez un mapplet en tant que règle, il est recommandé de spécifier une valeur par défaut au port de sortie de la transformation pour identifier les erreurs d'exécution. Par exemple, vous pouvez utiliser NULL comme la valeur par défaut à spécifier au port de sortie.

Création d'une règle dans Informatica Developer

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

Créez un 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 22

Fiches d'évaluation dans Informatica Developer

Ce chapitre comprend les rubriques suivantes :

- [Présentation des fiches d'évaluation dans Informatica Developer, 176](#)
- [Création d'une fiche d'évaluation, 176](#)
- [Exportation d'un fichier de ressource pour le lignage des fiches d'évaluation, 177](#)
- [Affichage du lignage des fiches d'évaluation dans Informatica Developer, 177](#)

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 23

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, 179](#)
- [Exécution d'un profil sur un mapplet ou un objet de mappage, 179](#)
- [Comparaison des profils pour les objets de mappage ou de mapplet, 180](#)
- [Génération d'un mappage à partir d'un profil, 180](#)

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 **Profilier 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 24

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, 182](#)
- [Glossaire de domaine de données dans Informatica Developer, 183](#)
- [Options de découverte des domaines de données dans Informatica Developer, 187](#)
- [Création d'un profil pour découvrir des domaines de données dans Informatica Developer, 190](#)
- [Modification d'un profil dans Informatica Developer, 190](#)
- [Exécution d'un profil pour effectuer une découverte des domaines de données dans Informatica Developer, 191](#)
- [Résultats de la découverte des domaines de données dans Informatica Developer, 191](#)

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 avec lequel vous voulez comparer les données de colonne et le nom de colonne, les options d'échantillonnage, les options d'exploration et l'environnement d'exécution. Vous pouvez choisir le nombre maximal de lignes sur lesquelles vous voulez effectuer la découverte de domaines de données. Vous pouvez choisir un critère de conformité pour la découverte de domaines de données. Vous pouvez exclure les valeurs Null de la découverte de domaines de données. 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.

Vous pouvez créer un profil avec une option d'échantillonnage et des filtres pour effectuer la découverte des domaines de données. Lorsque vous exécutez le profil, vous appliquez l'option d'échantillonnage et les filtres sur la source de données et générez un ensemble de données. Le processus de découverte des domaines de données utilise l'ensemble de données pour découvrir les domaines de données.

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'exclure des colonnes de la découverte de 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 maximale de la colonne.
Échelle	Échelle de la colonne.
Valeur Null autorisée	Indique une colonne pouvant contenir 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 indiquer si le profil doit traiter toutes les lignes de la source de données. Vous pouvez choisir des critères de conformité pour la correspondance du domaine de données et d'exclure les valeurs Null de la découverte de domaines de données.

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

Option	Description
Remplacer les options d'inférence par défaut	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 et les titres de colonne.
Nombre maximal de lignes à profiler	Nombre maximal de lignes sur lesquelles le profil peut s'exécuter. L'outil Developer tool choisit les lignes à partir de la première ligne de la source.
Pourcentage minimal de lignes	Pourcentage minimal de conformité des lignes dans l'ensemble de données requis pour une correspondance de domaine de données.
Nombre minimal de lignes	Nombre minimal de lignes dans l'ensemble de données requis pour une correspondance de domaine de données.
Exclure les valeurs Null de la découverte de domaines de données	Exclut les valeurs Null de l'ensemble de données pour la découverte de domaines de données.

Pourcentage de conformité minimal

Vous pouvez choisir un pourcentage minimal de lignes dans l'ensemble de données comme critère de conformité pour la découverte de domaines de données.

Le pourcentage de conformité correspond au rapport entre le nombre de lignes correspondantes et le nombre total de lignes.

Remarque: L'outil Developer tool 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.

Exemple

Vous disposez d'une source de données de 10 000 lignes parmi lesquelles 2 500 lignes contiennent des numéros de sécurité sociale dans la colonne Commentaires. Vous créez un profil de colonne à l'aide de la découverte de domaines de données et définissez un pourcentage minimal de lignes de 30 % comme critères de conformité. Lorsque vous exécutez le profil, les résultats du profil n'affichent pas les numéros de sécurité sociale comme un domaine de données inféré, car le critère de conformité minimum est de 30 % des lignes de la source de données, c'est-à-dire 3 000 lignes.

Nombre minimal de lignes conformes

Vous pouvez choisir un nombre minimal de lignes dans l'ensemble de données comme critère de conformité pour la découverte de domaines de données.

Exemple

Vous disposez d'une source de données de 10 000 lignes parmi lesquelles 3 lignes contiennent une adresse électronique dans la colonne Commentaires. Vous créez un profil de colonne à l'aide d'un profil de découverte de domaines de données et définissez le nombre minimal de lignes sur 1 comme critère de conformité. Lorsque vous exécutez le profil, les résultats affichent l'adresse électronique comme un domaine de données inféré comprenant trois lignes conformes ainsi que les autres domaines de données inférés.

Exclure les valeurs Null

Vous pouvez exclure les valeurs Null lorsque vous effectuez la découverte de domaines de données sur une source de données. Lorsque vous sélectionnez le pourcentage minimal de lignes avec l'option d'exclusion des valeurs Null, le pourcentage de conformité est le rapport entre le nombre de lignes correspondantes et le nombre total de lignes moins les valeurs Null de la colonne.

Le processus de découverte des domaines de données est différent lorsque vous sélectionnez l'option **Exclure les valeurs Null de la découverte de domaines de données** et plusieurs options d'échantillonnage ou des filtres.

Les scénarios suivants expliquent les résultats de la découverte de domaine de données lorsque vous sélectionnez l'option Exclure les valeurs Null avec une option d'échantillonnage et des filtres :

- Option d'échantillonnage **Toutes les lignes** et aucun filtre. La découverte de domaines de données ignore toutes les valeurs Null de la colonne.
- Une option d'échantillonnage et aucun filtre. La découverte de domaines de données ignore toutes les valeurs Null dans les données de l'échantillon et s'exécute sur le reste de ces données.
- Option d'échantillonnage **Toutes les lignes** et des filtres. La découverte de domaines de données ignore toutes les valeurs Null dans les données filtrées et s'exécute sur le reste de ces données.
- Une option d'échantillonnage et des filtres. La découverte de domaines de données ignore les valeurs Null dans les données filtrées de l'échantillon et s'exécute sur le reste de ces données.

Exemple

Vous disposez d'une source de données de 10 000 lignes parmi lesquelles 3 000 lignes comprennent des numéros de sécurité sociale dans la colonne Commentaires. Vous créez un profil de colonne avec la découverte de domaines de données et choisissez les options suivantes :

- Sélectionnez l'option **Exclure les valeurs Null de la découverte de domaines de données**.
- Sélectionnez l'option d'échantillonnage **Toutes les lignes**.
- Sélectionnez l'option **Pourcentage minimal de lignes** et configurez-la sur 12 %.

Lorsque vous exécutez le profil, il s'exécute sur l'ensemble des données et ignore les valeurs Null lors de la découverte de domaines de données.

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 contenant des types de données spécifiques et modifier les options de sélection de colonnes, de sélection de domaines de données et 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 des colonnes**, vous pouvez cliquer sur **Exclure des colonnes** pour configurer les options d'exclusion selon les 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 de domaine de données affichent les statistiques sur les colonnes qui correspondent aux domaines de données, y compris les critères de conformité pour la correspondance de domaine de données, et la correspondance des noms de colonne 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 tool. 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 que vous sélectionnez la vue Domaine de données , Groupe de domaines de données ou Colonnes .
Connexion	Nom de la connexion.
Statut	Statut d'inférence de la colonne.
% de conformité des données	Pourcentage minimal de conformité des lignes requis pour une correspondance de domaine de données.
Conforme au nombre de lignes	Nombre minimal de lignes requis pour une correspondance de domaine de données.
% Null	Pourcentage de valeurs Null de la colonne.
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 documenté	Type de données déclaré pour la colonne de l'objet de profil.
Exploration	Si cette option est sélectionnée, la vue d'exploration des lignes s'affiche.
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

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

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

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** ou **Conforme au nombre de lignes** 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

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

Par défaut, Informatica Developer affiche les domaines de données inférés dans les résultats de profil. 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**.

Informatica 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

Lorsque vous exportez les résultats de la découverte de domaines de données vers un fichier `.xlsx` à partir de l'outil Informatica Developer, vous pouvez enregistrer le fichier sur le serveur ou à un emplacement spécifique de la machine cliente.

1. Exécuter un profil pour effectuer la découverte de 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. Vous pouvez éventuellement utiliser 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 25

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, 195](#)
- [Processus de découverte des données d'entreprise, 196](#)
- [Options de profil pour la découverte des données d'entreprise, 197](#)
- [Création d'un profil de découverte des données d'entreprise dans Informatica Developer, 202](#)
- [Modification d'un profil, 203](#)
- [Exécution d'un profil de découverte des données d'entreprise, 204](#)
- [Découverte de clé étrangère, 205](#)
- [Analyse de jointure, 207](#)
- [Découverte du chevauchement, 209](#)
- [Fichiers de script DDL, 211](#)

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.

Vous pouvez sélectionner un profil de système d'exploitation dans Informatica Developer. Après avoir choisi un profil de système d'exploitation, le service d'intégration de données crée et exécute les profils de découverte de données d'entreprise selon les autorisations de l'utilisateur du système d'exploitation défini dans le profil du système d'exploitation.

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 indiquer si le profil doit traiter toutes les lignes de la source de données, puis choisir des critères de conformité pour la correspondance de domaine de données.

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.
Pourcentage minimal de lignes	Pourcentage minimal de conformité des lignes dans l'ensemble de données requis pour une correspondance de domaine de données. Le pourcentage de conformité correspond au rapport entre le nombre de lignes correspondantes et le nombre total de lignes. Remarque: L'outil Developer tool considère les valeurs Null comme des lignes qui ne correspondent pas.
Nombre minimal de lignes	Nombre minimal de lignes dans l'ensemble de données requis pour une correspondance de domaine de données.
Exclure les valeurs Null de la découverte de domaines de données	Exclut les valeurs Null de l'ensemble de données pour la découverte de domaines de données.
Toutes les lignes	Le profil s'exécute sur toutes les lignes de la source de données.

Option	Description
Échantillon 1re ligne	Nombre maximal de lignes sur lesquelles le profil peut s'exécuter. L'outil Developer tool choisit les lignes à partir de la première ligne de la source.
Excluez les types et les domaines de données approuvés de l'inférence du type et du domaine de données dans les exécutions de profil suivantes.	Exclut le type de données approuvé ou le domaine de données du type de données et l'inférence de domaine de données de l'exécution de profil suivante.

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> premières lignes	Nombre de lignes sur lesquelles vous voulez exécuter le profil de colonne. L'outil Developer tool 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 contenant un type de données approuvé émanant de l'inférence de type de données de l'exécution du profil de colonne.

Option d'environnement d'exécution

Choisissez l'option d'environnement d'exécution natif. Informatica Developer définit l'environnement d'exécution dans la définition de profil une fois que vous avez choisi l'environnement d'exécution. L'environnement d'exécution n'affecte pas les résultats de profil.

Le tableau suivant décrit les options d'environnement d'exécution pour un profil de découverte de données d'entreprise :

Option	Description
Natif	L'outil Developer tool envoie les tâches de profil au module de service de profilage. Celui-ci divise alors les tâches de profil en un ensemble de mappages. Le service d'intégration de données exécute les mappages et enregistre les résultats de profil dans l'entrepôt de profils.

Options d'inférence de clé primaire 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 primaire.
Nombre max. de colonnes clés	Nombre maximal de colonnes qui peuvent former une clé primaire.
Nombre max. de lignes	Nombre maximal de lignes sur lesquelles vous pouvez exécuter le profil.
Pourcentage minimal	Pourcentage de conformité minimal des données de colonne requises pour la correspondance de clé primaire.
Lignes de violation maximales	Nombre maximal de lignes avec violations de clés que le profil autorise lors de la définition des clés primaires.

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.

Vous pouvez inférer les clés étrangères dans Informatica Developer à l'aide de l'une des méthodes suivantes :

- Utilisez les valeurs par défaut.
- Configurez les options d'inférence de la clé étrangère.
- Utilisez la clé étrangère du fichier de configuration pour configurer les paramètres de conservation automatique.

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.

Options	Description
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 tool 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	<p>Nombre de clés primaires de rang supérieur utilisées dans l'inférence de clé étrangère lorsque l'outil Developer tool exécute un profil de clé étrangère sur toutes les sources de données. L'outil Developer tool 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.</p> <p>Le classement des clés inférées est basé sur le pourcentage de conformité décroissant arrondi à une seule décimale. Par exemple, l'outil Developer tool considère un pourcentage de conformité de 99,75 comme étant 99,8 et 99,74 comme étant 99,7.</p> <p>La valeur par défaut est 1. Définissez la valeur sur -1 si vous voulez que l'outil Developer tool utilise toutes les clés inférées dans l'inférence de clé étrangère.</p> <p>Remarque: Si les sources de données de clé primaire disposent de clés primaires approuvées, l'outil Developer tool n'utilise pas de clés primaires inférées pour l'inférence de clé étrangère.</p>
Nombre maximal de clés étrangères entre les objets de données	Nombre maximal de colonnes inférées renvoyées par l'outil Developer tool après l'exécution du profil et qui sont requises pour la découverte des clés étrangères.
Pourcentage de conformité minimal	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.

Paramètres de conservation automatique pour l'inférence de clé étrangère

Vous pouvez configurer les paramètres de conservation automatique pour inférer les relations de clé primaire et de clé étrangère sans intervention manuelle. Les paramètres de conservation automatique sont des attributs personnalisés définis par l'utilisateur que vous pouvez configurer pour identifier les relations des données en fonction de certaines conditions.

Lorsque les résultats de la découverte incluent un grand nombre de relations de clé primaire et de clé étrangère, vous pouvez avoir des difficultés à identifier les relations de données critiques parmi des centaines de relations de données. Vous pouvez également rencontrer des difficultés à conserver les relations en fonction de certaines conditions, comme la correspondance de données ou le type de données. Pour résoudre ce problème, vous pouvez configurer les paramètres de conservation automatique et exécuter le profil de découverte des données d'entreprise.

Si les sources de données ont plusieurs clés étrangères candidates et que vous souhaitez établir des règles pour choisir une clé étrangère candidate, vous pouvez effectuer les actions suivantes :

- Configurer les options **Nombre maximal de clés étrangères entre les objets de données** et **Pourcentage minimal de conformité** dans l'assistant profil de découverte des données d'entreprise.
- Configurer les pondérations et les scores des paramètres de conservation automatique dans le fichier `ForeignKeyConfig.xml`.

Un administrateur peut modifier et enregistrer un fichier de configuration de la clé étrangère. Configurer les paramètres de conservation automatique dans le fichier de configuration de la clé étrangère. L'algorithme interfère les relations de clé primaire et de clé étrangère entre plusieurs objets de données selon les paramètres de conservation automatique.

Le fichier de configuration de la clé étrangère `foreignkeyconfig.xml` est disponible dans le répertoire suivant :

```
<répertoire d'installation Informatica>\services\DataIntegrationService\modules  
\ProfilingService
```

Les paramètres de conservation automatique comprennent la correspondance de chevauchement des données, la correspondance du nom de colonne, la correspondance du type de relation et la correspondance de type de données.

Correspondance de chevauchement des données

La correspondance de chevauchement des données est l'estimation de chevauchement des valeurs entre la clé primaire et les clés étrangères. Vous pouvez définir la correspondance de chevauchement dans l'assistant profil de découverte des données d'entreprise avec l'option **Pourcentage minimal de conformité**. Par défaut, l'option **Pourcentage minimal de conformité** est définie sur 90.

Si la correspondance de chevauchement des données ne satisfait pas au pourcentage minimal de conformité, la clé étrangère n'est pas prise en compte pour la conservation automatique. Lorsque la correspondance de chevauchement des données satisfait au pourcentage minimal de conformité, les paramètres restants sont utilisés pour calculer le score ajusté.

Correspondance de nom

Le paramètre de correspondance de nom est un paramètre facultatif. Il utilise l'algorithme Distance de modification pour déterminer le niveau de conformité des noms des colonnes de clé primaire et de clé étrangère et définit le score entre 0 et 1. Définissez la pondération de la correspondance de nom sur 0 si vous ne voulez pas utiliser ce paramètre pour déterminer la relation entre la clé primaire et les clés étrangères.

Correspondance de type de relations

La correspondance de type de relations détermine le type de relations entre les colonnes de clé primaire et de clé étrangère et attribue un score fixe entre 0 et 1. La correspondance de type de relations est calculée en fonction du type de colonne de la colonne de clé étrangère.

Les correspondances de type de relations suivantes peuvent être définies dans le fichier

`ForeignKeyConfig.xml` :

- Relation clé primaire à clé étrangère, dans laquelle la colonne de clé étrangère est une colonne non-clé. La valeur par défaut de cette correspondance de type de relations est 1. Vous pouvez trouver ce type de relation dans de nombreuses sources de données.
- Relation clé primaire à clé primaire, dans laquelle la colonne de clé étrangère est une colonne de clé primaire. La valeur par défaut pour ce type de correspondance est 0,25. Ce type de relation se présente rarement, car il représente une table partitionnée verticalement.

- Relation clé primaire à séquence clé primaire, dans laquelle la colonne de clé étrangère est une colonne de clé primaire et le type de données de la colonne est un type de données de séquence. Par exemple, la colonne OrderID dans une table Order présente un type de données de séquence. La valeur par défaut pour ce type de relation est zéro, car les clés de séquence pourraient causer plusieurs clés étrangères faux positif, que l'algorithme clé primaire à clé primaire essaye d'éviter. Vous pouvez définir la correspondance de type de relations sur un score plus élevé si vous savez que la source de données contient quelques types de données de séquence.

Correspondance de type de données

La correspondance de type de données compare les types de données des colonnes de clé primaire avec les colonnes de clé étrangère et attribue un score de conformité fixe basé sur le niveau de correspondance des types de données des colonnes.

Le tableau suivant reprend les scores de correspondance de type de données fixes pour différentes combinaisons de clé primaire et de clés étrangères :

	Clé étrangère numérique	Clé étrangère de date	Clé étrangère de chaîne
Clé primaire numérique	1,0	0,5	0,0
Clé primaire de date	0,5	1,0	0,5
Clé primaire de chaîne	0,0	0,0	1,0

Vous pouvez modifier les scores de correspondance de type de données par défaut si nécessaire.

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 des données 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 que vous avez 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 **Sélectionner** 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 profils que vous voulez exécuter. Vous pouvez configurer les types de profils 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 « Exécuter la découverte des données d'entreprise »** pour les types de profils que vous voulez exécuter dans le cadre du profil de découverte des données d'entreprise. Le profilage de colonnes 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 modèle de profil de découverte des données d'entreprise à la fin** pour exécuter le profil une fois sa configuration terminée. Si vous activez 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. Ensuite, l'outil Developer exécute 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 l'affichage des résultats. Cette étape est obligatoire car l'importation des 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 contenant des types de données spécifiques et modifier les options de sélection de colonnes, de sélection de domaines de données et d'inférence.

1. Dans la vue **Explorateur d'objets**, sélectionnez le projet ou le dossier qui contient le profil que vous voulez modifier.
2. Cliquez sur **Équipe > Extraire** pour extraire le profil.
3. Double-cliquez sur le profil pour l'ouvrir.
4. Cliquez sur la vue **Propriétés**.

La vue Propriétés se trouve dans la vue Par défaut.
5. Cliquez sur **Profils** pour afficher les tâches de profil.

6. 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.
7. 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**.
8. Apportez les modifications nécessaires aux options de définition de profil.
9. Enregistrez les modifications.
10. Cliquez sur **Équipe > Archiver** pour archiver le profil.

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	Colonne que le profil considère comme une clé étrangère par rapport à la clé primaire parent dans la ligne actuelle.
% 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 l'exécution du profil. Si vous définissez une relation avant l'exécution du 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.

Propriété	Description
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.

- Vérifiez que les colonnes de jointure Gauche et Droite sont préfixées avec les noms d'objet de données corrects.
- 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.

- Dans la vue **Explorateur d'objets**, ouvrez le profil de découverte des données d'entreprise qui contient l'analyse de jointure.
- Exécutez le profil de jointure.
- Sélectionnez la vue **Résultats de jointure**.
- 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.
- 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.
% 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

Colonne de gauche	Colonne de droite
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 26

Résultats de la découverte des données d'entreprise

Ce chapitre comprend les rubriques suivantes :

- [Présentation des résultats de la découverte des données d'entreprise, 212](#)
- [Vue Relations, 213](#)
- [Vue Profilage de clé étrangère, 214](#)
- [Vue tabulaire, 217](#)
- [Vue Domaines de données, 219](#)
- [Vue Profil de colonne, 220](#)
- [Affichage des résultats de profil de colonne lors de l'exécution de la découverte des données d'entreprise, 221](#)
- [Affichage des résultats de la découverte des domaines de données lors de l'exécution de la découverte des données d'entreprise, 221](#)
- [Affichage du statut d'exécution de la découverte des données d'entreprise, 221](#)
- [Fichiers d'exportation de la découverte des données d'entreprise, 222](#)

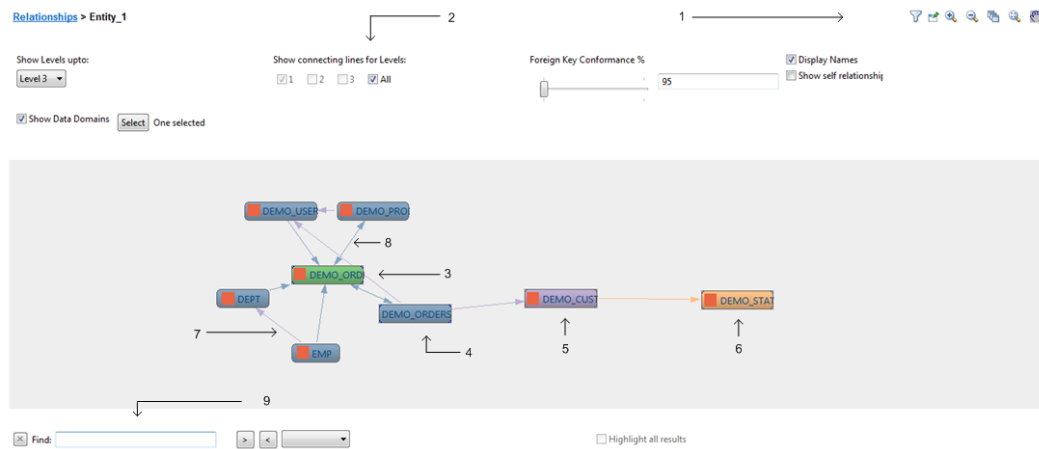
Présentation des résultats de la découverte des données d'entreprise

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.

L'image 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**.

Remarque: La vue **Profilage de clé étrangère** n'affiche aucun objet de données lorsque vous utilisez des sources de données Hive pour créer un profil de découverte de données d'entreprise.

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

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. Pour rejeter une relation de colonnes inférées dans la vue de graphique, sélectionnez l'objet de données, cliquez dessus avec le bouton droit de la souris, puis sélectionnez l'une des options suivantes :
 - **Rejeter toutes les relations avec des > Clés primaires inférées**. Choisissez cette option pour rejeter toutes les relations entre les colonnes contenant des clés primaires inférées dans l'objet de données vers les colonnes contenant des clés étrangères inférées dans les autres objets de données connectés.
 - **Rejeter toutes les relations avec des > Clés étrangères inférées**. Choisissez cette option pour rejeter toutes les relations entre les colonnes contenant des clés étrangères inférées dans l'objet de données vers les colonnes contenant des clés primaires inférées dans les autres objets de données connectés.
 - **Rejeter toutes les relations avec des > Clés primaires et étrangères inférées**. Choisissez cette option pour rejeter toutes les relations entre les colonnes contenant des clés primaires inférées dans l'objet de données vers les colonnes contenant des clés étrangères inférées dans les autres objets de données connectés, et toutes les relations entre les colonnes contenant des clés étrangères inférées dans l'objet de données vers les colonnes contenant des clés primaires inférées dans les autres objets de données connectés.
4. Dans la vue graphique, cliquez avec le bouton droit de la souris sur un objet de données et sélectionnez **Afficher les relations de colonnes**.
5. Sélectionnez la relation de l'objet de données que vous voulez conserver.
6. 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**.
7. Pour restaurer le statut inféré de la relation de colonne, cliquez avec le bouton droit de la souris et cliquez sur **Réinitialiser**.
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 une ligne et 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** ou **Conforme au nombre de lignes** 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 telles que les valeurs uniques, les valeurs Null, les types de données et les valeurs minimales et maximales de chaque colonne des 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 27

Bureau de glossaire métier dans Informatica Developer

Ce chapitre comprend les rubriques suivantes :

- [Recherche dans Business Glossary, 223](#)
- [Recherche d'un terme métier, 224](#)
- [Personnalisation des touches de raccourci pour rechercher un terme métier, 224](#)

Recherche dans Business Glossary

Recherchez la signification d'un nom d'objet de l'outil Developer en tant que terme d'entreprise dans le Business Glossary Desktop pour comprendre sa spécification dans l'entreprise et son implémentation actuelle.

Un glossaire d'entreprise est un ensemble de termes qui définissent des concepts dans un langage d'entreprise spécifique pour les utilisateurs professionnels. Un terme d'entreprise fournit la définition et l'utilisation par l'entreprise d'un concept. Le Business Glossary Desktop est un client qui se connecte au service Metadata Manager, qui héberge le glossaire d'entreprise. Utilisez le Business Glossary Desktop pour rechercher des termes d'entreprise dans un glossaire d'entreprise.

Si le Business Glossary Desktop 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 d'entreprise. Vous pouvez rechercher des noms d'objets dans les vues de l'outil Developer tool, telles que la vue **Explorateur d'objets**, ou bien des noms de colonnes, des profils et des ports de transformation dans l'éditeur.

Par exemple, un développeur souhaite rechercher dans un glossaire d'entreprise un terme d'entreprise qui correspond à l'objet de données Sales_Audit dans l'outil Developer. Le développeur souhaite afficher les détails du terme d'entreprise 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**.

INDEX

A

analyse de jointure
présentation [207](#)

C

conservation
 concepts [32](#)
 Informatica Analyst [75](#)
 Informatica Developer [172](#)
 processus [33](#)
 tâches [33](#)
création d'un profil de colonne
 profils [40](#)
création d'une règle d'expression
 règles [48](#)

D

découverte de clé étrangère
 présentation [205](#)
découverte de clé primaire
 présentation [153](#)
découverte de dépendance fonctionnelle
 présentation [155](#)
découverte de domaines de données
 présentation [29](#)
 présentation d'Informatica Analyst [109](#)
 présentation d'Informatica Developer [182](#)
 processus [31](#)
découverte de données
 présentation [22](#)
 processus [20](#)
découverte de données d'entreprise dans l'outil Analyst
 présentation [121](#)
découverte des données d'entreprise
 affichage des relations d'objet de données [214](#)
 exécution dans Informatica Analyst [124](#)
 modification [203](#)
 modification dans l'outil Analyst [126](#)
 présentation [195](#)
 processus [196](#)
 statut d'exécution [221](#)
 vue domaines de données [219](#)
 vue Profil de colonne [220](#)
 vue Profilage de clé étrangère [214](#)
 vue relations [213](#)
 vue tabulaire [217](#)
découverte des données d'entreprise dans l'outil Analyst
 conflit de type de données [130](#)
 processus [122](#)
 vue Profils [131](#)
 vue Résumé [128](#)

découverte du chevauchement
 exécution de [210](#)
 présentation [209](#)
 résultats [209](#)
domaine de données
 création à partir de résultats de profil dans Informatica Analyst [111](#)
 création à partir de résultats de profil dans Informatica Developer [184](#)
 création dans Informatica Analyst [110](#)
 création dans Informatica Developer [184](#)
 présentation [30](#)
 rechercher dans Informatica Developer [185](#)
domaines de données
 exportation [186](#)
 importation [185](#)
 rechercher dans Informatica Analyst [111](#)

E

exporter
 lignage des fiches d'évaluation vers XML [177](#)

F

fiche d'évaluation
 configuration des notifications [106](#)
 configuration des paramètres généraux des notifications [106](#)
fiches d'évaluation
 affichage [88](#)
 ajout de colonnes à une fiche d'évaluation [87](#)
 coût des données non valides [90](#)
 coût fixe [90](#)
 coût variable [90](#)
 création d'un groupe de mesures [91](#)
 définition des seuils [91](#)
 déplacement des scores [92](#)
 développement [93](#)
 exécution [88](#)
 graphe de tendance [93](#)
 Groupes de mesures [91](#)
 Informatica Analyst [83](#)
 Informatica Developer [176](#)
 mesures [90](#)
 modification [89](#)
 modification d'un groupe de mesures [92](#)
 notifications [104](#)
 poids métriques [90](#)
 présentation [28](#)
 processus de Informatica Analyst [84](#)
 Suppression d'un groupe de mesures [93](#)
filtres
 présentation [50](#)

G

- glossaire de domaine de données
 - Informatica Analyst [110](#)
 - Informatica Developer [183](#)
 - présentation [31](#)
- graphes de tendance
 - affichage [95](#)
 - coût [94](#)
 - exportation depuis Informatica Analyst [96](#)
 - score [94](#)
- groupe de domaines de données
 - création dans Informatica Analyst [110](#)
 - création dans Informatica Developer [183](#)
 - présentation [30](#)

I

- Informatica Analyst
 - présentation des profils de colonne [36](#), [69](#)
 - règles [45](#)
 - résultats de profil de colonne [55](#), [67](#)
 - verrouillage et gestion de version [40](#)
- Informatica Developer
 - présentation des profils [145](#)
 - règles [174](#)
 - vues de profils [147](#)

L

- lignage des fiches d'évaluation
 - affichage dans Informatica Analyst [107](#)
 - affichage dans Informatica Developer [177](#)

O

- objet de données fichier plat
 - synchronisation [43](#)
- objet de données table
 - synchronisation [44](#)
- objet de mappage
 - exécution d'un profil [179](#)
- objets de données avec fiche d'évaluation
 - Informatica Analyst [100](#)
- options de configuration
 - découverte des données d'entreprise dans l'outil Analyst [122](#)
- options de découverte des domaines de données
 - Informatica Developer [187](#)
- options de profil
 - découverte des données d'entreprise [197](#)

P

- profil
 - composants [23](#)
 - Formats Avro ou Parquet [163](#)
 - formats XML et JSON [162](#), [163](#)
- profil d'objet de données
 - commentaires [160](#)
- profil de clé étrangère
 - découverte [206](#)
- profil de colonne
 - exploration [74](#)
 - Informatica Developer [150](#)

- profil de colonne (*a continué*)
 - options [27](#)
 - présentation [26](#)
 - processus [37](#)
 - profil du système d'exploitation [39](#), [157](#)
- profil de découverte des données d'entreprise
 - création de scripts DDL [211](#)
 - exécution [204](#)
 - Scripts DDL [211](#)
- profilage
 - architecture [19](#)
 - présentation [17](#)
 - verrouillage et gestion de version [27](#)
- Profilage de mapplet et de mappage
 - Présentation [179](#)
- profils
 - création d'un filtre [50](#)
 - création d'un profil de colonne [40](#)
 - exécution [42](#), [68](#), [69](#), [117](#)
 - modification d'un filtre [53](#)
 - modification d'un profil de colonne [42](#)
- profils d'objet de données
 - création d'un profil unique [157](#)
 - création de plusieurs profils [158](#)
 - découverte des données d'entreprise [202](#)
 - présentation [149](#)
- Profils de mapplet et de mappage
 - Présentation [179](#)
- projets
 - terme métier Metadata Manager [142](#)

R

- recherche de découverte
 - prérequis [133](#)
- recherche de découverte dans l'outil Analyst
 - processus [134](#)
- rechercher
 - glossaire d'entreprise [223](#)
- règles
 - application d'une règle prédéfinie [46](#)
 - application dans Informatica Developer [175](#)
 - application dans PowerCenter Express [175](#)
 - création d'une règle d'expression [48](#)
 - création dans Informatica Developer [175](#)
 - expression [47](#)
 - prédéfinies [46](#)
 - prérequis [174](#)
- règles d'expression
 - processus [47](#)
- règles prédéfinies
 - processus [46](#)
- résultats de fiche d'évaluation
 - exportation [103](#)
 - exportation depuis Informatica Analyst [103](#)
 - exporter vers Excel [103](#)
- résultats de la découverte de domaines de données
 - exportation depuis Informatica Analyst [120](#)
- Résultats de la découverte de domaines de données
 - Informatica Analyst [118](#)
- résultats de la découverte des domaines de données
 - exportation dans Informatica Analyst [119](#)
 - exportation depuis Informatica Developer [194](#)
- Résultats de la découverte des domaines de données
 - Informatica Developer [191](#)
- résultats de la découverte des données d'entreprise
 - enregistrer en tant qu'image [216](#)

- résultats de la découverte des données d'entreprise (*a continué*)
 - exportation [222](#)
 - présentation [212](#)
- résultats de la découverte des données d'entreprise dans Informatica Analyst
 - présentation [127](#)
- résultats de la recherche de découverte
 - interface [136](#)
- résultats de la recherche de découverte dans l'outil Analyst
 - présentation [136](#)
- résultats de profil
 - ajout de commentaires dans Informatica Developer [160](#)
 - approbation de types de données [75](#)
 - approbation des domaines de données [118](#)
 - approbation des domaines de données dans Informatica Developer [193](#)
 - approbation des types de données dans Informatica Developer [172](#)
 - balises [81](#)
 - commentaires [80](#)
 - conservation des relations de colonne dans Informatica Developer [218](#)
 - développement [74](#)
 - Excel [77](#)
 - exportation [76](#)
 - exportation dans Informatica Developer [173](#)
 - exportation depuis Informatica Analyst [77](#)
 - formes de colonne [63](#)
 - rejet de types de données [76](#)
 - rejet de types de données dans l'outil Developer tool [172](#)
 - rejet des domaines de données [118](#)
 - rejet des domaines de données dans Informatica Developer [194](#)
 - résumé [71](#), [74](#)
 - termes d'entreprise [80](#)
 - types de données de colonne [61](#), [171](#)
 - valeurs de colonne [64](#)
 - vue de résumé [57](#)
 - vue détaillée [59](#)
- résultats de profil de colonne
 - Informatica Developer [168](#)

- résultats de profil de colonne (*a continué*)
 - profil de colonne [168](#)
- résultats de profil de colonne dans l'outil Analyst tool
 - interface [58](#), [72](#)
- résultats de profil de colonne dans l'outil Analyst tool
 - détails de colonne [58](#), [73](#)
 - interface [70](#)
 - résumé [56](#)
- résultats de profil de la découverte de domaines de données
 - Microsoft Excel [119](#)

T

- tableau de bord des fiches d'évaluation
 - Informatica Analyst [97](#)
- terme métier
 - recherche d'un terme métier [143](#)
- terme métier Metadata Manager
 - gestion des termes métier [143](#)
- projets [142](#)
- termes métier
 - personnalisation des touches de raccourci [224](#)
 - recherche [224](#)

V

- valeur hors norme
 - détection [63](#)
- volet des scores cumulatifs
 - Informatica Analyst [101](#)
- volet Fiches d'évaluation par projet
 - Informatica Analyst [98](#)
- volet Tendance d'exécution des fiches d'évaluation
 - Informatica Analyst [99](#)