



Informatica®

10.1.1 HotFix 2

Guide de profil

Informatica Guide de profil
10.1.1 HotFix 2
Novembre 2017

© Copyright Informatica LLC 2010, 2018

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.

Certaines parties de ce logiciel et/ou de cette documentation sont soumises à des droits d'auteur détenus par des tiers. Les notifications de tiers requises sont incluses avec le produit.

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: 2018-06-08

Sommaire

Préface.....	8
Ressources Informatica.	8
Informatica Network.	8
Base de connaissances Informatica.	8
Documentation Informatica.	9
Matrices de disponibilité de produit Informatica.	9
Informatica Velocity.	9
Informatica Marketplace.	9
Support client international Informatica.	9
 Partie I: Introduction aux profils.....	 10
 Chapitre 1: Introduction aux profils.....	 11
Présentation de l'introduction aux profils.	11
Processus de profilage.	12
Outils de profilage.	12
Composants de profil.	12
 Chapitre 2: Concepts de profil de colonne.....	 14
Présentation des concepts de profil de colonnePrésentation des profils.	14
Options de profil de colonne.	15
Verrous de profil de référentiel et gestion de profil avec versions.	15
Fiches d'évaluation.	16
 Chapitre 3: Concepts de conservation.....	 17
Présentation des concepts de conservation.	17
Tâches de conservation.	18
 Partie II: Profilage avec Informatica Analyst.....	 19
 Chapitre 4: Profils de colonne dans Informatica Analyst.....	 20
Présentation des profils de colonne dans Informatica Analyst.	20
Processus de profilage de colonne.	21
Options de profil.	21
Options d'échantillonnage.	22
Options de développement.	22
Environnement d'exécution.	23
Présentation des profils de système d'exploitation dans Informatica Analyst.	23
Sélection d'un profil de système d'exploitation.	23
Présentation des verrous des ressources du référentiel et du développement basé sur l'équipe.	23

Création d'un profil de colonne dans Informatica Analyst.	24
Modification d'un profil de colonne.	26
Exécution d'un profil.	26
Option de synchronisation.	27
Synchronisation d'un objet de données fichier plat.	27
Synchronisation d'un objet de données relationnel.	28
Chapitre 5: Règles dans Informatica Analyst.	29
Présentation des règles dans Informatica Analyst.	29
Règles d'un profil de colonne.	29
Règles prédéfinies.	30
Processus de règles prédéfinies.	30
Application d'une règle prédéfinie.	30
Règles d'expression.	31
Processus de règles d'expression.	31
Création d'une règle d'expression.	32
Chapitre 6: Résultats de profil de colonne dans Informatica Analyst.	34
Présentation des résultats de profil de colonne dans Informatica Analyst.	34
Vue de résumé.	35
Propriétés de la vue de résumé.	36
Filtres par défaut de la vue de résumé.	37
Vue détaillée.	37
Volets de la vue détaillée.	38
Statistiques.	39
Aperçu des données.	40
Types de données.	40
Valeurs hors norme.	41
Modèles.	42
Valeurs.	43
Types d'exécution de profil.	46
Dernière exécution de profil.	46
Exécution de profil historique.	46
Exécution de profil consolidée.	47
Sélection d'une exécution de profil.	47
Présentation des résultats de la comparaison de plusieurs profils.	48
Résultats de la comparaison de plusieurs profils.	48
Vue de résumé des résultats de la comparaison de profils	49
Vue détaillée des résultats de la comparaison de profils	52
Développement de profil de colonne.	53
Développement des données de ligne.	53
Application de filtres aux données de développement.	53
Conservation dans l'outil Analyst.	54

Approbation de types de données et de domaines de données.	54
Rejet de types de données et de domaines de données.	55
Fichiers d'exportation de profil de colonne dans Informatica Analyst.	55
Résultats de l'exportation de profil dans un fichier CSV.	55
Résultats de l'exportation de profil dans Microsoft Excel.	56
Exportation de résultats de profil depuis Informatica Analyst.	56

Chapitre 7: Fiches d'évaluation dans Informatica Analyst. 58

Présentation des fiches d'évaluation dans Informatica Analyst.	58
Processus de fiche d'évaluation de Informatica Analyst.	59
Création d'une fiche d'évaluation dans Informatica Analyst.	60
Ajouter des colonnes à une fiche d'évaluation existante.	61
Exécution d'une fiche d'évaluation.	63
Affichage d'une fiche d'évaluation.	63
Modification d'une fiche d'évaluation.	64
Mesures.	65
Poids métriques.	65
Valeur de la qualité des données.	65
Définition des seuils.	66
Groupes de mesures.	66
Création d'un groupe de mesures.	66
Déplacement des scores vers un groupe de mesures.	67
Modification d'un groupe de mesures.	67
Suppression d'un groupe de mesures.	68
Développement dans les colonnes.	68
Graphes de tendance.	68
Graphe de tendance de score.	69
Graphe de tendance de coût.	69
Affichage des graphes de tendances.	70
Exportation des graphes de tendance.	71
Fichier d'exportation de fiche d'évaluation dans Informatica Analyst.	72
Résultats de l'exportation de fiche d'évaluation dans Microsoft Excel.	72
Exportation des résultats de fiche d'évaluation à partir d'Informatica Analyst.	72
Notifications de la fiche d'évaluation.	73
Modèle de notification de message électronique.	73
Configuration des notifications de la fiche d'évaluation.	75
Configuration des paramètres généraux des notifications de la fiche d'évaluation.	75
Lignage des fiches d'évaluation.	76
Affichage du lignage des fiches d'évaluation dans Informatica Analyst.	76

Partie III: Profilage avec Informatica Developer..... 78

Chapitre 8: Profils d'objet de données..... 79

Profils de colonne dans Informatica Developer.	79
Options de filtrage.	80
Options d'échantillonnage.	80
Création d'un profil d'objet de données unique dans Informatica Developer.	80
Création de plusieurs profils d'objet de données dans Informatica Developer.	81
Synchronisation d'un objet de données de fichier plat.	82
Synchronisation d'un objet de données relationnel.	82

Chapitre 9: Profils de colonne dans des sources de données semi-structurées..... 83

Présentation de profils de colonne dans des sources de données semi-structurées.	83
Profil de colonne dans une source de données XML ou JSON.	84
Création d'un objet de données sur une source de données XML ou JSON.	84
Profil de colonne dans une source de données semi-structurée dans HDFS.	85
Profil de colonne sur une source de données XML ou JSON dans HDFS.	85
Profil de colonne d'une source de données Avro ou Parquet dans HDFS.	85
Création d'une connexion HDFS.	86
Création d'un objet de données de fichier complexe sur un fichier JSON ou XML dans HDFS.	86
Création d'un objet de données sur une source de données Avro ou Parquet.	87
Exécution d'un profil de colonne sur des sources de données semi-structurées.	88

Chapitre 10: Règles dans Informatica Developer..... 90

Présentation des règles dans Informatica DeveloperLignes directrices relatives aux règles.	90
Création d'une règle dans Informatica Developer.	91
Application d'une règle dans Informatica DeveloperApplication d'une règle.	91

Chapitre 11: Résultats de profil de colonne dans Informatica Developer..... 92

Résultats de profil de colonne dans Informatica DeveloperRésultats de profil de colonne.	92
Propriétés des valeurs de colonne.	93
Propriétés des formes de colonne.	94
Propriétés des statistiques de colonne.	94
Propriétés des types de données de colonne.	95
Conservation dans Informatica Developer Conservation dans Informatica Developer.	96
Approbation des types de donnéesApprobation des types de données dans Informatica Developer.	96
Rejet de types de donnéesRejet de types de données dans Informatica Developer.	96
Exportation des résultats de profil depuis Informatica Developer.	97

Chapitre 12: Fiches d'évaluation dans Informatica Developer.	98
Présentation des fiches d'évaluation dans Informatica Developer.	98
Création d'une fiche d'évaluation.	98
Exportation d'un fichier de ressource pour le lignage des fiches d'évaluation.	99
Affichage du lignage des fiches d'évaluation dans Informatica Developer.	99
 Chapitre 13: Profilage de mapplet et de mappage.	 101
Présentation du profilage de mapplet et de mappageProfils de mapplet et de mappage.	101
Exécution d'un profil sur un mapplet ou un objet de mappage.	101
Comparaison des profils pour les objets de mappage ou de mapplet.	102
Génération d'un mappage à partir d'un profil.	102
 Index.	 104

Préface

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

Utilisez les profils de colonne pour déterminer les caractéristiques des colonnes, telles que la fréquence de valeur, les pourcentages et les modèles, dans une source 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 aux profils

Cette partie contient les chapitres suivants :

- [Introduction aux profils, 11](#)
- [Concepts de profil de colonne, 14](#)
- [Concepts de conservation, 17](#)

CHAPITRE 1

Introduction aux profils

Ce chapitre comprend les rubriques suivantes :

- [Présentation de l'introduction aux profils, 11](#)
- [Processus de profilage, 12](#)
- [Outils de profilage, 12](#)
- [Composants de profil, 12](#)

Présentation de l'introduction aux profils

Créez et exécutez un profil 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 les outils Developer et Analyst :

- Effectuer le profilage de colonnes. Ce processus comprend la découverte du nombre de valeurs uniques, de valeurs Null et de modèles de données dans une colonne.
- Ajouter des règles aux profils de colonne.
- Conservez les types de données inférés dans les résultats de profil.
- Utilisez les fiches d'évaluation pour contrôler la qualité des données.
- Générer un objet de mappage à partir d'un profil.

Processus de profilage

Lorsque vous démarrez 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 utilisez différents profils pour différents types d'analyse de données, comme un profil de colonne. Vous devez découvrir et documenter les problèmes de qualité des données. Exécutez les tâches suivantes pour effectuer le profilage :

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. Vérifier les résultats de profilage.
3. Créer des données de référence.
4. Développer les résultats de profil.
5. Documenter les problèmes de données.
6. Créer et exécuter les règles.
7. Créer des fiches d'évaluation pour contrôler la qualité des données.

Outils de profilage

Vous utilisez plusieurs outils Informatica pour gérer le processus de profilage.

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

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 des profils dans cet outil pour rechercher et analyser les métadonnées d'une ou de plusieurs sources de données. Vous devez créer des profils à l'aide d'un assistant.

Informatica Analyst

Vous pouvez exécuter un profil de colonne dans l'outil Analyst. Après avoir exécuté un profil, vous pouvez développer les lignes dans une source de données.

Composants de profil

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.

Commentaire

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

Fiche d'évaluation

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

CHAPITRE 2

Concepts de profil de colonne

Ce chapitre comprend les rubriques suivantes :

- [Présentation des concepts de profil de colonne](#)[Présentation des profils, 14](#)
- [Options de profil de colonne, 15](#)
- [Verrous de profil de référentiel et gestion de profil avec versions, 15](#)
- [Fiches d'évaluation, 16](#)

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 3

Concepts de conservation

Ce chapitre comprend les rubriques suivantes :

- [Présentation des concepts de conservation, 17](#)
- [Tâches de conservation, 18](#)

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.

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 : Profilage avec Informatica Analyst

Cette partie contient les chapitres suivants :

- [Profils de colonne dans Informatica Analyst, 20](#)
- [Règles dans Informatica Analyst, 29](#)
- [Résultats de profil de colonne dans Informatica Analyst, 34](#)
- [Fiches d'évaluation dans Informatica Analyst, 58](#)

CHAPITRE 4

Profils de colonne dans Informatica Analyst

Ce chapitre comprend les rubriques suivantes :

- [Présentation des profils de colonne dans Informatica Analyst, 20](#)
- [Processus de profilage de colonne, 21](#)
- [Options de profil, 21](#)
- [Environnement d'exécution, 23](#)
- [Présentation des profils de système d'exploitation dans Informatica Analyst, 23](#)
- [Présentation des verrous des ressources du référentiel et du développement basé sur l'équipe, 23](#)
- [Création d'un profil de colonne dans Informatica Analyst, 24](#)
- [Modification d'un profil de colonne, 26](#)
- [Exécution d'un profil, 26](#)
- [Option de synchronisation, 27](#)

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 5

Règles dans Informatica Analyst

Ce chapitre comprend les rubriques suivantes :

- [Présentation des règles dans Informatica Analyst, 29](#)
- [Règles prédéfinies, 30](#)
- [Règles d'expression, 31](#)

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 6

Résultats de profil de colonne dans Informatica Analyst

Ce chapitre comprend les rubriques suivantes :

- [Présentation des résultats de profil de colonne dans Informatica Analyst, 34](#)
- [Vue de résumé, 35](#)
- [Vue détaillée, 37](#)
- [Statistiques, 39](#)
- [Types d'exécution de profil, 46](#)
- [Présentation des résultats de la comparaison de plusieurs profils, 48](#)
- [Développement de profil de colonne, 53](#)
- [Conservation dans l'outil Analyst, 54](#)
- [Fichiers d'exportation de profil de colonne dans Informatica Analyst, 55](#)

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.

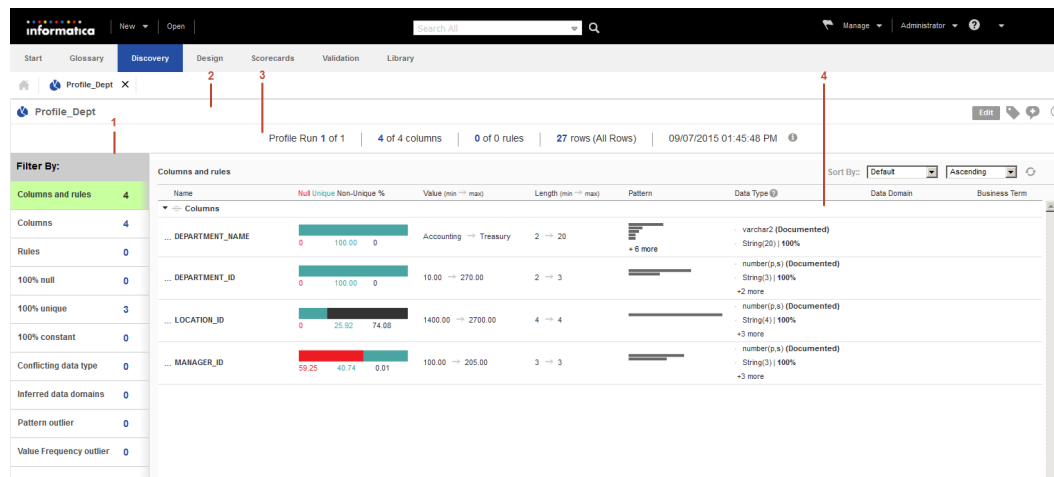
Remarque: Vous pouvez afficher et exécuter un profil sur des sources de données Avro, JSON, Parquet et XML. Vous pouvez afficher les résultats de profil de la dernière exécution, d'une exécution historique ou d'une exécution consolidée et comparer les résultats de deux exécutions de profils.

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 :

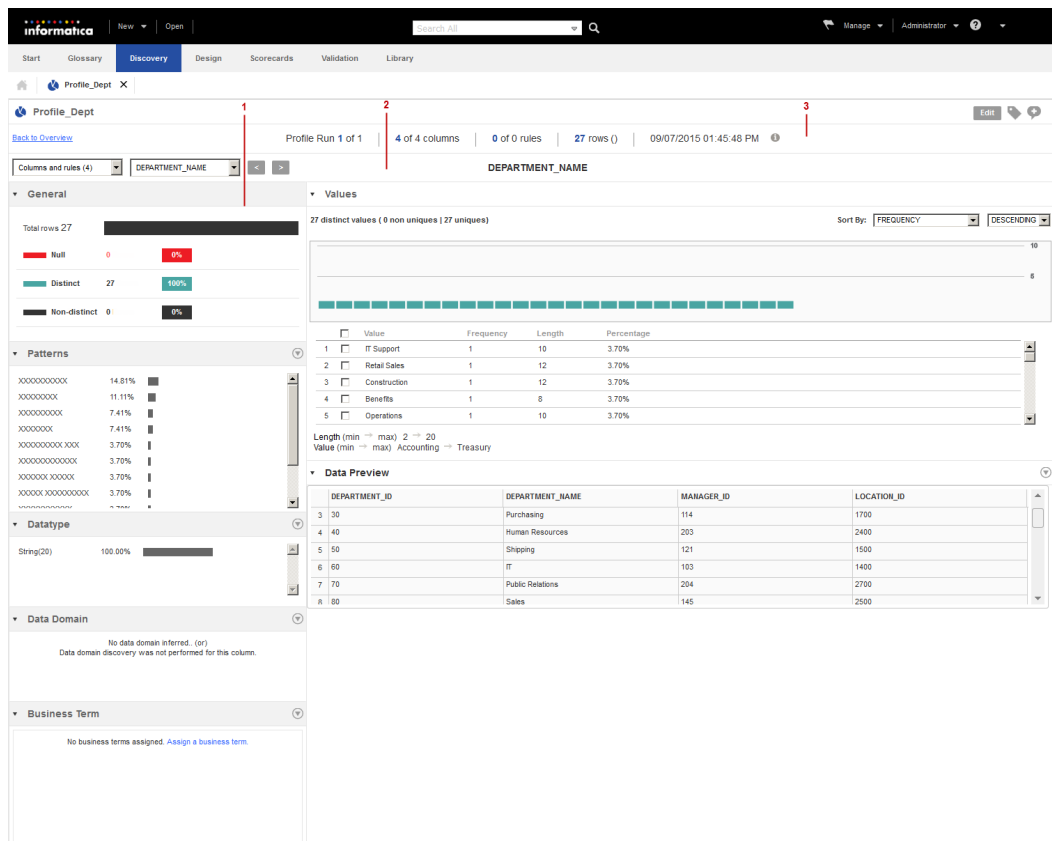
Option de filtre par défaut	Description
Colonnes et règles	Affiche les résultats de profil des colonnes source et de règle. Vous pouvez développer et réduire les colonnes source et de règle pour afficher les résultats.
Colonnes	Affiche les résultats de profil des colonnes source.
Règles	Affiche les résultats 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.

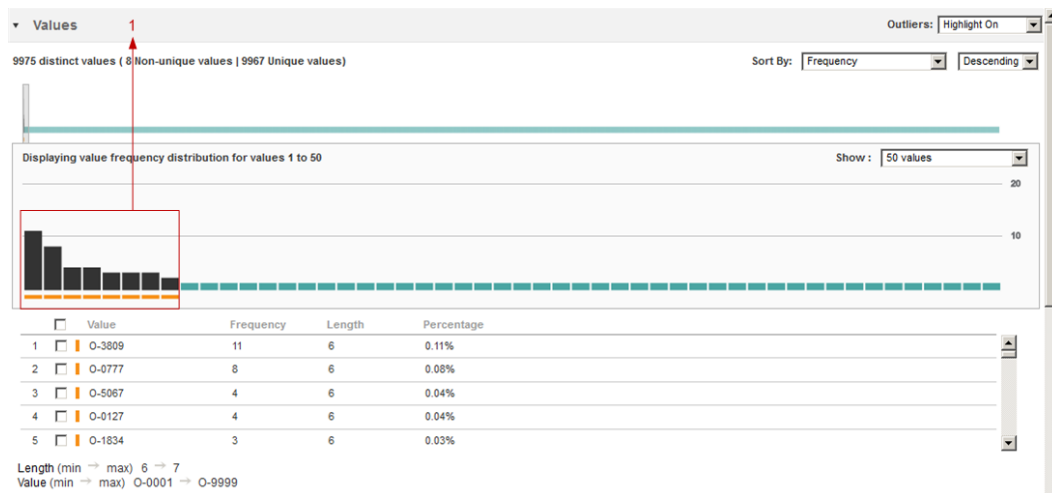
Valeurs hors norme

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

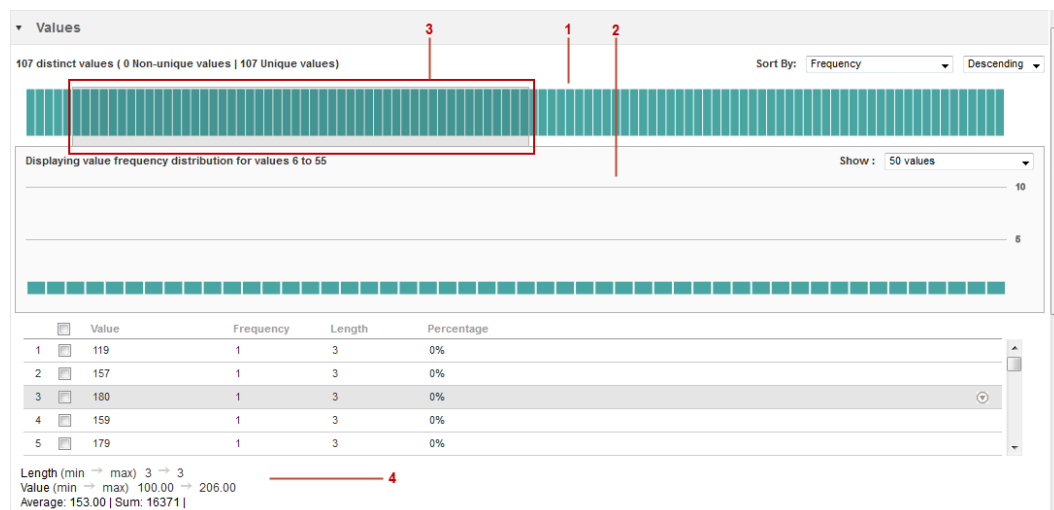
Les valeurs de colonne dans la vue détaillée comprennent les valeurs d'une colonne et la fréquence à laquelle elles s'affichent dans la colonne.

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.

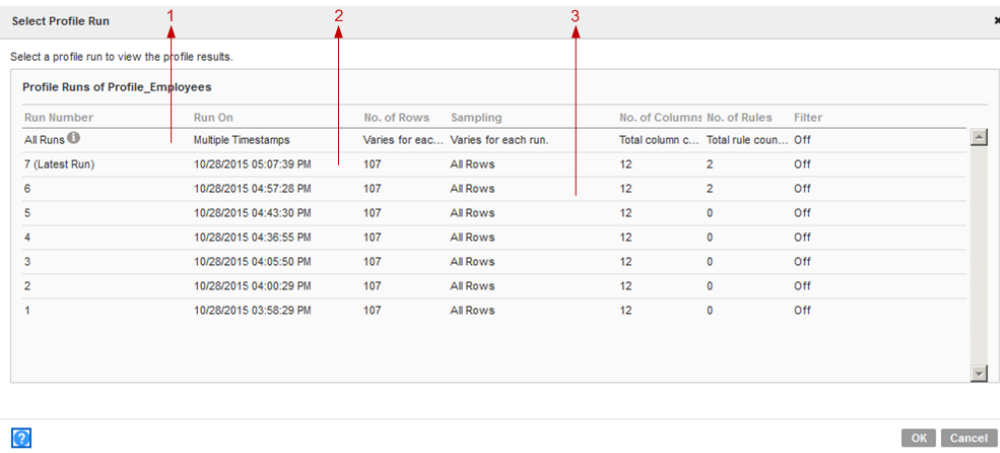
Sélection d'une exécution de profil

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**.

Compare Profile Runs

Select the profile runs that you want to compare.

Select a profile run

Run Number	Run On	No. of Rows Pr	Sampling	No. of Columns	No. of Rules Pr	Filter
3 (Latest Run)	09/09/2015 02:47:10 PM	27	No Sampling	4	0	Off
2	09/09/2015 02:45:24 PM	27	No Sampling	4	0	Off
1	09/07/2015 01:45:48 PM	27	No Sampling	4	0	Off

Select another profile run

Run Number	Run On	No. of Rows Pr	Sampling	No. of Columns	No. of Rules Pr	Filter
3 (Latest Run)	09/09/2015 02:47:10 PM	27	No Sampling	4	0	Off
2	09/09/2015 02:45:24 PM	27	No Sampling	4	0	Off
1	09/07/2015 01:45:48 PM	27	No Sampling	4	0	Off

OK Cancel

- 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.

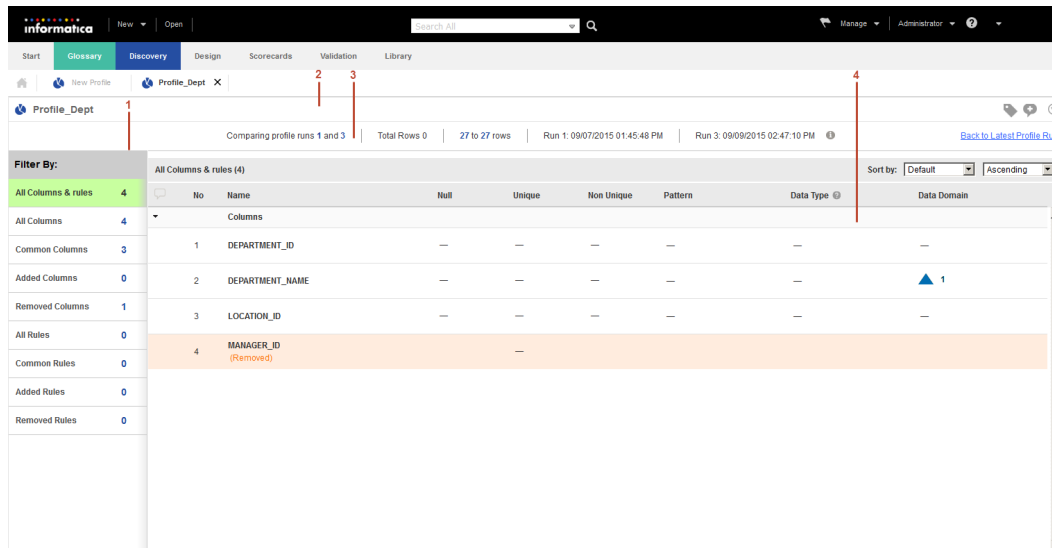
- Sélectionnez un premier profil dans le volet de l'**exécution A**, puis un deuxième dans le volet de l'**exécution B**.
- Cliquez sur **OK**.

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.

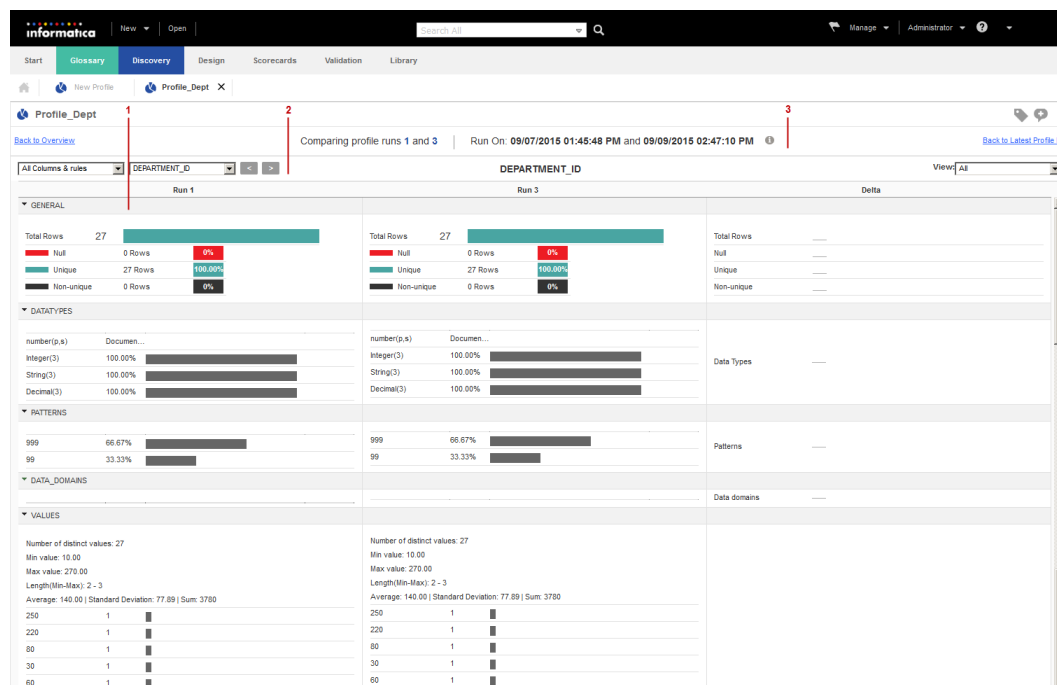
Par défaut, la vue de résumé affiche les résultats de profil pour toutes les colonnes source et virtuelles.

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 choisissiez 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 7

Fiches d'évaluation dans Informatica Analyst

Ce chapitre comprend les rubriques suivantes :

- [Présentation des fiches d'évaluation dans Informatica Analyst, 58](#)
- [Processus de fiche d'évaluation de Informatica Analyst, 59](#)
- [Création d'une fiche d'évaluation dans Informatica Analyst, 60](#)
- [Ajouter des colonnes à une fiche d'évaluation existante, 61](#)
- [Exécution d'une fiche d'évaluation, 63](#)
- [Affichage d'une fiche d'évaluation, 63](#)
- [Modification d'une fiche d'évaluation, 64](#)
- [Mesures, 65](#)
- [Groupes de mesures, 66](#)
- [Développement dans les colonnes, 68](#)
- [Graphes de tendance, 68](#)
- [Fichier d'exportation de fiche d'évaluation dans Informatica Analyst, 72](#)
- [Notifications de la fiche d'évaluation, 73](#)
- [Lignage des fiches d'évaluation, 76](#)

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 une 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.

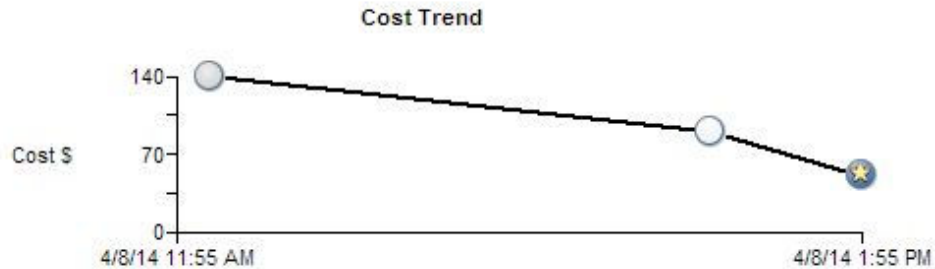
Graphe de tendance de coût

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

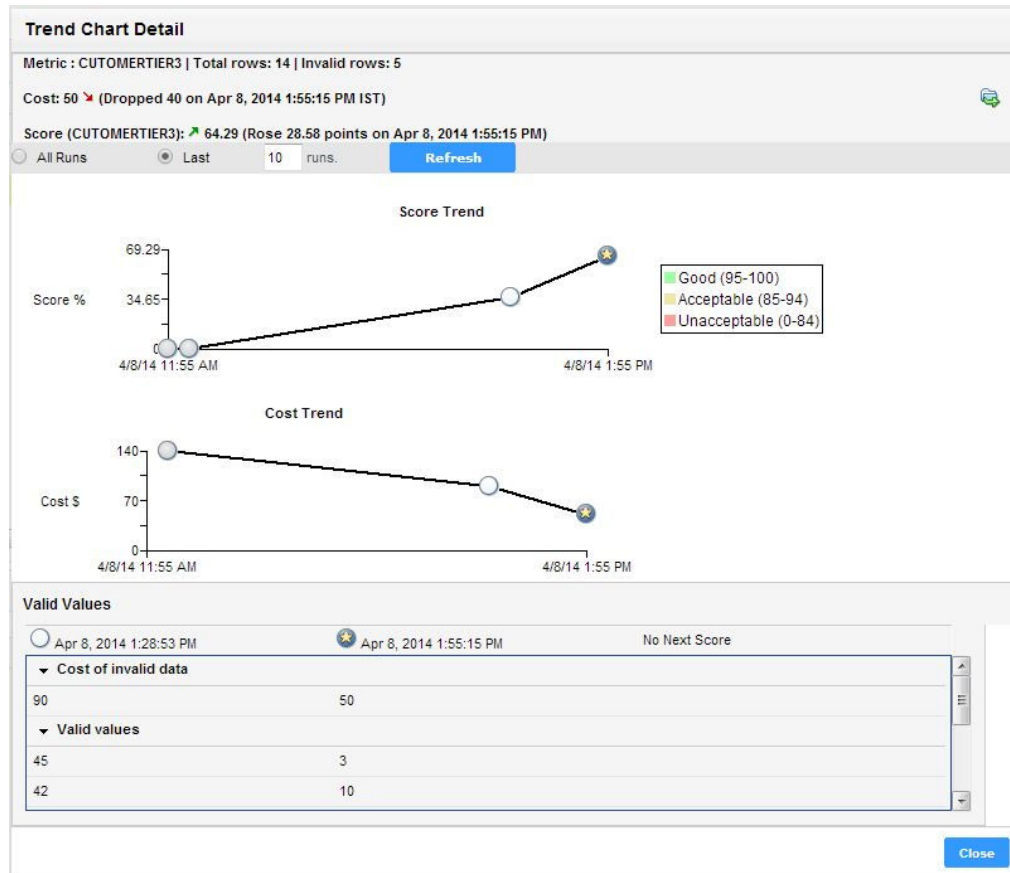
Vous pouvez afficher des graphes de tendance pour chaque mesure afin de surveiller la manière selon laquelle évolue le score ou le coût des données non valides dans le temps.

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 graphique de tendance**.
La boîte de dialogue **Détail de graphique de tendance** s'affiche.
3. Cliquez sur l'icône **Exporter les données**.
La boîte de dialogue **Exporter les données vers un fichier** s'ouvre.
4. Entrez un nom de fichier. Utilisez éventuellement le nom de fichier par défaut.
Le format de fichier par défaut est Microsoft Excel.
5. Sélectionnez la page de code du fichier.

6. Cliquez sur **OK**.

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

Vous pouvez exporter les résultats d'une fiche d'évaluation vers un fichier Microsoft Excel. L'outil Analyst 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 99](#).

Partie III : Profilage avec Informatica Developer

Cette partie contient les chapitres suivants :

- [Profils d'objet de données, 79](#)
- [Profils de colonne dans des sources de données semi-structurées, 83](#)
- [Règles dans Informatica Developer, 90](#)
- [Résultats de profil de colonne dans Informatica Developer, 92](#)
- [Fiches d'évaluation dans Informatica Developer, 98](#)
- [Profilage de mapplet et de mappage, 101](#)

CHAPITRE 8

Profils d'objet de données

Ce chapitre comprend les rubriques suivantes :

- [Profils de colonne dans Informatica Developer, 79](#)
- [Création d'un profil d'objet de données unique dans Informatica Developer, 80](#)
- [Création de plusieurs profils d'objet de données dans Informatica Developer, 81](#)
- [Synchronisation d'un objet de données de fichier plat, 82](#)
- [Synchronisation d'un objet de données relationnel, 82](#)

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.

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.

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**.

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.

CHAPITRE 9

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, 83](#)
- [Profil de colonne dans une source de données XML ou JSON, 84](#)
- [Profil de colonne dans une source de données semi-structurée dans HDFS, 85](#)
- [Création d'une connexion HDFS, 86](#)
- [Création d'un objet de données de fichier complexe sur un fichier JSON ou XML dans HDFS, 86](#)
- [Création d'un objet de données sur une source de données Avro ou Parquet, 87](#)
- [Exécution d'un profil de colonne sur des sources de données semi-structurées, 88](#)

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 10

Règles dans Informatica Developer

Ce chapitre comprend les rubriques suivantes :

- [Présentation des règles dans Informatica Developer](#)[Lignes directrices relatives aux règles, 90](#)
- [Création d'une règle dans Informatica Developer, 91](#)
- [Application d'une règle dans Informatica Developer](#)[Application d'une règle, 91](#)

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 11

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, 92](#)
- [Propriétés des valeurs de colonne, 93](#)
- [Propriétés des formes de colonne, 94](#)
- [Propriétés des statistiques de colonne, 94](#)
- [Propriétés des types de données de colonne, 95](#)
- [Conservation dans Informatica Developer](#)[Conservation dans Informatica Developer, 96](#)
- [Exportation des résultats de profil depuis Informatica Developer, 97](#)

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)	Minneapolis	Ann Arbor	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 States	United States	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 12

Fiches d'évaluation dans Informatica Developer

Ce chapitre comprend les rubriques suivantes :

- [Présentation des fiches d'évaluation dans Informatica Developer, 98](#)
- [Création d'une fiche d'évaluation, 98](#)
- [Exportation d'un fichier de ressource pour le lignage des fiches d'évaluation, 99](#)
- [Affichage du lignage des fiches d'évaluation dans Informatica Developer, 99](#)

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 13

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, 101](#)
- [Exécution d'un profil sur un mapplet ou un objet de mappage, 101](#)
- [Comparaison des profils pour les objets de mappage ou de mapplet, 102](#)
- [Génération d'un mappage à partir d'un profil, 102](#)

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.

INDEX

C

conservation
 concepts [17](#)
 Informatica Analyst [54](#)
 Informatica Developer [96](#)
 tâches [18](#)
création d'un profil de colonne
 profils [24](#)
création d'une règle d'expression
 règles [32](#)

E

exporter
 lignage des fiches d'évaluation vers XML [99](#)

F

fiche d'évaluation
 configuration des notifications [75](#)
 configuration des paramètres généraux des notifications [75](#)
fiches d'évaluation
 affichage [63](#)
 ajout de colonnes à une fiche d'évaluation [62](#)
 coût des données non valides [65](#)
 coût fixe [65](#)
 coût variable [65](#)
 création d'un groupe de mesures [66](#)
 définition des seuils [66](#)
 déplacement des scores [67](#)
 développement [68](#)
 exécution [63](#)
 graphe de tendance [68](#)
 Groupes de mesures [66](#)
 Informatica Analyst [58](#)
 Informatica Developer [98](#)
 mesures [65](#)
 modification [64](#)
 modification d'un groupe de mesures [67](#)
 notifications [73](#)
 poids métriques [65](#)
 présentation [16](#)
 processus de Informatica Analyst [59](#)
 Suppression d'un groupe de mesures [68](#)

G

graphes de tendance
 affichage [70](#)
 coût [69](#)
 exportation depuis Informatica Analyst [71](#)
 score [69](#)

I

Informatica Analyst
 présentation des profils de colonne [20](#), [48](#)
 règles [29](#)
 résultats de profil de colonne [34](#), [46](#)
 verrouillage et gestion de version [24](#)
Informatica Developer
 règles [90](#)

L

lignage des fiches d'évaluation
 affichage dans Informatica Analyst [76](#)
 affichage dans Informatica Developer [99](#)

O

objet de données fichier plat
 synchronisation [27](#)
objet de données table
 synchronisation [28](#)
objet de mappage
 exécution d'un profil [101](#)

P

profil
 composants [12](#)
 Formats Avro ou Parquet [85](#)
 formats XML et JSON [84](#), [85](#)
profil de colonne
 exploration [53](#)
 Informatica Developer [79](#)
 options [15](#)
 présentation [14](#)
 processus [21](#)
 profil du système d'exploitation [23](#)
profilage
 outils [12](#)
 processus [12](#)
 verrouillage et gestion de version [15](#)
Profilage de mapplet et de mappage
 Présentation [101](#)
profils
 création d'un profil de colonne [24](#)
 exécution [26](#), [47](#), [48](#)
 introduction [11](#)
 modification d'un profil de colonne [26](#)
profils d'objet de données
 création d'un profil unique [80](#)
 création de plusieurs profils [81](#)

Profils de mapplet et de mappage

Présentation [101](#)

R

règles

application d'une règle prédéfinie [30](#)

application dans Informatica Developer [91](#)

application dans PowerCenter Express [91](#)

création d'une règle d'expression [32](#)

création dans Informatica Developer [91](#)

expression [31](#)

prédéfinies [30](#)

prérequis [90](#)

règles d'expression

processus [31](#)

règles prédéfinies

processus [30](#)

résultats de fiche d'évaluation

exportation [72](#)

exportation depuis Informatica Analyst [72](#)

exporter vers Excel [72](#)

résultats de profil

approbation de types de données [54](#)

approbation des types de données dans Informatica Developer [96](#)

développement [53](#)

Excel [56](#)

résultats de profil (*a continué*)

exportation [55](#)

exportation dans Informatica Developer [97](#)

exportation depuis Informatica Analyst [56](#)

formes de colonne [42](#)

rejet de types de données [55](#)

rejet de types de données dans l'outil Developer tool [96](#)

résumé [50](#), [53](#)

types de données de colonne [40](#), [95](#)

valeurs de colonne [43](#)

vue de résumé [36](#)

vue détaillée [38](#)

résultats de profil de colonne

Informatica Developer [92](#)

profil de colonne [92](#)

résultats de profil de colonne dans l'outil Analyst tool

interface [37](#), [51](#)

résultats de profil de colonne dans l'outil Analyst tool

détails de colonne [37](#), [52](#)

interface [49](#)

résumé [35](#)

V

valeur hors norme

détection [42](#)