



Informatica®
10.4.0

Handbuch für Datenerkennung

Diese Software und die Dokumentation werden nur im Rahmen eines eigenen Lizenzvertrags zur Verfügung gestellt, der Beschränkungen für die Verwendung und Weitergabe enthält. Ohne ausdrückliche schriftliche Genehmigung der Informatica LLC darf kein Teil dieses Dokuments zu irgendeinem Zweck vervielfältigt oder übertragen werden, unabhängig davon, auf welche Art und Weise oder mit welchen Mitteln (elektronisch, mechanisch, durch Fotokopieren, Aufzeichnen usw.) dies geschieht.

Informatica und das Informatica-Logo sind Marken oder eingetragene Marken der Informatica LLC in den Vereinigten Staaten von Amerika und zahlreichen anderen Ländern der Welt. Eine aktuelle Liste der Informatica-Marken ist im Internet auf <https://www.informatica.com/trademarks.html> verfügbar. Alle weiteren Produkt- und Firmennamen sind möglicherweise Markennamen oder Warenzeichen der jeweiligen Eigentümer.

Den RECHTEN DER REGIERUNG DER VEREINIGTEN STAATEN unterliegende Programme, Software, Datenbanken und zugehörige Dokumentation und technische Daten, die an Kunden der Regierung der Vereinigten Staaten geliefert werden, sind "kommerzielle Computersoftware" oder "kommerzielle technische Daten" gemäß der anwendbaren Beschaffungsverordnung der Vereinigten Staaten (Federal Acquisition Regulation – FAR) und der ergänzenden Bestimmungen der spezifischen Behörde. Damit unterliegen die Nutzung, das Kopieren, die Offenlegung, das Modifizieren und die Anpassung den im anwendbaren Regierungsvertrag gemachten Einschränkungen und Lizenzbedingungen und, soweit im Rahmen der Bedingungen des Regierungsvertrags und der in FAR 52.227-19 aufgeführten Rechte anwendbar, der Lizenz für die kommerzielle Computersoftware.

Teile dieser Software und/oder Dokumentationen unterliegen dem Urheberrecht Dritter. Die erforderlichen Hinweise auf Drittanbieter sind im Lieferumfang des Produkts enthalten.

Die in dieser Dokumentation enthaltenen Informationen können jederzeit ohne vorherige Ankündigung geändert werden. Wenn Sie Probleme in dieser Dokumentation finden, melden Sie sie uns unter infa_documentation@Informatica.com.

Informatica-Produkte unterliegen einer Gewährleistung gemäß den Geschäftsbedingungen der Vereinbarungen, unter denen sie bereitgestellt werden. INFORMATICA STELLT DIE INFORMATIONEN IN DIESEM DOKUMENT OHNE MÄNGELGEWÄHR UND OHNE AUSDRÜCKLICHE ODER STILLSCHWEIGENDE GEWÄHRLEISTUNG JEDLICHER ART ZUR VERFÜGUNG. DIES GILT EINSCHLIESSLICH FÜR GEWÄHRLEISTUNGEN DER MARKTGÄNGIGKEIT, DER EIGNUNG FÜR EINEN BESTIMMTEN ZWECK UND GEWÄHRLEISTUNGEN ODER ZUSICHERUNGEN ÜBER DIE NICHTVERLETZUNG VON RECHTEN DRITTER.

Inhalt

Einleitung	13
Informatica-Ressourcen.	13
Informatica-Netzwerk.	13
Informatica-Wissensdatenbank.	13
Informatica-Dokumentation.	14
Informatica-Produktverfügbarkeitsmatrizen.	14
Informatica Velocity.	14
Informatica Marketplace.	14
Globaler Kundensupport von Informatica.	14
 Teil I: Einführung in Data Discovery.....	15
 Kapitel 1: Einführung in Profiling.	16
Profiling - Übersicht.	16
Profiling-Architektur.	17
Data Discovery-Prozess.	19
 Kapitel 2: Data Discovery.....	21
Data Discovery - Übersicht.	21
Profil- und Analysetypen.	21
Profiling-Komponenten.	22
Profilergebnisse.	23
 Kapitel 3: Spaltenprofilkonzepte.....	25
Spaltenprofilkonzepte - Übersicht.	25
Spaltenprofiloptionen.	26
Repository-Profilsperrern und Verwaltung versionierter Profile.	26
Scorecards.	27
 Kapitel 4: Konzepte der Datendomänenerkennung.....	28
Konzepte der Datendomänenerkennung - Übersicht.	28
Datendomänen.	29
Datendomänenengruppen.	29
Datendomänenglossar.	30
Datendomänenenerkennungsprozess.	30
Datendomänenerkennung auf der Spark-Engine.	31
 Kapitel 5: Wiederherstellungskonzepte.....	32
Wiederherstellungskonzepte - Übersicht.	32
Wiederherstellung für Analysten und Entwickler.	32

Wiederherstellungsaufgaben.	33
Teil II: Data Discovery mit Informatica Analyst.	35
Kapitel 6: Spaltenprofile in Informatica Analyst.	36
Spaltenprofile in Informatica Analyst - Übersicht.	36
Spalten-Profilingprozess.	37
Profil-Optionen.	37
Stichprobenoptionen.	38
Drilldown-Optionen.	39
Laufzeitumgebung.	39
Native Umgebung.	39
Hadoop-Umgebung.	39
Betriebssystemprofile in Informatica Analyst - Übersicht.	40
Auswählen eines Betriebssystemprofils.	41
Repository-Objektsperren und teambasierte Entwicklung - Übersicht.	41
Erstellen eines Spaltenprofils in Informatica Analyst.	41
Bearbeiten eines Spaltenprofils.	43
Ausführen eines Profils.	44
Ausführen eines Profils auf der Spark-Engine.	44
Synchronisierungsoption.	44
Synchronisieren eines Einfachdatei-Datenobjekts in Informatica Analyst.	45
Synchronisieren eines relationalen Datenobjekts in Informatica Analyst.	47
Kapitel 7: Regeln in Informatica Analyst.	48
Regeln in Informatica Analyst - Übersicht.	48
Vordefinierte Regeln.	49
Vordefinierte Regeln-Prozess.	49
Anwenden einer vordefinierten Regel.	49
Ausdrucksregeln.	50
Erstellen einer Ausdrucksregel.	50
Erstellen einer Ausdrucksregel unter Verwendung der Regelspezifikation.	52
Kapitel 8: Filter in Informatica Analyst.	53
Filter in Informatica Analyst - Übersicht.	53
Erstellen eines Filters.	53
Erstellen eines einfachen Filters.	54
Erstellen eines erweiterten Filters.	55
Erstellen eines SQL-Filters.	56
Verwalten von Filtern.	56
Kapitel 9: Spaltenprofilerggebnisse in Informatica Analyst.	58
Spaltenprofilerggebnisse in Informatica Analyst - Übersicht.	58

Zusammenfassungsansicht.	59
Eigenschaften der Zusammenfassungsansicht.	60
Standardfilter in der Zusammenfassungsansicht.	61
Detailansicht.	61
Bereiche der Detailansicht.	62
Statistiken.	63
Datenvorschau.	64
Datentypen.	64
Sonderfälle.	65
Muster.	66
Werte.	67
Typen von Profilausführungen.	70
Aktuelle Profilausführung.	70
Historische Profilausführung.	70
Zusammengefasste Profilausführung.	71
Auswählen einer Profilausführung.	71
Vergleichen mehrerer Profilergebnisse - Übersicht.	72
Vergleichen mehrerer Profilergebnisse.	72
Zusammenfassungsansicht für den Profilergebnisvergleich	73
Detailansicht des Profilergebnisvergleichs	76
Spaltenprofil-Drilldown.	77
Drilldown an Zeilendaten.	77
Filter auf Drilldown-Daten anwenden.	77
Wiederherstellung im Analyst-Tool.	78
Genehmigen von Datentypen und Datendomänen.	78
Ablehnen von Datentypen und Datendomänen.	79
Exportdateien des Spaltenprofils in Informatica Analyst.	79
Profilieren von Exportergebnissen in eine CSV-Datei.	79
Profilieren von Exportergebnissen in Microsoft Excel.	80
Exportieren von Profilergebnissen aus Informatica Analyst.	80

Kapitel 10: Geschäftsbegriffe, Kommentare und Tags in Informatica Analyst. 82

Geschäftsbegriffe, Kommentare und Tags in Informatica Analyst - Übersicht.	82
Geschäftsbegriffe.	82
Zuweisen von Geschäftsbegriffen zu Spalten.	83
Kommentare.	83
Hinzufügen von Kommentaren zu einem Profil oder Spalten.	83
Tags.	84
Zuweisen von Tags zu einem Profil oder Spalten.	84

Kapitel 11: Scorecards in Informatica Analyst. 85

Scorecards in Informatica Analyst - Übersicht.	85
--	----

Informatica Analyst Scorecard-Prozess.	86
Erstellen einer Scorecard in Informatica Analyst.	87
Hinzufügen von Spalten zu vorhandenen Scorecards.	88
Hinzufügen von Spalten zu einer vorhandenen Scorecard.	89
Ausführen einer Scorecard.	90
Anzeigen einer Scorecard.	90
Bearbeiten einer Scorecard.	91
Maße.	92
Metrikgewichtungen.	92
Wert der Datenqualität.	92
Definieren von Schwellenwerten.	93
Metrikgruppen.	93
Erstellen einer Metrikgruppe.	93
Verschieben von Scores in eine Metrikgruppe.	94
Bearbeiten einer Metrikgruppe.	94
Löschen einer Metrikgruppe.	95
Spalten-Drilldown.	95
Trenddiagramme.	95
Score-basiertes Trenddiagramm.	96
Kostenbasiertes Trenddiagramm.	96
Anzeigen von Trenddiagrammen.	97
Exportieren von Trenddiagrammen.	98
Scorecard-Dashboard in Informatica Analyst.	99
Scorecards nach Projekt.	100
Trend bei der Scorecard-Ausführung.	101
Datenobjekte mit Scorecards.	102
Trend kumulativer Metriken.	103
Scorecard-Exportdateien in Informatica Analyst.	104
Exportieren von Scorecard-Ergebnissen aus Informatica Analyst.	105
Scorecard-Exportergebnisse in Microsoft Excel.	106
Scorecard-Benachrichtigungen.	106
Vorlage für Benachrichtigungs-E-Mail.	107
Einrichten der Scorecard-Benachrichtigungen.	108
Konfigurieren von globalen Einstellungen für Scorecard-Benachrichtigungen.	108
Scorecard-Herkunft.	109
Anzeigen der Scorecard-Herkunft in Informatica Analyst.	110
Kapitel 12: Datenqualitätsfunktionen in Informatica Analyst.	111
Datendomänenerkennung in Informatica Analyst - Übersicht.	111
Datendomänglossar in Informatica Analyst.	111
Erstellen einer Datendomänengruppe in Informatica Analyst.	112
Erstellen einer Datendomäne in Informatica Analyst.	112
Erstellen einer Datendomäne aus Profilergebnissen in Informatica Analyst.	113

Suchen von Datendomänengruppen in Informatica Analyst.	113
Optionen der Datendomänenerkennung in Informatica Analyst.	114
Datendomänen-Spaltenauswahl in Informatica Analyst.	114
Datendomänenauswahl in Informatica Analyst.	115
Ableitungsoptionen für Datendomänen in Informatica Analyst.	115
Erstellen eines Spaltenprofils zur Ausführung der Datendomänenerkennung in Informatica Analyst	117
Bearbeiten eines Spaltenprofils und einer Datendomänenerkennung in Informatica Analyst.	118
Ausführen eines Profils zur Durchführung von Datendomänenerkennung.	119
Datendomänenerkennungsergebnisse in Informatica Analyst.	119
Genehmigen von Datendomänen.	120
Ablehnen von Datendomänen.	120
Datendomänenerkennungs-Exportdateien in Informatica Analyst.	121
Datendomänenerkennungsergebnisse in Microsoft Excel.	121
Exportieren von Datendomänenerkennungsergebnissen aus Informatica Analyst.	121
Kapitel 13: Enterprise-Erkennung in Informatica Analyst.	122
Enterprise-Erkennung in Informatica Analyst - Übersicht.	122
Enterprise-Erkennungsprozess in Informatica Analyst.	123
Konfigurationsoptionen für Enterprise-Erkennung.	123
Einstellungen für Datendomänenerkennung.	123
Spaltenprofileinstellungen.	124
Erstellen eines Enterprise-Erkennungsprofils in Informatica Analyst.	125
Bearbeiten der Enterprise-Erkennungsoptionen.	127
Kapitel 14: Enterprise-Erkennungsergebnisse in Informatica Analyst.	128
Ergebnisse der Enterprise-Erkennung in Informatica Analyst - Übersicht.	128
Zusammenfassungsansicht.	128
Profilergebnisse Zusammenfassungsansicht.	129
Anzeigen von Datendomänenerkennungsergebnissen.	130
Anzeigen von Spaltenprofilergebnissen.	130
Datentypkonflikt.	131
Anzeigen von Datentypkonflikten.	131
Profilansicht.	131
Anzeigen von Profileigenschaften.	132
Kapitel 15: Erkennungssuche in Informatica Analyst.	133
Erkennungssuche in Informatica Analyst – Übersicht.	133
Voraussetzungen der Erkennungssuche.	134
Prozess der Erkennungssuche in Informatica Analyst.	134
Optionen für die Erkennungssuche.	135
Kriterien der Erkennungssuche.	136
Suchen nach einem Objekt.	136
Ergebnisse der Erkennungssuche in Informatica Analyst.	136

Bereich für die Erkennungssuchergebnisse.	137
Filtern von Ergebnissen der Erkennungssuche.	138
Match-Typen.	138
Direkter Match.	138
Indirekte Übereinstimmung.	139
Anzeigen der Match-Informationen.	139
Öffnen von Objekten über die Ergebnisse der Erkennungssuche.	139
Zugehörige Ressourcen.	140
Verbundene Objekte für jeden Objekttyp.	140
Anzeigen verbundener Objekte.	141
Häufig gestellte Fragen (FAQ).	141
Kapitel 16: Business Glossary Desktop in Informatica Analyst.	143
Geschäftsbegriffe.	143
Geschäftsbegriffe verwalten im Unternehmensglossar des Metadata Manager.	144
Suchen eines Geschäftsbegriffs im Unternehmensglossar-Desktop.	144
Teil III: Data Discovery mit Informatica Developer.	145
Kapitel 17: Informatica Developer-Profile.	146
Informatica Developer-Profile - Übersicht.	146
Informatica Developer-Profilansichten.	147
Repository-Objektsperren und teambasierte Entwicklung mit versionierten Objekten.	148
Kapitel 18: Datenobjektprofile.	149
Datenobjektprofile - Übersicht.	149
Spaltenprofile in Informatica Developer.	150
Filteroptionen.	151
Stichprobenoptionen.	152
Laufzeitumgebung.	153
Native Umgebung.	153
Hadoop-Umgebung.	153
Primärschlüssel-Erkennung.	154
Rückschlusseigenschaften von Primärschlüsseln.	155
Gefolgerte Primärschlüsseigenschaften.	155
Schlüsselverstoß-Eigenschaften.	156
Funktionale Abhängigkeitserkennung.	156
Funktionaler Abhängigkeitsrückschluss - Eigenschaften.	156
Gefolgerte funktionale Abhängigkeit - Eigenschaften.	157
Funktionale Abhängigkeitsverstöße - Eigenschaften.	157
Betriebssystemprofile in Informatica Developer.	158
Auswählen eines Betriebssystemprofils.	158
Erstellen eines einzelnen Datenobjektprofils in Informatica Developer.	158

Erstellen mehrerer Datenobjektprofile in Informatica Developer.	159
Bearbeiten eines Profils.	160
Synchronisierungsoption.	160
Synchronisieren eines Einfachdatei-Datenobjekts in Informatica Developer.	160
Synchronisieren eines relationalen Datenobjekts in Informatica Developer.	162
Kommentare.	163
Hinzufügen von Kommentaren in Informatica Developer.	163
Kapitel 19: Spaltenprofile in halbstrukturierten Datenquellen.	164
Spaltenprofile in halbstrukturierten Datenquellen - Übersicht.	164
JSON- und XML-Datenobjekte.	165
Erstellen eines Datenobjekts aus einer JSON- oder XML-Datenquelle.	165
Komplexe Dateidatenobjekte für halbstrukturierte Datenquellen in HDFS.	166
Komplexes Dateidatenobjekt aus einer JSON- oder XML-Datenquelle in HDFS.	166
Komplexes Dateidatenobjekt aus einer Avro- oder Parquet-Datenquelle in HDFS.	166
Erstellen einer HDFS-Verbindung.	167
Erstellen eines komplexen Dateidatenobjekts anhand einer JSON- oder XML-Datei in HDFS.	167
Erstellen eines komplexen Dateidatenobjekts aus einer Avro- oder Parquet-Datenquelle.	168
Erstellen eines Spaltenprofils in einer halbstrukturierten Datenquelle.	169
Kapitel 20: Regeln in Informatica Developer.	171
Regeln in Informatica Developer - Übersicht.	171
Erstellen einer Regel in Informatica Developer.	172
Anwenden einer Regel in Informatica Developer.	172
Kapitel 21: Mapplet- und Mapping-Profilings.	173
Mapplet- und Mapping-Profilings - Übersicht.	173
Ausführen eines Profils für ein Mapplet- oder Mapping-Objekt.	173
Vergleichen von Profilen für Mapping- oder Mapplet-Objekte.	174
Generieren eines Mapping aus einem Profil.	174
Kapitel 22: Spaltenprofilerggebnisse in Informatica Developer.	176
Spaltenprofilerggebnisse in Informatica Developer.	176
Spaltenwert-Eigenschaften.	177
Spaltenmuster-Eigenschaften.	177
Spaltenstatistik-Eigenschaften.	177
Eigenschaften des Spaltendatentyps.	178
Wiederherstellung in Informatica Developer.	179
Genehmigen von Datentypen.	179
Ablehnen von Datentypen.	179
Exportieren von Profilergebnissen aus Informatica Developer.	180

Kapitel 23: Scorecards in Informatica Developer.....	181
Scorecards in Informatica Developer - Übersicht.	181
Erstellen einer Scorecard.	181
Exportieren einer Ressourcendatei zur Scorecard-Herkunftskontrolle.	182
Anzeigen der Scorecard-Herkunft aus Informatica Developer.	182
 Kapitel 24: Datendomänenerkennung in Informatica Developer.....	 184
Datendomänenerkennung in Informatica Developer - Übersicht.	184
Datendomänglossar in Informatica Developer.	185
Erstellen einer Datendomängengruppe in Informatica Developer.	185
Erstellen einer Datendomäne in Informatica Developer.	185
Erstellen einer Datendomäne aus Profilergebnissen in Informatica Developer.	186
Suchen von Datendomänen in Informatica Developer.	187
Importieren von Datendomänen.	187
Exportieren von Datendomänen.	188
Optionen der Datendomänenerkennung in Informatica Developer.	189
Datenqualitäts-Funktionen in Informatica Developer.	189
Datendomänen-Spaltenauswahl in Informatica Developer.	189
Ableitungsoptionen für Datendomänen in Informatica Developer.	190
Erstellen eines Profils zur Ausführung der Datendomänenerkennung in Informatica Developer. . .	191
Bearbeiten eines Profils in Informatica Developer.	192
Ausführen eines Profils zur Ausführung der Datendomänenerkennung in Informatica Developer. .	193
Datendomänenerkennungsergebnisse in Informatica Developer.	193
Anzeigen nach Datendomängengruppen.	194
Anzeigen nach Spalten.	194
Überprüfen der Ergebnisse.	194
Genehmigen von Datendomänen.	195
Ablehnen von Datendomänen.	195
Exportieren von Datendomänenerkennungsergebnissen aus Informatica Developer.	196
 Kapitel 25: Enterprise-Erkennung in Informatica Developer.....	 197
Enterprise-Erkennung in Informatica Developer - Übersicht.	197
Enterprise-Erkennungsprozess.	198
Profilooptionen für Enterprise-Erkennung.	199
Datendomänenauswahl für die Enterprise-Erkennung.	199
Stichprobenoptionen des Spaltenprofils für die Enterprise-Erkennung.	200
Laufzeitumgebung Option.	200
Ableitungsoptionen der Primärschlüssel für die Enterprise-Erkennung.	201
Ableitungsoptionen der Fremdschlüssel für die Enterprise-Erkennung.	201
Automatische Wiederherstellungsparameter für die Fremdschlüsselableitung.	202
Erstellen eines Enterprise-Erkennungsprofils in Informatica Developer.	204
Bearbeiten eines Profils.	205

Ausführen eines Enterprise-Erkennungsprofils.	206
Fremdschlüsselerkennung.	207
Definieren übergeordneter und Kind-Objektbeziehungen.	207
Erkennen von Fremdschlüsselbeziehungen zwischen Datenobjekten	207
Ergebnisse der Fremdschlüsselanalyse.	208
Join-Analyse.	208
Erstellen eines Join-Profils.	209
Ergebnisse der Join-Analyse.	209
Exportieren von Join-Profilergebnissen in eine Datei.	210
Überschneidungserkennung.	210
Ergebnisse der Überschneidungserkennung.	211
Erkennen von sich überschneidenden Daten.	212
DDL-Skriptdateien.	212
Erstellen von DDL-Skripts aus einem Enterprise-Erkennungsprofil.	212
Synchronisieren eines Enterprise-Erkennungsprofils.	213
Synchronisieren eines Enterprise-Erkennungsprofils.	213

Kapitel 26: Ergebnisse der Enterprise-Erkennung. 215

Enterprise-Erkennungsergebnisse - Übersicht.	215
Beziehungsansicht.	216
Suchen nach einem Datenobjekt.	216
Navigieren zur Fremdschlüsselprofilansicht.	217
Fremdschlüssel-Profiling-Ansicht.	217
Anzeigen von Datenobjektbeziehungen.	217
Vergrößern und Verkleinern der Ansicht.	218
Suchen nach einem Datenobjekt.	218
Anzeigen von Spaltenbeziehungen.	219
Speichern des Entitätsbeziehungsdiagramms als Symbol.	219
Anzeigen von Datenobjektprofilergebnissen aus der Ansicht Fremdschlüssel-Profiling heraus	219
Tabellenansicht.	220
Tabellendetailbereich.	220
Überprüfen der Enterprise-Erkennungsergebnisse.	220
Wiederherstellen von Spaltenbeziehungen.	221
Übergeben der Ergebnisse an das Modellrepository.	221
Datendomänenansicht.	222
Anzeigen von Datendomänenenerkennungsergebnissen.	222
Überprüfen der Datendomänenenerkennungsergebnisse.	222
Zeilen-Drilldown.	222
Anzeigen von Datenobjektprofilergebnissen aus der Datendomänenansicht.	223
Spaltenprofilansicht.	223
Anzeigen von Datenobjektprofilergebnissen.	223
Anzeigen von Spaltenprofilergebnissen während der Ausführung der Enterprise-Erkennung.	223

Anzeigen von Datendomänenenerkennungsergebnissen während der Ausführung der Enterprise-Erkennung.	224
Anzeigen des Laufzeitstatus der Enterprise-Erkennung.	224
Enterprise-Erkennungs-Exportdateien.	225
Exportieren der Enterprise-Erkennungsergebnisse.	225
Kapitel 27: Business Glossary Desktop in Informatica Developer.	226
Business Glossary-Suche.	226
Suchen einer Geschäftsbedingung.	227
Anpassen der Abkürzungstasten zum Nachschlagen einer Geschäftsbedingung.	227
Anhang A: Funktionsunterstützung auf Basis der Profiling-Warehouse-Verbindung.	228
Unterstützung der Profiling-Funktion.	228
Index.	230

Einleitung

Im *Handbuch zu Informatica Data Discovery* erfahren Sie, wie Sie Profile erstellen und ausführen können, um den Inhalt, die Qualität und die Struktur von Datenquellen zu analysieren. Erfahren Sie, wie Sie Data Discovery durchführen können, um die Metadaten von Quellsystemen zu ermitteln, die Beziehungen zwischen Spalten aus einer oder mehreren Datenquellen enthalten. Mit dem Developer Tool und dem Analyst Tool können Sie ein Spaltenprofil, ein Datendomänenerkennungsprofil oder ein Enterprise-Erkennungsprofil erstellen, verwalten und ausführen.

Das Handbuch ist für Datenanalysten und Entwickler vorgesehen.

Informatica-Ressourcen

Informatica stellt Ihnen über das Informatica-Netzwerk und andere Online-Portale zahlreiche Produktressourcen zur Verfügung. Nutzen Sie die Ressourcen, um Ihre Informatica-Produkte und -Lösungen optimal zu nutzen und von anderen Informatica-Benutzern und Fachspezialisten zu lernen.

Informatica-Netzwerk

Das Informatica-Netzwerk bietet Zugriff auf zahlreiche Ressourcen, darunter die Informatica-Wissensdatenbank und der globale Kundensupport von Informatica. Um auf das Informatica-Netzwerk zuzugreifen, besuchen Sie <https://network.informatica.com>.

Als Mitglied des Informatica-Netzwerks haben Sie die folgenden Optionen:

- Durchsuchen Sie die Wissensdatenbank nach Produktressourcen.
- Zeigen Sie Informationen zur Produktverfügbarkeit an.
- Erstellen und überprüfen Sie Ihre Supportfälle.
- Ihr lokales Informatica-Netzwerk für Benutzergruppen suchen und mit anderen Benutzern zusammenarbeiten.

Informatica-Wissensdatenbank

In der Informatica-Wissensdatenbank finden Sie Produktressourcen wie beispielsweise praktische Anleitungen, Best Practices, Videotutorials und Antworten auf häufig gestellte Fragen.

Zum Durchsuchen der Wissensdatenbank besuchen Sie <https://search.informatica.com>. Wenn Sie Fragen, Kommentare oder Ideen zur Wissensdatenbank haben, wenden Sie sich per E-Mail an das Team der Informatica-Wissensdatenbank unter KB_Feedback@informatica.com.

Informatica-Dokumentation

Verwenden Sie das Informatica-Dokumentationsportal, um in einer umfangreichen Dokumentationsbibliothek nach aktuellen und neuen Produktversionen zu suchen. Um das Dokumentationsportal zu erkunden, besuchen Sie <https://docs.informatica.com>

Wenn Sie Fragen, Kommentare oder Ideen zur Produktdokumentation haben, wenden Sie sich an das Informatica-Dokumentationsteam unter infa_documentation@informatica.com

Informatica-Produktverfügbarkeitsmatrizen

Produktverfügbarkeitsmatrizen (PAMs) geben die Versionen der Betriebssysteme, Datenbanken und Typen von Datenquellen und Zielen an, die in einer Produktversion unterstützt werden. Sie können die Informatica-PAMs unter <https://network.informatica.com/community/informatica-network/product-availability-matrices> durchsuchen.

Informatica Velocity

Informatica Velocity ist eine Sammlung von Tipps und Best Practices, die von den Professionellen Informatica-Diensten entwickelt wurden und auf praktischen Erfahrungen aus Hunderten von Datenmanagementprojekten basieren. Informatica Velocity umfasst das gesammelte Wissen von Informatica-Beratern, die mit Unternehmen auf der ganzen Welt zusammenarbeiten, um erfolgreiche Datenmanagementlösungen zu planen, zu entwickeln, bereitzustellen und zu warten.

Die Informatica Velocity-Ressourcen finden Sie unter <http://velocity.informatica.com>. Wenn Sie Fragen, Anregungen oder Ideen zu Informatica Velocity haben, wenden Sie sich an die professionellen Informatica-Dienste unter ips@informatica.com.

Informatica Marketplace

Informatica Marketplace ist ein Forum, das Lösungen zur Erweiterung und Verbesserung Ihrer Informatica-Implementierungen bereitstellt. Nutzen Sie die zahlreichen Lösungen von Informatica-Entwicklern und -Partnern im Marketplace, um Ihre Produktivität zu steigern und die Implementierungsdauer Ihrer Projekte zu verkürzen. Den Informatica Marketplace finden Sie unter <https://marketplace.informatica.com>.

Globaler Kundensupport von Informatica

Sie können sich telefonisch oder über das Informatica-Netzwerk an ein Global Support-Center wenden.

Die Telefonnummer des globalen Kundensupports von Informatica vor Ort finden Sie auf der Informatica-Website unter folgender Verknüpfung:

<https://www.informatica.com/services-and-training/customer-success-services/contact-us.html>.

Um im Informatica-Netzwerk nach Online-Supportressourcen zu suchen, besuchen Sie <https://network.informatica.com> und wählen Sie die eSupport-Option aus.

Teil I: Einführung in Data Discovery

Dieser Teil enthält die folgenden Kapitel:

- [Einführung in Profiling, 16](#)
- [Data Discovery, 21](#)
- [Spaltenprofilkonzepte, 25](#)
- [Konzepte der Datendomänenerkennung, 28](#)
- [Wiederherstellungskonzepte, 32](#)

KAPITEL 1

Einführung in Profiling

Dieses Kapitel umfasst die folgenden Themen:

- [Profiling - Übersicht, 16](#)
- [Profiling-Architektur, 17](#)
- [Data Discovery-Prozess, 19](#)

Profiling - Übersicht

Verwenden Sie das Profiling, um nach Inhalt, Qualität und Struktur von Datenquellen einer Anwendung, eines Schemas oder eines Unternehmens zu suchen. Der Inhalt der Datenquelle schließt Werthäufigkeiten und Datentypen ein. Die Struktur der Datenquelle schließt Schlüssel und funktionale Abhängigkeiten ein.

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

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

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

- Durchführen von Spalten-Profiling. Der Prozess schließt das Erkennen der Anzahl an eindeutigen Werten, Null-Werten und Datenmustern in einer Spalte ein.
- Ausführen der Datendomänenerkennung. Sie können kritische Datenmerkmale innerhalb eines Unternehmens erkennen.
- Pflegen Sie Profilergebnisse, einschließlich Datentypen, Datendomänen, Primär- und Fremdschlüssel.
- Erstellen von Scorecards, um die Datenqualität zu überwachen.
- Wählen Sie ein Betriebssystemprofil zum Erstellen und Ausführen von Spaltenprofilen, Enterprise-Erkennungsprofilen und Scorecards basierend auf den Berechtigungen des im Betriebssystemprofil angegebenen Benutzers aus.
- Verwenden Sie Repository-Objektsperren, damit andere Benutzer keine Arbeiten überschreiben können.
- Verwenden Sie das Versionsverwaltungssystem, um mehrere Versionen eines Profils zu speichern.

- Erstellen und Zuweisen von Tags zu Datenobjekten.
- Schlagen Sie die Bedeutung von Objektnamen als Geschäftsbegriffe in Business Glossary Desktop nach. Beispielsweise können Sie die Bedeutung eines Spaltennamens oder eines Profilnamens nachschlagen, um dessen geschäftliche Anforderungen und gegenwärtige Implementierung zu verstehen.

Die folgenden Aufgaben können im Developer Tool ausgeführt werden.

- Erkennen des Grads der potenziellen Joins zwischen zwei Datenspalten in einer Datenquelle.
- Bestimmen des Prozentsatzes sich überschneidender Daten zwischen Spaltenpaaren einer Datenquelle oder mehreren Datenquellen.
- Vergleichen der Ergebnisse des Spalten-Profiling.
- Generieren eines Mapping-Objekts aus einem Profil.
- Erkennen von Primärschlüsseln in einer Datenquelle.
- Erkennen von Fremdschlüsseln in einem Satz von einer oder mehreren Datenquellen.
- Erkennen von funktionalen Abhängigkeiten zwischen Spalten in einer Datenquelle.
- Ausführen von Datenerkennungs-Tasks über eine große Anzahl von Datenquellen über mehrere Verbindungen hinweg. Zu den Datenerkennungs-Tasks gehören Spaltenprofil, Rückschluss von Primärschlüssel- und Fremdschlüssel-Beziehungen, Datendomänenerkennung und Generierung einer konsolidierten grafischen Übersicht über die Beziehungen zwischen den Daten.

Die folgenden Aufgaben können im Analyst Tool ausgeführt werden:

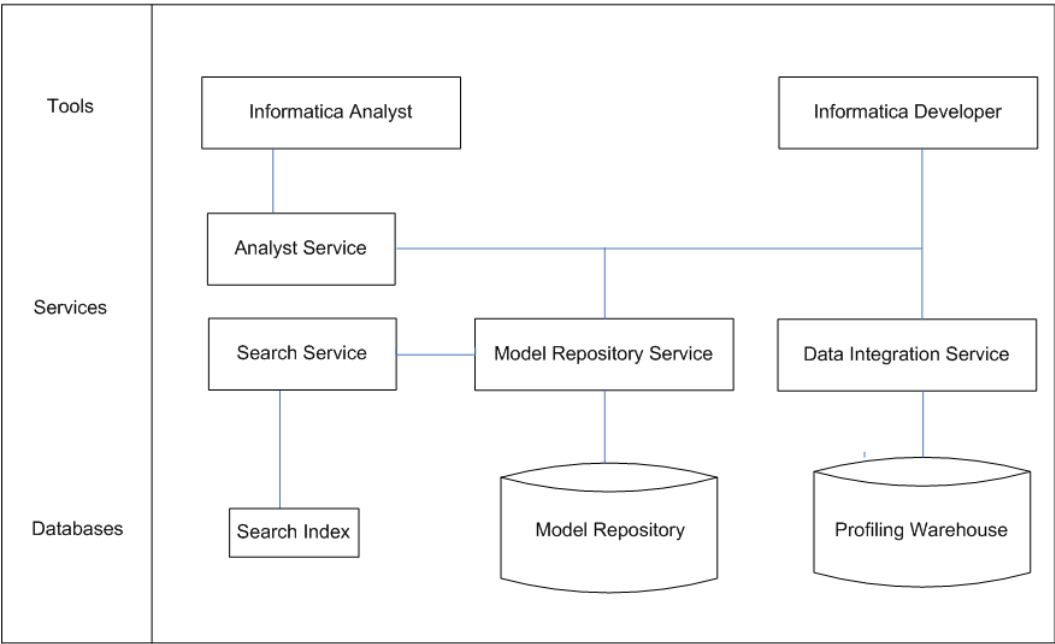
- Ausführen der Enterprise-Erkennung auf vielen Datenquellen über mehrere Verbindungen hinweg. Sie können eine konsolidierte Zusammenfassung für die Erkennungsergebnisse der Spalte Metadaten und Datendomänen anzeigen.
- Ausführen der Erkennungssuche, um zu ermitteln, wo sich die Daten und Metadaten im Unternehmen befinden. Sie können nach bestimmten Objekten wie Datenobjekten, Regeln und Profilen suchen. Bei der Erkennungssuche werden Objekte gesucht und Beziehungen zu anderen Objekten in den Datenbanken und Schemata des Unternehmens ermitteln.
- Anzeigen der Profilergebnisse für eine historische Profilausführung.
- Vergleichen Sie die Profilergebnisse für zwei Profilausführungen in einem Spaltenprofil.
- Die Scorecard-Herkunft kann für jede Scorecard-Metrik oder -Metrikgruppe angezeigt werden.
- Zeigen Sie das Scorecard-Dashboard an.
- Hinzufügen von Kommentaren zu einem Profil oder zu Spalten in einem Profil.
- Zuweisen von Tags zu einem Profil oder zu Spalten in einem Profil.
- Zuweisen von Geschäftsbegriffen zu Spalten in einem Profil.

Profiling-Architektur

Die Profiling-Architektur umfasst Tools, Dienste und Datenbanken. Die Komponenten der Tools bestehen aus Client-Anwendungen. Die Dienstkomponente weist Anwendungsdienste auf, die zum Verwalten der Tools,

Durchführen der Datenintegrationsaufgaben und Verwalten der Metadaten von Profilobjekten erforderlich sind. Die Datenbankenkomponente besteht aus dem Modellrepository und dem Profiling Warehouse.

Die folgende Abbildung zeigt die Architekturkomponenten für das Profiling:



Wenn Sie ein Profil ausführen, erhält der Analyst-Dienst oder das Developer Tool die Profildefinition vom Modellrepository-Dienst. Der Analyst-Dienst oder das Developer Tool rufen dann das Profiling-Plugin im Datenintegrationsdienst auf. Das Profiling-Plugin verarbeitet danach den Profiljob und schickt den Job an den Datenintegrationsdienst. Der Datenintegrationsdienst generiert die Profiling-Ergebnisse. Der Datenintegrationsdienst schreibt die Profiling-Ergebnisse dann in das Profiling Warehouse.

Die Erkennungssuche verwendet den Suchdienst. Bei dem Suchvorgang des Suchdienstes wird nicht das Modellrepository oder das Profiling Warehouse verwendet, sondern der Suchindex. Der Suchdienst generiert den Suchindex basierend auf dem Inhalt im Modellrepository und Profiling Warehouse. Der Suchdienst enthält Extraktionen zum Extrahieren von Inhalt aus jedem Repository.

In der folgenden Tabelle werden die Architekturkomponenten beschrieben:

Komponente	Beschreibung
Informatica Analyst	Eine webbasierte Client-Anwendung, die Sie verwenden können, um Daten und Metadaten von Datenquellen zu erkennen und zu analysieren und um Berichte für diese zu erstellen.
Informatica Developer	Eine Client-Anwendung, die Sie zur Durchführung erweiterter Datenerkennung verwenden können, beispielsweise zur Primärschlüsselerkennung, Fremdschlüsselerkennung und Enterprise-Erkennung.
Analyst-Dienst	Ein Anwendungsdienst, der das Analyst Tool ausführt und Verbindungen zwischen Dienstkomponenten und Benutzern des Analyst Tools verwaltet.
Suchdienst	Ein Anwendungsdienst, der Suchläufe im Analyst Tool verwaltet. Der Suchdienst gibt standardmäßig Suchergebnisse aus dem Modellrepository zurück, z. B. Datenobjekte, Profile, Mapping-Spezifikationen, Referenztabellen, Regeln und Scorecards.

Komponente	Beschreibung
Suchindex	Ein Dateisystem in einem benutzerdefinierten Verzeichnis, das indizierten Inhalt speichert, der vom Suchdienst aus dem Modellrepository und dem Profiling Warehouse extrahiert wird.
Modellrepository-Dienst	Ein Anwendungsdienst, der das Modellrepository verwaltet.
Datenintegrationsdienst	Ein Anwendungsdienst, der Datenintegrationsaufgaben für das Analyst Tool, das Developer Tool und externe Clients durchführt.
Modellrepository	Eine relationale Datenbank, in der die Metadaten für Projekte gespeichert werden, die im Analyst Tool oder Developer Tool erstellt wurden.
Profiling-Warehouse	Eine Datenbank, in der Profiling-Informationen gespeichert werden, beispielsweise Profilergebnisse und Scorecard-Ergebnisse.

Data Discovery-Prozess

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

Für die verschiedenen Arten der Datenanalyse werden verschiedene Profile verwendet, beispielsweise ein Spaltenprofil, eine Primärschlüsselerkennung, Fremdschlüsselerkennung und Datendomänenerkennung. Sie decken Datenqualitätsprobleme auf und dokumentieren sie. Führen Sie folgende Aufgaben aus, um eine Datenerkennung durchzuführen:

1. Finden und Analysieren des Inhalts der Daten in den Datenquellen. Umfasst Datentypen, Werthäufigkeit, Musterhäufigkeit und Datenstatistiken, wie Mindestwert und Maximalwert.
2. Erkennen der Datenstruktur. Umfasst Schlüssel, funktionale Abhängigkeiten und Fremdschlüssel.
3. Überprüfen und validieren Sie Profilergebnisse.
4. Drilldown der Profilergebnisse.
5. Stellen Sie Profilergebnisse wieder her.
6. Erstellen von Referenzdaten.
7. Dokumentieren von Datenproblemen.
8. Erstellen und Ausführen von Regeln.
9. Erstellen von Scorecards, um die Datenqualität zu überwachen.

Mit den folgenden Tools können Sie den Erkennungsvorgang verwalten:

Informatica Administrator

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

Informatica Developer

Erstellen Sie Profile in diesem Tool und führen Sie sie aus, um die Metadaten von einer oder mehreren Datenquellen zu analysieren und auch die Beziehungen zwischen Spalten zu erkennen. Profile werden mit einem Assistenten erstellt.

Informatica Analyst

Sie können ein Spaltenprofil ausführen und eine Datendomänenenerkennung sowie eine Enterprise-Erkennung bei Datenobjekten im Analyst-Tool durchführen. Nach der Ausführung eines Profils können Sie einen Drilldown auf Datenzeilen in einer Datenquelle durchführen.

KAPITEL 2

Data Discovery

Dieses Kapitel umfasst die folgenden Themen:

- [Data Discovery - Übersicht, 21](#)
- [Profil- und Analysetypen, 21](#)
- [Profiling-Komponenten, 22](#)
- [Profilergebnisse, 23](#)

Data Discovery - Übersicht

Data Discovery ist die Suche nach Metadaten von Quellsystemen, die Inhalt und Struktur einschließen. Inhalte beziehen sich auf Datenwerte, Häufigkeiten und Datentypen. Struktur beinhaltet Kandidatenschlüssel, Primärschlüssel, Fremdschlüssel und funktionale Abhängigkeiten. Sie können Profile erstellen und ausführen, um den Inhalt und Struktur der Datenquellen zu erkunden.

Sie können ein Profil zum Analysieren von Daten in einem einzelnen Datenobjekt oder über mehrere Datenobjekte hinweg definieren. Sie können dem Profil Kommentare hinzufügen, um den Profiling-Prozess effektiv nachzuverfolgen.

Zum Evaluieren der Datenstruktur und Überprüfen, dass Datenspalten die erwarteten Informationstypen enthalten, müssen Sie Profile ausführen. Für Datenzeilen in profilierten Daten ist Drilldown möglich. Wenn die Profilergebnisse auf Probleme in den Daten hinweisen, können Sie Regeln anwenden, um den Ergebnissatz zu reparieren. Sie können Scorecards erstellen, um die Datenqualität vor und nach Anwenden der Regeln zu verfolgen und zu messen. Wenn die Metadaten der externen Quelle eines Profils oder einer Scorecard geändert werden, können Sie die Änderungen mit dem jeweiligen Datenobjekt synchronisieren.

Profil- und Analysetypen

Erstellen Sie ein Profil bezogen auf den Analysetyp, den Sie durchführen müssen. Der Profiltyp, den Sie erstellen, entspricht dem Analysetyp, den Sie durchführen. Um eine Primärschlüsselanalyse durchzuführen, erstellen Sie ein Primärschlüsselprofil.

Sie können die folgenden Profile erstellen, um Datenanalyse und Datenerkennung durchzuführen.

Spaltenprofil

Analysieren der Datenqualität in ausgewählten Spalten in einer Tabelle oder Datei. Profile für die Spaltenanalyse können Sie sowohl im Analyst Tool als auch im Developer Tool definieren.

Datendomänenerkennung

Erkennt wichtige Dateneigenschaften innerhalb eines Unternehmens. Die Datendomänenerkennung identifiziert alle Datendomänen, die mit einer Spalte auf der Basis der Spaltenwerte oder des Namens verknüpft sind. Als Teil des Erkennungsprozesses können Sie Datenregeln und Spaltennamenregeln manuell erstellen, um zu überprüfen, ob ein Wert oder ein Spaltenname zu einer Datendomäne gehört. Sie können dann diese Regeln beim Erstellen einer Datendomäne zuweisen. Sie können auch Datendomänen aus den Werten und Mustern in den Spaltenprofilergebnissen erstellen.

Primärschlüsselprofil

Erkennt Primärschlüssel-Beziehungen zwischen Spalten in einer Tabelle oder Datei. Im Developer Tool können Sie Profile für die Primärschlüsselanalyse definieren.

Funktionales Abhängigkeitsprofil

Erkennt funktionale Abhängigkeiten zwischen Spalten in einer Tabelle oder Datei. Mit dem Developer Tool können Sie Profile für die funktionale Abhängigkeitsanalyse definieren.

Fremdschlüsselprofil

Erkennt Fremdschlüssel-Beziehungen zwischen Spalten über mehrere Tabellen oder mehrere Dateien hinweg. Mit dem Developer Tool können Sie Profile für die Fremdschlüsselanalyse definieren.

Join-Profil

Bestimmt den Grad der Überschneidung zwischen Spalten in einer Datenquelle bzw. über mehrere Datenquellen hinweg. Im Developer Tool können Sie Profile für die Join-Analyse definieren. Die Ergebnisse werden im Venn-Diagramm angezeigt.

Überschneidungserkennung

Bestimmt den Prozentsatz sich überschneidender Daten zwischen Spaltenpaaren einer Datenquelle oder mehreren Datenquellen. Sie können die Überschneidungserkennungsaufgabe über den Editor im Developer-Tool ausführen. Sie können die Ergebnisse validieren und in Form eines Venn-Diagramms anzeigen.

Enterprise-Erkennung

Führt mehrere Datenerkennungsaufgaben über eine große Anzahl von Datenquellen aus und erzeugt eine konsolidierte Zusammenfassung der Profilergebnisse. Beinhaltet die Ausführung von Spaltenprofil, Datendomänenerkennung und Erkennung von Primärschlüssel- und Fremdschlüssel-Beziehungen. Die Enterprise-Erkennung automatisiert die Profilierung für eine große Anzahl von Datenquellen.

Hinweis: Änderungen, die Sie im Analyst Tool an Profilen vornehmen, werden erst im Developer Tool angezeigt, wenn Sie die Verbindung zwischen Developer Tool und Modellrepository aktualisiert haben.

Profiling-Komponenten

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

Ein Profil hat folgende Komponenten:

Filter

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

Regel

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

Tag

Metadaten, die ein Objekt im Modellrepository anhand der Unternehmensanwendung definieren. Erstellen Sie Tags, um Objekte entsprechend ihrer Nutzung im Unternehmen zu gruppieren. Verwenden Sie das Analyst Tool, um einem Profil oder Spalten im Profil Tags zuzuweisen.

Kommentar

Beschreibung des Profils. Mithilfe von Kommentaren können Sie mit anderen Benutzern des Analyst oder Developer Tools Profilinformationen gemeinsam nutzen. Verwenden Sie das Analyst Tool, um einem Profil oder Spalten im Profil Kommentare hinzuzufügen.

Scorecard

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

Profilergebnisse

Nachdem Sie ein Profil ausgeführt haben, können Sie die Profilergebnisse anzeigen. Sie können eine Zusammenfassung der Werte, Muster und Statistiken für Spalten und Regeln im Profil anzeigen. Sie können Eigenschaften für die Spalten und Regeln in dem Profil anzeigen. Die Vorschau der Profildaten ist ebenfalls möglich.

Die folgende Tabelle beschreibt die Profilergebnisse für jeden Profiltyp:

Profiltyp	Ergebnisse
Spaltenprofil	<ul style="list-style-type: none">- Anzahl und Prozentsatz der eindeutigen, nicht eindeutigen und Nullwerte in Spalten sowie der abgeleiteten Datentypen für Spaltenwerte.- Häufigkeit und Charaktermuster von Datenwerten in einer ausgewählten Spalte und eine statistische Zusammenfassung für die Spalte.- Datentypen, die mittels Analyse von Spaltendaten abgeleitet werden.- Dokumentierte Datentypen für die Daten.- Maximal- und Mindestwerte.- Datum und Uhrzeit der Profilausführung.
Primärschlüsselprofil	<ul style="list-style-type: none">- Anzahl und Prozentsatz von eindeutigen, doppelten und Nullwerten für gefolgerte Primärschlüsselkandidaten.- Anzahl der Schlüsselverstöße in den gefolgerten Primärschlüsselkandidaten.
Funktionales Abhängigkeitsprofil	<ul style="list-style-type: none">- Gefolgerte Funktionsabhängigkeiten- Anzahl der funktionalen Abhängigkeitsverstöße.
Fremdschlüsselprofil	<ul style="list-style-type: none">- Primär- und Fremdschlüsselspalten, die das von Ihnen definierte Primärschlüssel-Fremdschlüssel-Rückschlusskriterium erfüllen.- Anzahl der Datenwerte, die zwischen den Primär- und Fremdschlüsseln ein Match ergeben, angegeben in Prozent.- Der vor dem Ausführen des Profils für die Primär- und Fremdschlüsselspalten definierte Beziehungstyp.

Profiltyp	Ergebnisse
Join-Profil	<ul style="list-style-type: none"> - Venn-Diagramm, das die Beziehungen zwischen Spalten zeigt. - Anzahl und Prozentsatz der verwaisten, Null- und Joined-Werte in Spalten.
Überschneidungserkennung	<ul style="list-style-type: none"> - Prozentsatz von Überschneidungen zwischen zwei Spalten. - Venn-Diagramm, das die Überschneidungen zwischen Spalten zeigt.
Datendomänenerkennung	<ul style="list-style-type: none"> - Spaltenname und Daten, die vordefinierten Datendomänen entsprechen. - Datendomänengruppe, zu der die Spalte gehört, sowie der jeweilige Datentyp.
Enterprise-Erkennung	<ul style="list-style-type: none"> - Spaltenprofilergebnisse - Datendomänenerkennungsergebnisse. - Primärschlüssel-Erkennungsergebnisse. - Fremdschlüsselprofilergebnisse, sowohl als grafische als auch tabellarische Ansicht.

Mithilfe der Berichterstellungs-Tools von Fremdanbietern können Sie die Profilergebnisse aus dem Profiling-Warehouse abrufen. Informatica stellt mehrere Profilansichten bereit, die Sie für die Profilstatistiken anpassen können, die Sie lesen möchten. Diese Ansichten beruhen auf gängigen Arten von Profilstatistiken und Profilergebnisanalysen.

KAPITEL 3

Spaltenprofilkonzepte

Dieses Kapitel umfasst die folgenden Themen:

- [Spaltenprofilkonzepte - Übersicht, 25](#)
- [Spaltenprofiloptionen, 26](#)
- [Repository-Profilsperrungen und Verwaltung versionierter Profile, 26](#)
- [Scorecards, 27](#)

Spaltenprofilkonzepte - Übersicht

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

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

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

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

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

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

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

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

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

Spaltenprofiloptionen

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

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

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

Repository-Profilesperren und Verwaltung versionierter Profile

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

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

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

Scorecards

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

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

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

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

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

KAPITEL 4

Konzepte der Datendomänenenerkennung

Dieses Kapitel umfasst die folgenden Themen:

- [Konzepte der Datendomänenenerkennung - Übersicht, 28](#)
- [Datendomänen, 29](#)
- [Datendomänenengruppen, 29](#)
- [Datendomänenenglossar, 30](#)
- [Datendomänenenerkennungsprozess, 30](#)
- [Datendomänenenerkennung auf der Spark-Engine, 31](#)

Konzepte der Datendomänenenerkennung - Übersicht

Sie müssen die Bedeutung der kritischen Quelldaten erkennen und verstehen, damit Sie Maßnahmen ergreifen können, um effektiv mit ihnen zu arbeiten. Bei der Datendomänenenerkennung wird die funktionelle Bedeutung von Daten in den Datenquellen auf der Basis der Datensemantik erkannt.

Erstellen Sie ein Profil zur Ausführung der Datendomänenenerkennung, damit Sie kritische Datenmerkmale innerhalb eines Unternehmens erkennen können. Dann können Sie weitere Datenmanagementrichtlinien anwenden, wie beispielsweise Datenqualität oder Datenmaskierung. Erkennen Sie zum Beispiel Produktcodes oder Beschreibungen, um zu analysieren, welche Datenqualitätstandardisierungs- oder Parsing-Regeln Sie anwenden müssen, damit die Daten sinnvoll und vertrauenswürdig sind. Ein anderes Beispiel ist die Suche nach sensiblen Kundendaten wie beispielsweise Kreditkartennummern, E-Mail-IDs und Telefonnummern. Diese Daten können Sie anschließend maskieren, um sie zu schützen.

Ein Profil zur Ausführung der Datendomänenenerkennung kann sowohl im Analyst als auch im Developer Tool erstellt und ausgeführt werden. Sie können ein Profil zur Ausführung der Datendomänenenerkennung basierend auf den folgenden Regeln definieren:

- Datenregel: Sucht nach Spalten mit Daten, die einer bestimmten, in der Regel definierten Logik entsprechen.
- Spaltennamenregel: Sucht nach Spalten, die einer bestimmten, in der Regel definierten Spaltennamenlogik entsprechen.

Sie können Datendomänen aus den Werten und Mustern in den Spaltenprofilerggebnissen erstellen. Sie können dann diese Datendomänen verwenden, um kritische Daten über mehrere Datensysteme oder Unternehmen zu erkennen.

Sie können ein Profil mit einer Stichprobenoption und Filtern erstellen, um Datendomänenerkennung durchzuführen. Beim Ausführen des Profils wenden Sie die Stichprobenoption und die Filter auf die Datenquelle an und erzeugen einen Datensatz. Der Datendomänenerkennungs-Prozess verwendet den Datensatz zur Erkennung von Datendomänen.

Datendomänen

Ein Datendomäne ist ein vordefiniertes oder benutzerdefiniertes Model Repository-Objekt basierend auf der Semantik von Spaltendaten oder Spaltenname. So könne beispielsweise Sozialversicherungsnummer, Kreditkartennummer, E-Mail-ID und Telefonnummer individuelle Datendomänen sein.

Ein Datendomäne hilft Ihnen, wichtige Daten in einer Datenquelle zu finden, die sonst unentdeckt bleiben. Zum Beispiel: Sie verfügen über Altdatensysteme, die Sozialversicherungsnummern in einem Kommentarfeld enthalten. Sie müssen diese Informationen finden und schützen, bevor Sie sie in neue Datensysteme verschieben.

Sie können einen Mindestprozentsatz oder eine Mindestanzahl an Quellzeilen als Konformitätskriterium für den Datendomänenabgleich auswählen. Sie können darüber hinaus Nullwerte ausschließen, wenn Sie die Datendomänenerkennung in einem Spaltenprofil durchführen.

Sie können logische Datendomänen in Datendomänengruppen anordnen. In einem Datendomänenglossar werden alle Datendomänen und Datendomänengruppen aufgelistet. Verwenden Sie das Menü Einstellungen im Developer Tool zum Importieren und Exportieren von Datendomänen in das bzw. aus dem Datendomänenglossar.

Mithilfe von Regeln definieren Sie Daten- und Spaltennamenmuster, die Quelldaten und Metadaten abgleichen. Wenn Sie ein Datendomäne erstellen, kopiert das Analyst oder Developer Tool die zugehörigen Regeln und andere abhängige Objekte in das Datendomänenglossar. Verwenden Sie das Developer Tool zum Verwalten von Datendomänen, einschließlich Importieren und Exportieren von Datendomänen in das bzw. aus dem Datendomänenglossar. Sie können das Developer Tool auch nutzen, um die Regellogik der Datendomänen zu verwalten.

Hinweis: Möglicherweise möchten Sie alle Datendomänenregeln in einem einzelnen Projekt oder Ordner speichern. Dieser Schritt hilft, wenn Sie Datendomänen exportiert haben und dann die Regeln und andere zugehörige Datenobjekte bearbeiten müssen.

Datendomänengruppen

Mithilfe von Datendomänengruppen können Sie Datendomänen in bestimmte Gruppen einordnen. Zum Beispiel können Sie die Datendomänen `first_name`, `last_name`, und `account_number` unter der Datendomänengruppe Personal Health Information (PHI) zusammenfassen.

Sie können eine Datendomänengruppe Personal Identifiable Informations (PII) erstellen, die die Sozialversicherungsnummer, den Vornamen und den Nachnamen enthält. Ein Datendomäne kann Bestandteil mehreren Datendomänengruppen sein. So kann zum Beispiel die Sozialversicherungsnummer zu den beiden Datendomänengruppen Payment Card Industry (PCI) und PII gehören. Datendomänengruppen können Datendomänen enthalten, aber keine anderen Datendomänengruppen.

Hinweis: Wenn Sie die Datendomänendatei `Informatica_IDE_DataDomain.xml` nach der Installation importieren, werden im Datendomänenglossar vordefinierte Datendomänengruppen und Datendomänen angezeigt. Sie können dann nach Bedarf weitere Datendomänengruppen erstellen. Zum Anzeigen und Ändern

der mit den Datendomänen verknüpften Regeln importieren Sie die Datei
`Informatica_IDE_DataDomainRule.xml`.

Datendomänenglossar

Das Datendomänenglossar ist ein Container für alle Domänengruppen und Datendomänen. Sie können das Datendomänenglossar zum Erstellen, Verwalten und Entfernen von Datendomänen und Datendomänengruppen nutzen.

Sie können im Datendomänenglossar nach bestimmten Domänen und Domänengruppen suchen. Darüber hinaus können Sie Datendomänen in eine XML-Datei exportieren und aus einer XML-Datei in das Datendomänenglossar importieren. Das Datendomänenglossar enthält Kopien der Regeln und alle mit den Datendomänen verknüpften Referenzdaten. Die Regeln im Datendomänenglossar können nicht bearbeitet werden.

Sie können das Datendomänenglossar über das Menü Einstellungen im Developer Tool und über das Menü Manager im Analyst Tool ansehen. Legen Sie mit der Model Repository Service-Berechtigung **Datendomänen verwalten** fest, wer Datendomänen und Datendomänengruppen erstellen, bearbeiten und löschen darf.

Datendomänenerkennungsprozess

Sie können ein Profil zur Ausführung der Datendomänenerkennung im Analyst Tool oder Developer Tool entsprechend Ihrer Arbeitsplatzrolle erstellen und ausführen. Nachdem Sie die Optionen der Datendomänenerkennung konfiguriert und das Profil ausgeführt haben, können Sie eine Überprüfung und ein Drilldown der Ergebnisse durchführen. Wenn Sie die Datendomänenerkennung aus dem Editor heraus ausführen, können Sie dem Datenmodell die Ergebnisse hinzufügen.

Führen Sie die folgenden Schritte aus, um eine Datendomänenerkennung durchzuführen:

1. Erstellen oder Importieren von Datendomänen und Domänengruppen.
2. Optionale Konsolidierung von Datendomänen unter entsprechenden Domänengruppen.
3. Erstellen eines Profils zur Ausführung der Datendomänenerkennung. Sie beginnen mit der Auswahl, ob Sie ein Spaltenprofil zusammen mit der Datendomänenerkennung ausführen möchten, oder nur die Datendomänenerkennung.
4. Wählen Sie die Spalten, Domänen und die geeigneten Stichprobenoptionen aus.
5. Führen Sie das Profil aus.
6. Führen Sie eine Überprüfung und einen Drilldown der Profilergebnisse durch und fügen Sie einem Datenmodell die Ergebnisse nach Bedarf hinzu.

Datendomänenerkennung auf der Spark-Engine

Wenn Sie ein Profil ausführen, um die Datenerkennung auf der Spark-Engine durchzuführen, werden auf dem Hadoop-Cluster Referenztabelle bereitgestellt. Um sicherzustellen, dass auf dem Cluster für alle Datendomänen Referenztabelle bereitgestellt werden, können Sie die folgenden Schritte durchführen:

Voraussetzungen:

Sie müssen über die Berechtigung verfügen, sich als HDFS-Benutzer auszugeben, wenn Sie eine Datendomänenerkennung durchführen.

Herunterladen der JDBC-JAR-Dateien

1. Holen Sie sich die JDBC-JAR-Dateien der Referenzdatenbank, die Sie verwenden. Sie können die Dateien von der Website des Datenbank-anbieters herunterladen.
2. Kopieren Sie die heruntergeladenen Dateien an den folgenden Speicherort: `<INFA_HOME>/externaljdbcjars`

Konfigurieren der benutzerdefinierten Eigenschaften für den Data Integration Service

1. Starten Sie Informatica Administrator und wählen Sie anschließend **Data Integration Service** im **Domain Navigator**.
2. Klicken Sie auf die Option **Benutzerdefinierte Eigenschaften** in der Registerkarte **Eigenschaften**.
3. Legen Sie die folgenden benutzerdefinierten Eigenschaften fest, um Referenztabelle für die Datendomänen bereitzustellen:

Eigenschaftsname	Eigenschaftswert
<code>AdvancedProfilingServiceOptions.ProfilingSparkReferenceDataHDFSDir</code>	<code>hdfs:// <Namenode>:<Port>/ tmp/cms</code>
<code>ExecutionContextOptions.SparkRefTableHadoopConnectorArgs</code>	<code>--connect <JDBC thin driver connection URL></code>

4. Stellen Sie sicher, dass das Verzeichnis `hdfs://<Namenode>:<Port>/tmp/cms` auf dem Cluster vorhanden ist. Wenn das Verzeichnis nicht vorhanden ist, erstellen Sie das Verzeichnis `hdfs://<Namenode>:<Port>/tmp/cms` oder ein benutzerdefiniertes Verzeichnis, in dem Sie die Daten bereitstellen möchten. Standardmäßig werden die Referenzdaten im Verzeichnis `hdfs://<Namenode>:<Port>/tmp/cms` bereitgestellt.
5. Starten Sie den Datenintegrationsdienst neu.
6. Öffnen Sie das Analyst Tool oder das Developer Tool und stellen Sie sicher, dass Sie ein erstes Profil mit allen Datendomänen ausführen, um die Referenzdaten bereitzustellen.

Hinweis: Wenn Sie in der ersten Profilausführung nicht alle Datendomänen auswählen und anschließend in der nächsten Profilausführung weitere Datendomänen auswählen, kann die Profilausführung fehlschlagen.

KAPITEL 5

Wiederherstellungskonzepte

Dieses Kapitel umfasst die folgenden Themen:

- [Wiederherstellungskonzepte - Übersicht, 32](#)
- [Wiederherstellung für Analysten und Entwickler, 32](#)
- [Wiederherstellungsaufgaben, 33](#)

Wiederherstellungskonzepte - Übersicht

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

Sie können die folgenden abgeleiteten Profilergebnisse wiederherstellen:

- Datentypen
- Datendomänen
- Primärschlüssel
- Fremdschlüssel

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

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

Wiederherstellung für Analysten und Entwickler

Als Datenanalytiker oder Datenverwalter können Sie die Spaltenprofilergebnisse und Datendomänenenerkennungsergebnisse im Analyst-Tool wiederherstellen. Sie können die Profilergebnisse

wiederherstellen, um genaue Profilinformatoren für die Erkennungssuche und die weitere Validierung der Datenobjekte bereitzustellen.

Als Entwickler oder Datenarchitekt können Sie Spaltenprofilerggebnisse sowie Erkennungsergebnisse von Datendomänen und Primär- und Fremdschlüsseln im Developer-Tool wiederherstellen.

Wiederherstellungsbeispiele

Wenn Sie die Enterprise-Erkennung als Entwickler ausführen, verarbeitet das Developer-Tool die ausgewählten Datendomänen für den gesamten Dataset. Diese Aktion kann mehrere Datendomäneninferenzen zur Folge haben, wie beispielsweise die Telefonnummerdaten, die als Datendomäne der Sozialversicherungsnummer abgeleitet sind. Mehrere Datendomäneninferenzen treten auf, wenn Teile der Daten in einer Spalte mit anderen Datendomänen übereinstimmen. Beispielsweise weist eine 10-stellige Telefonnummer, der eine Ziffer fehlt, möglicherweise dasselbe Muster wie eine Sozialversicherungsnummer auf. Diese Begebenheit deutet auf potenzielle Datenqualitätsprobleme in einer Spalte oder ein übereinstimmendes Muster in mehreren Datendomänen hin. In diesem Fall leitet das Developer-Tool möglicherweise die Datendomäne der Telefonnummer und die der Sozialversicherungsnummer ab. Sie können die Profilergebnisse so wiederherstellen, dass Sie die geeignetste Datendomäne auswählen und genehmigen können. Im Beispiel ist die Telefonnummer die relevante Datendomäne, da die Inferenz der Datendomäne der Sozialversicherungsnummer aufgrund eines Datenqualitätsproblems entstanden ist.

Beim Ausführen der Enterprise-Erkennung leitet das Developer Tool unter Umständen mehrere Datentypen, wie z. B. „Datum/Uhrzeit“, „Zeichenfolge“ und „Varchar“, für eine Datenspalte ab. Als Datenarchitekt möchten Sie unter Umständen den Datentyp „Datum/Uhrzeit“, der als wichtigster Datentyp für eine Datumsspalte fungiert, auswählen und genehmigen.

Die Enterprise-Erkennung im Developer Tool leitet möglicherweise alle Datenobjektbeziehungen basierend auf den Spaltendaten ab. Einige dieser Datenobjektbeziehungen beinhalten unerwünschte Datenobjektbeziehungen in den erkannten Kandidatenschlüsseln. Das Developer Tool leitet beispielsweise Spalten ab, die eine Sequenz als mögliche Schlüssel darstellen und Beziehungen zu anderen Tabellen mit ähnlichen Spalten erkennen. Diese Datenobjektbeziehungen führen möglicherweise nicht zu gültigen Beziehungen in der Datenbank. In solchen Fällen können Sie die passendsten abgeleiteten Profile als Teil der Wiederherstellung bewerten, überprüfen und genehmigen.

Wiederherstellungsaufgaben

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

Sie können die folgenden Wiederherstellungsaufgaben im Analyst-Tool ausführen:

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

Sie können die folgenden Wiederherstellungsaufgaben im Developer-Tool ausführen:

- Genehmigen Sie die abgeleiteten Datentypen für mehrere Spalten oder lehnen Sie sie ab.

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

Teil II: Data Discovery mit Informatica Analyst

Dieser Teil enthält die folgenden Kapitel:

- [Spaltenprofile in Informatica Analyst, 36](#)
- [Regeln in Informatica Analyst, 48](#)
- [Filter in Informatica Analyst, 53](#)
- [Spaltenprofilerggebnisse in Informatica Analyst, 58](#)
- [Geschäftsbegriffe, Kommentare und Tags in Informatica Analyst, 82](#)
- [Scorecards in Informatica Analyst, 85](#)
- [Datenqualitätsfunktionen in Informatica Analyst, 111](#)
- [Enterprise-Erkennung in Informatica Analyst, 122](#)
- [Enterprise-Erkennungsergebnisse in Informatica Analyst, 128](#)
- [Erkennungssuche in Informatica Analyst, 133](#)
- [Business Glossary Desktop in Informatica Analyst, 143](#)

KAPITEL 6

Spaltenprofile in Informatica Analyst

Dieses Kapitel umfasst die folgenden Themen:

- [Spaltenprofile in Informatica Analyst - Übersicht, 36](#)
- [Spalten-Profilingprozess, 37](#)
- [Profil-Optionen, 37](#)
- [Laufzeitumgebung, 39](#)
- [Betriebssystemprofile in Informatica Analyst - Übersicht, 40](#)
- [Repository-Objektsperren und teambasierte Entwicklung - Übersicht, 41](#)
- [Erstellen eines Spaltenprofils in Informatica Analyst, 41](#)
- [Bearbeiten eines Spaltenprofils, 43](#)
- [Ausführen eines Profils, 44](#)
- [Ausführen eines Profils auf der Spark-Engine, 44](#)
- [Synchronisierungsoption, 44](#)

Spaltenprofile in Informatica Analyst - Übersicht

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

Sie haben die Möglichkeit, breite Tabellen und Einfachdateien mit maximal 1.000 Spalten zu profilieren. Wenn Sie ein Profil erstellen oder ausführen, können Sie alle Spalten oder jede Spalte für ein Profil auswählen. Sie können alle Spalten auswählen, für die ein Drilldown durchgeführt werden soll, und Werthäufigkeiten für diese Spalten anzeigen. Sie können in der Datei mit Trennzeichen keine Spalte zur Ausführung eines Profils auswählen, wenn der Spaltenname länger als 245 Zeichen ist.

Sie können kein Profil auf halbstrukturierte Datenquellen auf der Spark-Engine ausführen.

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

- Im Arbeitsbereich **Bibliothek** mit der rechten Maustaste auf das Datenobjekt klicken, um ein Profil zu erstellen.

- Die Standardoptionen verwenden, um ein standardmäßiges Spaltenprofil zu erstellen.
- Die Einstellungen für das Profil anpassen, um ein benutzerdefiniertes Profil zu erstellen.

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

Spalten-Profilingprozess

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

Die folgenden Schritte beschreiben den Spalten-Profilingprozess:

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

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

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

Profil-Optionen

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

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

Stichprobenoptionen

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

Die folgende Tabelle beschreibt die Stichprobenoptionen für ein Profil:

Option	Beschreibung
Alle Zeilen	Führt ein Profil für alle Zeilen im Datenobjekt aus. Unterstützt in der nativen Laufzeitumgebung und in der Blaze- und Spark-Laufzeitumgebung.
Stichprobe erste <number> Zeilen	Führt ein Profil auf Stichprobenzeilen aus, die sich am Anfang der Zeilen des Datenobjekts befinden. Sie können maximal 2.147.483.647 Zeilen auswählen. Unterstützt in der nativen Laufzeitumgebung und in der Blaze-Laufzeitumgebung.
Zufällige Stichprobe <number> Zeilen	Führt ein Profil auf einer zufällig ausgewählten Anzahl von Zeilen im Datenobjekt aus. Sie können maximal 2.147.483.647 Zeilen auswählen. Unterstützt in der nativen Laufzeitumgebung und in der Blaze-Laufzeitumgebung.
Zufällige Stichprobe (auto)	Führt ein Profil auf den Stichprobenzeilen aus, die auf der Grundlage der Anzahl der Zeilen im Datenobjekt berechnet werden. Unterstützt in der nativen Laufzeitumgebung und in der Blaze-Laufzeitumgebung.
Limit n <number> Zeilen	Führt ein Profil basierend auf der Anzahl der Zeilen im Datenobjekt aus. Wenn Sie ein Profil in der Hadoop-Validierungsumgebung ausführen, sammelt die Spark-Engine Stichproben von mehreren Partitionen des Datenobjekts und überträgt die Stichproben in einen einzigen Knoten, um den Stichprobenumfang zu berechnen. Die Stichprobenoption „Limit n“ unterstützt nur Oracle-, SQL Server- und DB2-Datenbanken. Sie können den erweiterten Filter nicht zusammen mit der Stichprobenoption „Limit n“ anwenden. Unterstützt in der Spark-Laufzeitumgebung.
Zufallsprozentsatz	Führt ein Profil auf einem Prozentsatz von Zeilen im Datenobjekt aus. Unterstützt in der Spark-Laufzeitumgebung.
Genehmigte Datentypen und Datendomänen in den nachfolgenden Profilausführungen aus der Datentyp- und Datendomänenableitung ausschließen	Schließt den genehmigten Datentyp oder die genehmigte Datendomäne bei der nächsten Profilausführung aus der Datentyp- und Datendomänenableitung aus.

Nachdem Sie das Profil für eine zufällige Stichprobe von Zeilen ausgeführt haben, wählt der zufällige Stichproben-Algorithmus die Zeilen zufällig in dem Datenobjekt aus, auf dem das Profil ausgeführt werden soll. Wenn Sie eine zufällige Stichproben-Option für Spaltenprofile auswählen, führt das Analyst Tool ein Drilldown für die Staging-Daten durch. Dies kann sich auf die Drilldown-Leistung auswirken. Wenn Sie eine zufällige Stichproben-Option für Datendomänenenerkennungs-Profile auswählen, führt das Analyst Tool ein Drilldown auf Live-Daten aus.

Drilldown-Optionen

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

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

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

Laufzeitumgebung

Sie können „Nativ“ oder „Hadoop“ als Laufzeitumgebung für ein Spaltenprofil auswählen. Sie können in der Hadoop-Laufzeitumgebung die Blaze-Engine oder die Spark-Engine auswählen. Informatica Analyst legt die Laufzeitumgebung in der Profildefinition fest, nachdem Sie eine Laufzeitumgebung ausgewählt haben.

Native Umgebung

Wenn Sie ein Profil in der nativen Laufzeitumgebung ausführen, übermittelt das Analyst Tool die Profiljobs an das Profilerstellungsdienst-Modul. Das Profilerstellungsdienst-Modul teilt die Profiljobs dann in mehrere Zuordnungen auf. Der Datenintegrationsdienst führt diese Zuordnungen auf demselben Computer aus, auf dem der Datenintegrationsdienst ausgeführt wird, und schreibt die Profilergebnisse in das Profiling-Warehouse. Standardmäßig werden alle Profile in der nativen Laufzeitumgebung ausgeführt.

Sie können native Quellen verwenden, um Profile in der nativen Umgebung zu erstellen und auszuführen. Bei einer nativen Datenquelle handelt es sich um eine Nicht-Hadoop-Quelle, wie z. B. eine Einfachdatei, eine relationale oder Mainframe-Quelle. Sie können ein Profil auch in einer Zuordnungsspezifikation oder einer logischen Datenquelle mit einer Hive- oder HDFS-Datenquelle in der nativen Umgebung ausführen.

Hadoop-Umgebung

Sie können die Blaze-Engine oder die Spark-Engine auswählen, um die Profile in der Hadoop-Laufzeitumgebung auszuführen.

Nach Auswahl der Blaze- oder Spark-Engine können Sie eine Hadoop-Verbindung auswählen. Der Datenintegrationsdienst übergibt die Profilogik an die Blaze- oder Spark-Engine auf dem Hadoop-Cluster, um Profile auszuführen.

Wenn Sie ein Profil in der Hadoop-Umgebung ausführen, übermittelt das Developer Tool die Profiljobs an das Profilerstellungsdienst-Modul. Das Profilerstellungsdienst-Modul teilt die Profiljobs dann in mehrere Zuordnungen auf. Der Datenintegrationsdienst übergibt die Zuordnungen über die Hadoop-Verbindung an die Hadoop-Umgebung. Die Blaze-Engine oder die Spark-Engine verarbeitet die Zuordnungen und der Datenintegrationsdienst schreibt die Profilergebnisse in das Profiling-Warehouse.

Spaltenprofile für Sqoop-Datenquellen

Sie können ein Spaltenprofil für Datenobjekte ausführen, die Sqoop verwenden. Nachdem Sie Hadoop als Validierungsumgebung ausgewählt haben, können Sie die Blaze-Engine oder die Spark-Engine auf der Hadoop-Verbindung auswählen, um die Spaltenprofile auszuführen.

Beim Ausführen eines Spaltenprofils für ein logisches oder benutzerdefiniertes Datenobjekt können Sie das num-mappers-Argument konfigurieren, um Parallelismus zu erreichen und die Leistung zu optimieren. Sie müssen darüber hinaus das split-by-Argument konfigurieren, um die Spalte anzugeben, auf deren Grundlage Sqoop die Arbeitseinheiten aufteilen muss.

Verwenden Sie folgende Syntax:

```
--split-by <column_name>
```

Wenn der Primärschlüssel keine gleichmäßige Verteilung der Werte zwischen dem minimalen und maximalen Bereich aufweist, können Sie das split-by-Argument zur Angabe einer anderen Spalte mit gleichmäßiger Datenverteilung konfigurieren, um die Arbeitseinheiten aufzuteilen.

Wenn Sie die split-by-Spalte nicht definieren, teilt Sqoop die Arbeitseinheiten basierend auf folgenden Kriterien auf:

- Wenn das Datenobjekt einen einzigen Primärschlüssel enthält, verwendet Sqoop den Primärschlüssel als split-by-Spalte.
- Wenn das Datenobjekt einen zusammengesetzten Primärschlüssel enthält, verarbeitet Sqoop zusammengesetzte Primärschlüssel standardmäßig ohne das split-by-Argument. Weitere Informationen finden Sie in der Sqoop-Dokumentation.
- Wenn ein Datenobjekt zwei Tabellen mit einer identischen Spalte enthält, müssen Sie die split-by-Spalte mit einem für Tabellen qualifizierten Namen definieren. Wenn beispielsweise CUSTOMER als Tabellename und FULL_NAME als Spaltenname verwendet werden, definieren Sie die split-by-Spalte folgendermaßen:

```
--split-by CUSTOMER.FULL_NAME
```
- Wenn das Datenobjekt keinen Primärschlüssel enthält, wird für den Wert des m-Arguments und des num-mappers-Arguments standardmäßig 1 verwendet.

Wenn Sie Cloudera Connector Powered by Teradata oder Hortonworks Connector for Teradata verwenden und die Tabelle keinen Primärschlüssel enthält, wird das split-by-Argument benötigt.

Betriebssystemprofile in Informatica Analyst - Übersicht

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

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

Auswählen eines Betriebssystemprofils

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

1. Klicken Sie im Kopfzeilenbereich von Informatica Analyst auf **<Benutzername> > Einstellungen**.
Das Dialogfeld **Einstellungen** wird angezeigt.
2. Wählen Sie ein Betriebssystemprofil aus. Klicken Sie auf **Speichern**.

Repository-Objektsperren und teambasierte Entwicklung - Übersicht

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

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

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

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

Erstellen eines Spaltenprofils in Informatica Analyst

Sie können ein benutzerdefiniertes Profil oder ein Standardprofil erstellen. Wenn Sie ein benutzerdefiniertes Profil erstellen, können Sie die Spalten, Stichprobenzeilen und Drilldown-Optionen konfigurieren. Wenn Sie ein

Standardprofil erstellen, wird das Spaltenprofil und die Datendomänenenerkennung für den gesamten Datensatz mit allen Datendomänen ausgeführt.

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

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

Der Assistent **Neues Profil** wird eingeblendet.

2. Die Option **Einzelquelle** ist standardmäßig ausgewählt. Klicken Sie auf **Weiter**.
3. Geben Sie im Bildschirm **Allgemeine Eigenschaften angeben** einen Namen und eine optionale Beschreibung für das Profil ein. Wählen Sie im Feld „Speicherort“ das Projekt oder den Ordner aus, in dem das Profil erstellt werden soll. Klicken Sie auf **Weiter**.
4. Klicken Sie im Bildschirm **Quelle auswählen** auf **Auswählen**, um ein Datenobjekt auszuwählen, oder auf **Neu**, um ein Datenobjekt zu importieren. Klicken Sie auf **Weiter**.

- Wählen Sie im Dialogfeld **Datenobjekt auswählen** ein Datenobjekt aus. Klicken Sie auf **OK**. Im Bereich „Eigenschaften“ werden die Eigenschaften des ausgewählten Datenobjekts angezeigt. Im Bereich „Datenvorschau“ werden die Spalten im Datenobjekt angezeigt.
- Im Dialogfeld **Neues Datenobjekt** können Sie eine Verbindung, ein Schema, eine Tabelle oder eine Ansicht auswählen, um ein Profil dafür zu erstellen. Zudem haben Sie die Möglichkeit, einen Speicherort auszuwählen und einen Ordner zu erstellen, um das Datenobjekt zu importieren. Klicken Sie auf **OK**.

5. Wählen Sie im Bildschirm **Quelle auswählen** die Spalten aus, für die ein Profil ausgeführt werden soll. Wählen Sie optional **Name** aus, um alle Spalten auszuwählen. Klicken Sie auf **Weiter**.

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

6. Legen Sie im Bildschirm **Einstellungen angeben** die Ausführung eines Spaltenprofils, einer Datendomänenenerkennung oder eines Spaltenprofils und einer Datendomänenenerkennung fest. Standardmäßig ist die Option für das Spaltenprofil ausgewählt.

- Wählen Sie **Spaltenprofil ausführen** zum Ausführen eines Spaltenprofils aus.
- Wählen Sie **Datendomänenenerkennung ausführen** zur Durchführung der Datendomänenenerkennung aus. Wählen Sie im Bereich **Datendomäne** die zu erkennenden Datendomänen und ein Konformitätskriterium aus und im Dialogfeld **Spaltenauswahl für Datendomänenenerkennung bearbeiten** die Spalten für die Datendomänenenerkennung.
- Wählen Sie **Spaltenprofil ausführen** und **Datendomänenenerkennung ausführen** aus, um das Spaltenprofil und die Datendomänenenerkennung auszuführen. Wählen Sie die Optionen für Datendomänen im Bereich **Datendomäne** aus.

Hinweis: Die Spalten, die Sie auswählen, werden standardmäßig für das Spaltenprofil und die Datendomänenenerkennung verwendet. Klicken Sie auf **Bearbeiten**, um Spalten für die Datendomänenenerkennung auszuwählen bzw. ihre Auswahl aufzuheben.

- Wählen Sie Daten, Spalten oder Daten und Spalten für die Datendomänenenerkennung aus.
- Wählen Sie eine Stichprobenoption aus. Sie können **Alle Zeilen (vollständige Analyse)**, **Erste Stichprobe**, **Zufallsstichprobe**, **Zufallsstichprobe (auto)**, **Limit n** oder **Zufallsprozentsatz** als Stichprobenoption im Bereich **Profil ausführen für** auswählen. Die Stichproben-Option gilt für das Spaltenprofil und die Datendomänenenerkennung.

- Wählen Sie eine Drilldown-Option aus. Sie können im Bereich **Drilldown** die Drilldown-Option **Live** oder **Zwischengespeichert** auswählen bzw. **Aus**, um den Drilldown zu deaktivieren. Optional können Sie auf **Spalten auswählen** klicken, um Spalten für den Drilldown auszuwählen. Sie können die Datentyp- und Datendomäneninferenz für Spalten mit einem genehmigten Datentyp oder einer genehmigten Datendomäne weglassen.
 - Wählen Sie **Nativ**, **Blaze** oder **Spark** als Laufzeitumgebung aus. Klicken Sie bei Verwendung der Option **Blaze** oder **Spark** auf **Auswählen**, um eine Hadoop-Verbindung im Dialogfeld **Hadoop-Verbindung auswählen** auszuwählen.
7. Klicken Sie auf **Weiter**.
Der Bildschirm **Regeln und Filter angeben** wird geöffnet.
8. Im Bildschirm **Regeln und Filter angeben** können Sie die folgenden Aufgaben durchführen:
- Erstellen, bearbeiten oder löschen einer Regel. Sie können vorhandene Regeln auf das Profil anwenden.
 - Erstellen, bearbeiten oder löschen eines Filters.
- Hinweis:** Wenn Sie eine Scorecard für dieses Profil erstellen, können Sie die Filter, die Sie für das Profil anlegen, wiederverwenden.
9. Klicken Sie auf **Speichern und beenden**, um das Profil zu erstellen, oder auf **Speichern und ausführen**, um das Profil zu erstellen und auszuführen.

Bearbeiten eines Spaltenprofils

Sie können ein Spaltenprofil ändern, nachdem Sie es ausführen.

1. Wählen Sie im Arbeitsbereich **Bibliothek** das Projekt mit dem Profil aus oder wählen Sie das Profil im Bereich **Objekte** aus.
2. Klicken Sie auf den Profilnamen.
Die Zusammenfassungsansicht wird im Arbeitsbereich **Erkennung** angezeigt.
3. Falls das Versionsverwaltungssystem aktiviert ist, klicken Sie auf **Aktionen** > **Auschecken**, um das Profil auszuchecken.
4. Klicken Sie auf **Aktionen** > **Profil bearbeiten**.
Der **Profilassistent** wird angezeigt.
5. Wählen Sie je nach den von Ihnen gewünschten Änderungen eine der folgenden Seitenoptionen aus:
 - **Allgemeine Eigenschaften angeben.** Ändern Sie die allgemeinen Eigenschaften, wie z. B. Name, Beschreibung und Speicherort.
 - **Quelle auswählen.** Wählen Sie eine andere passende Datenquelle und Spalten aus, für die Sie das Profil ausführen.
 - **Einstellungen angeben.** Führen Sie das Spaltenprofil oder das Spaltenprofil und die Datendomänenenerkennung aus. Wählen Sie die zu erkennenden Datendomänen aus und bearbeiten Sie die Datendomänenenerkennungs-, Stichproben- und Drilldown-Optionen.
 - **Regeln und Filter angeben.** Erstellen, bearbeiten oder löschen Sie Regeln und Filter.
6. Klicken Sie auf **Speichern und beenden**, um die Bearbeitung des Profils abzuschließen, oder auf **Speichern und ausführen**, um das Profil zu bearbeiten und auszuführen.
7. Falls das Versionsverwaltungssystem aktiviert ist, müssen Sie die folgenden Aufgaben durchführen:

- Klicken Sie auf **Speichern und beenden**, um die Bearbeitung des Profils abzuschließen.
- Klicken Sie in der Zusammenfassungsansicht auf **Einchecken**, um das Profil einzuchecken.
- Klicken Sie auf **Aktionen > Profil ausführen**, um das Profil auszuführen.

Ausführen eines Profils

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

1. Wählen Sie im Arbeitsbereich **Bibliothek** im Bereich „Projekte“ das Projekt bzw. den Ordner aus, in dem das Profil enthalten ist. Alternativ können Sie das Profil im Bereich **Objekte** auswählen.
2. Klicken Sie auf **Aktionen > Öffnen**.
Die Zusammenfassungsansicht wird im Arbeitsbereich **Erkennung** angezeigt.
3. Klicken Sie auf **Aktionen > Profil ausführen**.
Das Analyst Tool führt eine Profilausführung durch und zeigt die Profilergebnisse in der Zusammenfassungsansicht an.
4. Klicken Sie in der Zusammenfassungsansicht auf eine Spalte, um die Spaltenergebnisse anzuzeigen.
Die Detailansicht wird angezeigt.

Ausführen eines Profils auf der Spark-Engine

Wenn Sie ein Profil mit der JDBC-Verbindung auf der Spark-Engine ausführen, schlägt die Profilausführung fehl.

Bevor Sie das Profil auf der Spark-Engine ausführen, führen Sie die folgenden Schritte aus:

1. Erstellen Sie eine JDBC-Warehouse-Verbindung.
2. Beziehen Sie die Data Direct-JAR-Dateien der Referenzdatenbank, die Sie verwenden, um Daten zu extrahieren.
3. Kopieren Sie die Dateien an die folgenden Speicherorte: <INFA_HOME>/externaljdbcjars

Synchronisierungsoption

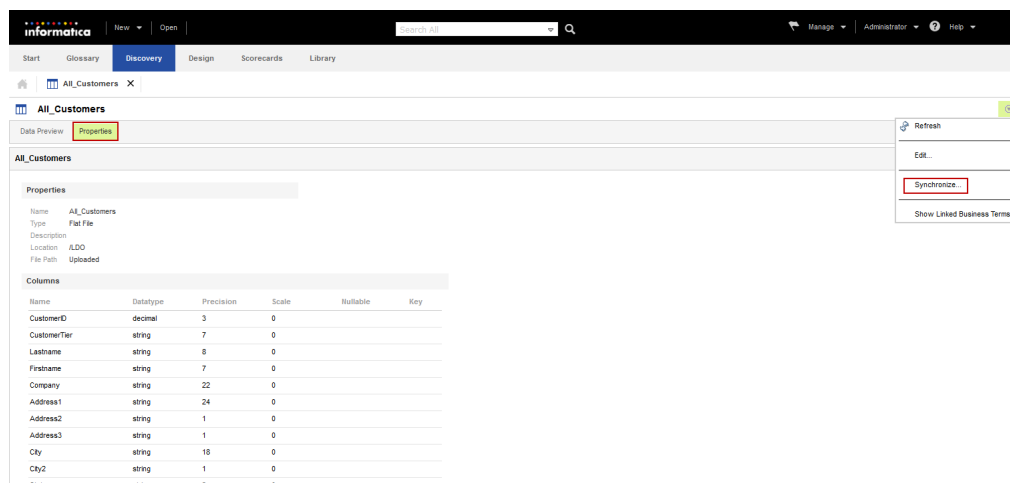
Wenn Sie die Metadaten einer externen Datenquelle ändern, werden die Datenobjekt-Metadaten im Modellrepository standardmäßig nicht aktualisiert. Verwenden Sie die Synchronisierungsoption, um die Datenobjekt-Metadaten mit den Metadaten der Datenquelle zu synchronisieren. Sie können die Synchronisierungsoption für Spaltenprofile, Enterprise-Erkennungsprofile und Scorecards verwenden. Die externe Datenquelle kann eine relationale Datenquelle oder eine Einfachdatei-Datenquelle sein.

Synchronisieren eines Einfachdatei-Datenobjekts in Informatica Analyst

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

1. Öffnen Sie den Arbeitsbereich **Bibliothek**.
2. Wählen Sie im Abschnitt **Projekte** ein Einfachdatei-Datenobjekt aus einem Projekt aus.
Im Analyst Tool werden die Eigenschaften für die Tabelle auf der Registerkarte **Eigenschaften** angezeigt.
3. Klicken Sie im Aktionsmenü auf **Synchronisieren**.

Die folgende Abbildung zeigt die Registerkarte „Eigenschaften“ und die Option „Synchronisieren“ im Menü „Aktionen“:



Der Assistent **Einfachdatei synchronisieren** wird angezeigt.

4. Sie können wahlweise nach einem Speicherort suchen oder einen Netzwerkpfad eingeben, um die Einfachdatei zu importieren.
 - Klicken Sie zum Suchen nach einem Speicherort auf **Datei auswählen**, um die Einfachdatei aus einem Verzeichnis auszuwählen, auf das der Computer Zugriff hat.
 - Wählen Sie zum Eingeben eines Netzwerkpfads **Netzwerkpfad eingeben** aus und konfigurieren Sie den Pfad und den Namen der Datei.

Die folgende Abbildung zeigt den Assistenten zum Synchronisieren von Einfachdateien:

Synchronize Flat File: Step 1 of 5

Specify a location to import the flat file from and specify how to import the flat file.

☒ Browse and Upload:

[Browse...](#)

No file selected.

☐ Enter a Network Path:

☐ Hadoop File System

Description

Upload files from a local machine. Recommended for smaller files up to 10 MB. The Analyst tool uploads a copy of the file to the node on which the Analyst Service runs. Upload the file again if you modify the file.



Back

Next

Finish

Cancel

5. Klicken Sie auf **Weiter**.
6. Importieren Sie wahlweise eine delimitierte Einfachdatei oder eine Einfachdatei mit fester Breite.
 - Um eine mit Trennzeichen getrennte Einfachdatei zu importieren, akzeptieren Sie die Option **Delimitiert**.
 - Um eine Einfachdatei mit fester Breite zu importieren, wählen Sie die Option **Mit fester Breite** aus.
7. Klicken Sie auf **Weiter**.
8. Konfigurieren Sie die Optionen für die delimitierte Einfachdatei bzw. die Einfachdatei mit fester Breite.
9. Klicken Sie auf **Weiter**.
10. Optional können Sie die Spaltenattribute ändern.
11. Klicken Sie auf **Weiter**.
12. Übernehmen Sie den Standardnamen oder geben Sie einen anderen Namen für die Einfachdatei ein.
13. Sie können auch eine Beschreibung eingeben.
14. Klicken Sie auf **Fertig stellen**.

Anhand einer Synchronisierungsnachricht werden Sie aufgefordert, die Aktion zu bestätigen.
15. Klicken Sie auf **Ja**, um die Einfachdatei zu synchronisieren.

Eine Nachricht wird angezeigt, die den Abschluss der Synchronisierung bestätigt. Klicken Sie auf **Details anzeigen**, um Details zu den Änderungen der Metadaten anzuzeigen.
16. Klicken Sie auf **OK**.

Synchronisieren eines relationalen Datenobjekts in Informatica Analyst

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

1. Öffnen Sie den Arbeitsbereich **Bibliothek**.
2. Wählen Sie im Abschnitt **Projekte** ein Tabellendatenobjekt aus einem Projekt aus.
Im Analyst Tool werden die Eigenschaften für die Tabelle auf der Registerkarte **Eigenschaften** angezeigt.
3. Klicken Sie im Aktionsmenü auf **Synchronisieren**.
Sie werden aufgefordert, die Aktion zu bestätigen.
4. Um den Synchronisierungsprozess abzuschließen, klicken Sie auf **Ja**.
Eine Statusnachricht für den Synchronisierungsprozess wird angezeigt.
5. Eine Nachricht wird angezeigt, die den Abschluss der Synchronisierung bestätigt.
Klicken Sie auf **Details anzeigen**, um Details zu den Änderungen der Metadaten anzuzeigen.
6. Klicken Sie auf **OK**.

KAPITEL 7

Regeln in Informatica Analyst

Dieses Kapitel umfasst die folgenden Themen:

- [Regeln in Informatica Analyst - Übersicht, 48](#)
- [Vordefinierte Regeln, 49](#)
- [Ausdrucksregeln, 50](#)

Regeln in Informatica Analyst - Übersicht

Eine Regel ist eine Geschäftslogik, die Bedingungen definiert, welche auf Quelldaten angewendet werden, wenn ein Spaltenprofil ausgeführt wird. Sie können eine Regel zum Profil hinzufügen, um Daten zu validieren.

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

Wenn Sie ein Spaltenprofil erstellen oder bearbeiten, können Sie eine Regel erstellen und dem Profil hinzufügen oder eine vorhandene Regel auf das Profil anwenden. Sie können Ausdrucksregeln oder vordefinierte Regeln in einem Spaltenprofil verwenden.

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

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

Vordefinierte Regeln

Vordefinierte Regeln sind Regeln, die im Developer Tool erstellt oder mit dem Developer Tool und dem Analyst Tool zur Verfügung gestellt wurden. Wenden Sie vordefinierte Regeln auf die Spaltenprofile an, um Quelldaten zu ändern oder zu validieren.

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

Vordefinierte Regeln-Prozess

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

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

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

Anwenden einer vordefinierten Regel

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

1. Wählen Sie im Arbeitsbereich **Bibliothek** das Projekt mit dem Profil aus oder wählen Sie das Profil im Bereich **Objekte** aus.
2. Klicken Sie auf **Aktionen > Öffnen**, um das Profil zu öffnen.
Die Zusammenfassungsansicht wird im Arbeitsbereich **Erkennung** angezeigt.
3. Klicken Sie auf **Aktionen > Profil bearbeiten**.
Der **Profilassistent** wird angezeigt.
4. Klicken Sie auf **Regeln und Filter angeben**.
5. Klicken Sie im Bildschirm **Regeln und Filter angeben** auf **Aktionen > Vorhandene Regel anwenden** im Bereich **Regeln**.
Das Dialogfeld **Regelassistenten anwenden** wird geöffnet.
6. Wählen Sie eine Regel aus und klicken Sie auf **Weiter**.
7. Klicken Sie auf **Hinzufügen**.
Das Dialogfeld **Spalten für Eingabeport auswählen** wird angezeigt.
8. Wählen Sie ein Feld und eine Eingabespalte aus. Klicken Sie auf **OK**.
Die Eingabe- und Ausgabespalten werden im Dialogfeld **Regelassistenten anwenden** angezeigt.
9. Klicken Sie im Dialogfeld **Regelassistenten anwenden** auf **OK**.
Die Regel wird im Bildschirm **Regeln und Filter angeben** angezeigt.

Ausdrucksregeln

Ausdrucksregeln verwenden Ausdrucksfunktionen und Spalten zum Definieren der Regellogik. Erstellen Sie Ausdrucksregeln und fügen Sie sie zu einem Spaltenprofil im Analyst Tool hinzu.

Ausdrucksregeln verwenden Sie zum Ändern oder Validieren von Werten für Spalten in einem Spaltenprofil. Sie können eine oder mehrere Ausdrucksregeln für die Verwendung in einem Profil erstellen.

Ausdrucksfunktionen sind SQL-ähnliche Funktionen zum Umwandeln von Quelldaten. Sie können die Ausdrucksregellogik mit folgenden Funktionstypen erstellen:

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

Sie können die folgenden Methoden verwenden, um eine Ausdrucksregel zu erstellen:

- Profil-Assistent. Wenn Sie ein Spaltenprofil erstellen oder bearbeiten, können Sie Ausdrucksregeln im Profil-Assistenten erstellen und anwenden. Sie können die Regel zu einer wiederverwendbaren Regel heraufstufen und in mehreren Profilen verwenden.
- Regelspezifikation. Sie können eine Regelspezifikation im Analyst Tool konfigurieren und die Regelspezifikation im Spaltenprofil verwenden. Wenn Sie eine Regelspezifikation konfigurieren, übersetzen Sie die Anforderungen einer Geschäftsregel in mindestens eine Regelanweisung. Die Regelanweisungen stellen die Logik für das Festlegen dar, ob ein Datensatz der Geschäftsregel entspricht. Generieren Sie ein Mapplet anhand der Regelspezifikation, und verwenden Sie das Mapplet in den Spaltenprofilen, die Sie im Developer Tool erstellen.

Den Ausdrucks-Editor können Sie verwenden, um Ausdrucksfunktionen hinzuzufügen, Spalten als Eingabe in die Funktionen zu konfigurieren, den Ausdruck zu validieren und den Rückgabebetyp, die Präzision und die Skala zu konfigurieren. Nachdem Sie eine Ausdrucksregel erstellt und validiert haben, können Sie den Präzisionswert der Ausgaberegelspalte bearbeiten. Der Genauigkeitswert der Ausgaberegelspalte ist standardmäßig auf 10 festgelegt. Er wird abgeschnitten, wenn die Ausgaberegelspalte den festgelegten Genauigkeitswert überschreitet.

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

Erstellen einer Ausdrucksregel

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

1. Öffnen Sie ein Profil.

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



Das Dialogfeld **Neue Regel** wird geöffnet.

5. Geben Sie im Dialogfeld **Neue Regel** einen Namen und eine optionale Beschreibung für die Regel ein. Sie können eine Regel im Bereich „Funktionen“ oder im Bereich „Spalten“ erstellen.
 - Wählen Sie im Bereich „Funktionen“ eine Funktionskategorie aus und klicken Sie auf die Schaltfläche mit den nach rechts weisenden Pfeilen (>>). Geben Sie im Dialogfeld Parameter an und klicken Sie auf **OK**.
Die Funktion wird zusammen mit den Spalten und Werten im Bereich „Ausdruck“ angezeigt.
 - Wählen Sie im Bereich „Spalten“ eine Spalte aus und klicken Sie auf die Schaltfläche mit den nach rechts weisenden Pfeilen (>>). Die Spalte wird im Bereich „Ausdruck“ angezeigt. Fügen Sie Funktionen, Ausdrücke und Werte hinzu, um eine Regel zu erstellen.
6. Klicken Sie zum Überprüfen der Regel auf **Validieren**.
7. Sie können die Regel optional als wiederverwendbare Regel deklarieren und den Projekt- und Ordnerspeicherort konfigurieren. Wenn Sie eine Regel als wiederverwendbare Regel definieren, können Sie oder andere Benutzer die Regel als vordefinierte Regel in einem anderen Profil verwenden.
8. Klicken Sie auf **OK**.

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

Erstellen einer Ausdrucksregel unter Verwendung der Regelspezifikation

Sie können die Regelspezifikation verwenden, um eine Ausdrucksregel in Informatica Analyst zu erstellen. Sie können die Regel zu Spaltenprofilen hinzufügen, um Daten zu validieren.

1. Klicken Sie im Kopfzeilenbereich auf **Neu > Regelspezifikation**.
Der Assistent für **Neue Regelspezifikation** wird angezeigt.
2. Geben Sie im Assistenten für **Neue Regelspezifikation** einen Namen und eine optionale Beschreibung für die Regel ein.
3. Klicken Sie im Feld **Speicherort** auf **Durchsuchen**, um das Projekt oder den Ordner auszuwählen, in dem Sie die Regel speichern möchten.
4. Klicken Sie auf **Fortfahren**.
Die Regelspezifikation wird im Arbeitsbereich **Design** angezeigt.
5. Um die Eigenschaften für die Regel einzugeben, wählen Sie das Achteck der obersten Ebene in der Regel aus, und klicken Sie auf **Eigenschaften**.
6. Um einen primären Regelsatz zu konfigurieren, klicken Sie in der Regel auf das Rechteck der nächsten Ebene.
7. Um die Eingaben für den Regelsatz einzugeben, klicken Sie auf **Eigenschaften > Eingaben**.
Das Dialogfeld **Eingabenverwaltung** wird angezeigt.
8. Klicken Sie im Dialogfeld **Eingabenverwaltung** auf **Eingabe hinzufügen** und geben Sie einen Namen, einen Datentyp, eine maximale Länge und eine Beschreibung für die Eingabe ein. Optional können Sie mehrere Eingaben vornehmen.
9. Klicken Sie auf **OK**.
Die Eingaben werden im Abschnitt **Eigenschaften** angezeigt.
10. Um eine Regellogik zu definieren, klicken Sie auf **Regellogik**, geben Sie einen Operator und eine Bedingung ein und wählen Sie eine Aktion aus der Liste **Aktion** aus.
11. Optional können Sie bei Bedarf mehrere Regelsätze eingeben.
12. Um die Regel zu validieren, klicken Sie auf das Symbol **Validieren** ().
13. Um die Regelspezifikation in Spaltenprofilen zu speichern und zu verwenden, klicken Sie auf **Speichern und beenden**.
14. Um die Regel zu speichern und weiter daran zu arbeiten, klicken Sie auf **Speichern und fortfahren**.
15. Um die Regelspezifikation im Developer Tool zu verwenden, klicken Sie auf das Symbol **Regel generieren** () , um ein Mapplet zu erzeugen.

Das Analyst Tool erstellt ein Mapplet im Modellrepository. Validieren Sie das Mapplet als Regel, und verwenden Sie dann das Mapplet in den im Developer Tool erstellten Spaltenprofilen.

KAPITEL 8

Filter in Informatica Analyst

Dieses Kapitel umfasst die folgenden Themen:

- [Filter in Informatica Analyst - Übersicht, 53](#)
- [Erstellen eines Filters, 53](#)
- [Verwalten von Filtern, 56](#)

Filter in Informatica Analyst - Übersicht

Sie können einen Filter erstellen, mit dem Sie eine den Filterkriterien entsprechende Teilmenge der ursprünglichen Quelldaten erzeugen können. Anschließend können Sie ein Profil für die gefilterten Daten ausführen.

Sie können einen Filter zur Anzeige der Profilergebnisse erstellen, die den Filterkriterien entsprechen. Sie können die Profilergebnisse mit den Standardfiltern anzeigen, die in der Zusammenfassungsansicht zur Verfügung stehen.

Erstellen eines Filters

Sie können einen Filter erstellen, mit dem Sie eine den Filterkriterien entsprechende Teilmenge der ursprünglichen Quelldaten erzeugen können.

1. Öffnen Sie ein Profil.
2. Klicken Sie in der Zusammenfassungsansicht auf **Aktionen > Profil bearbeiten**.
Der **Profilassistent** wird angezeigt.
3. Klicken Sie auf **Regeln und Filter angeben**.
4. Klicken Sie im Bereich **Filter** auf **Aktionen > Filter hinzufügen**.
Das Dialogfeld **Neuer Filter** wird geöffnet.
5. Erstellen Sie einen einfachen, erweiterten oder einen SQL-Filter.

Hinweis: Geben Sie für einen einfachen oder erweiterten Filter in einer Datumsspalte die Bedingung im Format JJJJ/MM/TT HH:MM:SS ein.

Im Bereich **Datenvorschau** wird die den Filterkriterien entsprechende Teilmenge der ursprünglichen Datenquelle angezeigt.

6. Klicken Sie auf **OK**.

Der Bildschirm **Regeln und Filter angeben** wird mit dem Filter im Bereich **Filter** angezeigt.

Erstellen eines einfachen Filters

Sie können einen einfachen Filter mit Bedingungsoperatoren erstellen, wie z. B. =, !=, >, <. Verwenden Sie den Filter, um eine Teilmenge der ursprünglichen Datenquelle zu erstellen.

1. Klicken Sie im Dialogfeld **Neuer Filter** auf **Einfach**.

Die folgende Abbildung zeigt die Optionen, mit denen Sie im Dialogfeld **Neuer Filter** einen einfachen Filter erstellen können:

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

Name*:

Description:

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

Columns	Operator	Values(s)
-Select-	-Select-	+

Filter Preview

2. Geben Sie einen Namen und eine optionale Beschreibung ein.
3. Wählen Sie eine Spalte aus.
4. Wählen Sie einen Bedingungsoperator aus.
5. Geben Sie einen Wert ein.
6. Klicken Sie optional auf das Pluszeichen (+), um weitere Filter hinzuzufügen.
7. Klicken Sie auf **OK**.

Die Seite **Regeln und Filter angeben** wird mit dem Filter im Bereich „Filter“ angezeigt.

Erstellen eines erweiterten Filters

Sie können einen erweiterten Filter mit Ausdrücken, wie z. B. AND, OR und NOT, erstellen, um eine Teilmenge der ursprünglichen Datenquelle zu erhalten.

1. Klicken Sie im Dialogfeld **Neuer Filter** auf **Erweitert**.

Die folgende Abbildung zeigt die Optionen für erweiterte Filter im Dialogfeld **Neuer Filter**.

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

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

Erstellen eines SQL-Filters

Sie können einen SQL-Filter mit SQL-Abfragen erstellen. Sie können einen SQL-Filter für relationale Datenquellen erstellen.

1. Klicken Sie im Dialogfeld **Neuer Filter** auf **SQL**.

Die folgende Abbildung zeigt die Optionen für SQL-Filter im Dialogfeld **Neuer Filter**:

2. Geben Sie einen Namen und eine optionale Beschreibung für den SQL-Filter ein.
3. Geben Sie im Textfeld eine SQL-Abfrage ein oder fügen Sie sie dort ein.
4. Klicken Sie auf **Validieren**, um den SQL-Ausdruck zu überprüfen.
5. Klicken Sie auf **OK**.

Die Seite **Regeln und Filter angeben** wird mit dem SQL-Filter im Bereich „Filter“ angezeigt.

Verwalten von Filtern

Sie können Filter bearbeiten und löschen.

1. Wählen Sie im Arbeitsbereich **Bibliothek** das Projekt mit dem Profil aus oder wählen Sie das zu filternde Profil im Bereich **Objekte** aus.
2. Öffnen Sie ein Profil.
3. Klicken Sie in der Zusammenfassungsansicht auf **Aktionen > Profil bearbeiten**, um den **Profilassistenten** zu öffnen.
4. Klicken Sie auf **Regeln und Filter angeben**.

5. Wählen Sie im Bereich „Filter“ einen Filter aus und klicken Sie auf **Aktionen > Filter bearbeiten**.
Das Dialogfeld **Filter bearbeiten** wird angezeigt.
6. Bearbeiten Sie die Filtereinstellungen und klicken Sie auf **OK**.
7. Wählen Sie einen zu löschenden Filter aus, indem Sie auf **Aktionen > Filter löschen** klicken.

KAPITEL 9

Spaltenprofilerggebnisse in Informatica Analyst

Dieses Kapitel umfasst die folgenden Themen:

- [Spaltenprofilerggebnisse in Informatica Analyst - Übersicht, 58](#)
- [Zusammenfassungsansicht, 59](#)
- [Detailansicht, 61](#)
- [Statistiken, 63](#)
- [Typen von Profilausführungen, 70](#)
- [Vergleichen mehrerer Profilergebnisse - Übersicht, 72](#)
- [Spaltenprofil-Drilldown, 77](#)
- [Wiederherstellung im Analyst-Tool, 78](#)
- [Exportdateien des Spaltenprofils in Informatica Analyst, 79](#)

Spaltenprofilerggebnisse in Informatica Analyst - Übersicht

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

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

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

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

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

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

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

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

Zusammenfassungsansicht

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

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

Filter By:	Columns and rules	Name	Null Unique Non-Unique %	Value (min -> max)	Length (min -> max)	Pattern	Data Type	Data Domain	Business Term
Columns and rules 4	Columns	DEPARTMENT_NAME	0 100.00 0	Accounting -> Treasury	2 -> 20	+ 6 more	varchar2 (Documented) String(20) 100%		
Rules 0		DEPARTMENT_ID	0 100.00 0	10.00 -> 270.00	2 -> 3		number(p,s) (Documented) String(3) 100%		
100% null 0		LOCATION_ID	0 25.92 74.08	1400.00 -> 2700.00	4 -> 4		number(p,s) (Documented) String(4) 100%		
100% unique 3		MANAGER_ID	59.25 40.74 0.01	100.00 -> 205.00	3 -> 3		number(p,s) (Documented) String(3) 100%		
100% constant 0									
Conflicting data type 0									
Inferred data domains 0									
Pattern outlier 0									
Value Frequency outlier 0									

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

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

Eigenschaften der Zusammenfassungsansicht

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

In der folgenden Tabelle werden die Zusammenfassungseigenschaften der Profilergebnisse beschrieben:

Eigenschaft	Beschreibung
Name	Zeigt den Namen der Spalte oder Regel im Profil an.
Null Eindeutig Nicht eindeutig %	Zeigt die eindeutigen, nicht eindeutigen und Nullwerte als Prozentsätze für eine Spalten- oder Regelausgabe an. Sie können die Werte in einem horizontalen Balkendiagramm anzeigen.
Muster	Zeigt die Mehrfachmuster in der Spalte als horizontale Balkendiagramme an. Sie können die Musterzeichen und die Anzahl ähnlicher Muster in einer Spalte als Prozentsatz anzeigen, wenn Sie den Mauszeiger über das Balkendiagramm bewegen.
Wert	Zeigt die Mindest- und Maximalwerte in der Spalten- oder Regelausgabe an.
Länge	Zeigt die minimale und maximale Länge der Werte in der Spalten- oder Regelausgabe an.
Datentyp	<p>Zeigt den dokumentierten Datentyp der Spalte oder Regel an. Zeigt die abgeleiteten Datentypen an, wenn Sie den Mauszeiger über das Feld bewegen. Das Analyst Tool kann die folgenden Datentypen ableiten:</p> <ul style="list-style-type: none"> - Zeichenfolge - Varchar - Dezimal - Ganzzahl - Datum <p>Sie können auch den Konformitätsprozentsatz basierend auf den abgeleiteten Datentypen anzeigen.</p> <p>Hinweis: Das Analyst Tool kann den Datentyp nicht aus den Werten einer numerischen Spalte ableiten, deren Genauigkeit größer als 38 ist. Das Analyst Tool kann den Datentyp nicht aus den Werten einer Zeichenfolgenspalte ableiten, deren Genauigkeit größer als 255 ist. Wenn Sie für eine Datumsspalte ein Spaltenprofil mit einem Jahreswert vor 1800 erstellen, wird als abgeleiteter Datentyp unter Umständen eine Zeichenfolge mit fester Länge angezeigt. Ändern Sie den Standardwert für den Parameter Jahreszahlenminimum bei Bedarf in InferDateTimeConfig.xml.</p>
Datendomäne	Zeigt die Namen der mit der Spalte verknüpften Datendomänen mit dem Konformitätsprozentsatz und der Anzahl an konformen Zeilen an.
Geschäftsbegriff	Zeigt den der Spalte zugewiesenen Geschäftsbegriff an.

Standardfilter in der Zusammenfassungsansicht

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

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

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

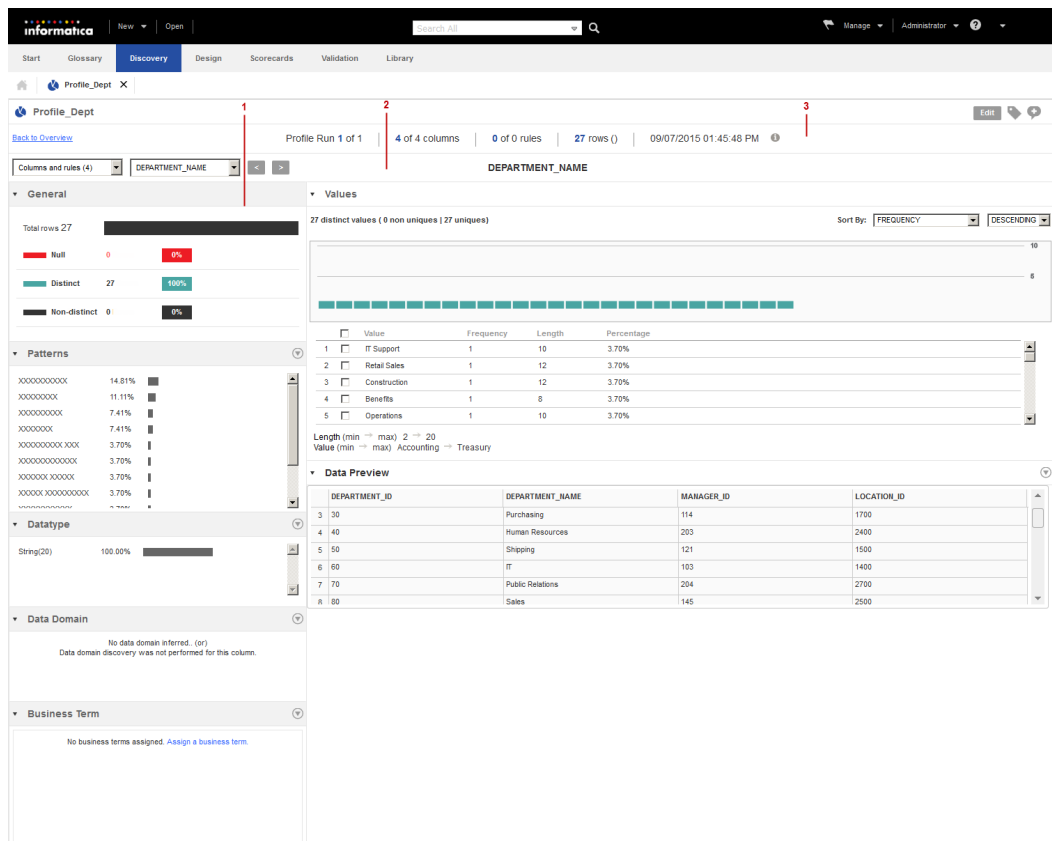
Standardfilteroption	Beschreibung
Spalten und Regeln	Zeigt die Profilergebnisse für die Quell- und Regelspalten an. Sie können die Quell- und Regelspalten zur Anzeige der Ergebnisse erweitern und reduzieren.
Spalten	Zeigt die Profilergebnisse für die Quellspalten an.
Regeln	Zeigt die Profilergebnisse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

Detailansicht

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

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

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



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

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

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

Bereiche der Detailansicht

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

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

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

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

Statistiken

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

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

Datenvorschau

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

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

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

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

Datentypen

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

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

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

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

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

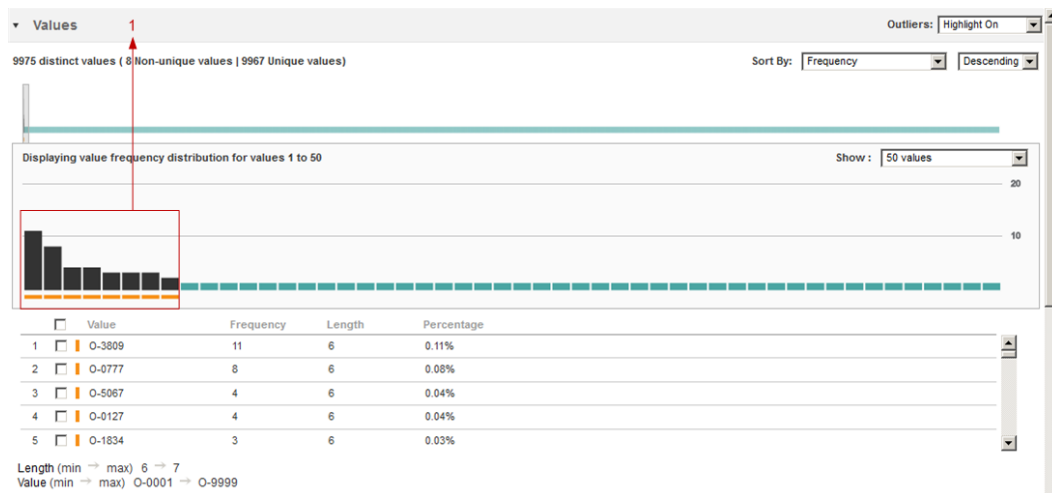
Sonderfälle

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

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

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

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



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

Ausführen eines Sonderfalls

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

1. Klicken Sie in der Zusammenfassungsansicht auf **Aktionen > Sonderfall erkennen**.
Der Wert von „Sonderfall des Musters“ und „Sonderfall der Werthäufigkeit“ im Bereich **Filtern nach** ändert sich von „N/V“ in die Anzahl der erkannten Sonderfälle.
2. Klicken Sie im Bereich **Filtern nach** auf **Sonderfall für Muster**.
Die Spalten mit Sonderfällen bei Mustern werden in der Zusammenfassungsansicht angezeigt.
3. Klicken Sie in der Spalte **Filtern nach** auf **Sonderfall bei Werten und Häufigkeiten**.
Die Spalten mit Sonderfällen bei Werten oder Häufigkeiten werden in der Zusammenfassungsansicht angezeigt.
4. Wählen Sie in der Detailansicht in der Dropdown-Liste für Sonderfälle die Option **Markierung Ein** aus.
Die Sonderfälle werden im Bereich „Werte“ als vertikale Balken angezeigt, die orange unterstrichen sind.
5. Klicken Sie in der Dropdown-Liste „Sonderfälle“ auf **Filter**, um nur Sonderfallwerte anzuzeigen.

Muster

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

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

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

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

Die Eigenschaften für die Spaltenmuster sind in folgender Tabelle beschrieben:

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

Die folgende Tabelle beschreibt die Musterzeichen und was diese darstellen:

Zeichen	Beschreibung
'B' oder 'b' oder ''	Stellt ein Leerzeichen dar.
'C' oder 'c'	Stellt ein beliebiges Zeichen dar.
'L' oder 'l'	Stellt ein beliebiges alphabetisches Zeichen in Kleinschreibung dar.
'T' oder 't'	Stellt ein Tabulatorzeichen dar.
'U' oder 'u'	Stellt ein beliebiges alphabetisches Zeichen in Großschreibung dar.
9	Stellt ein beliebiges numerisches Zeichen dar. Informatica Analyst zeigt separat bis zu drei Zeichen im Format „9“ an. Das Tool zeigt mehr als drei Zeichen als Wert in Klammern an. Das Format „9(8)“ stellt beispielsweise einen numerischen Wert mit 8 Ziffern dar.
'X' oder 'x'	Stellt ein beliebiges alphabetisches Zeichen dar. Informatica Analyst zeigt separat bis zu drei Zeichen im Format „X“ an. Das Tool zeigt mehr als drei Zeichen als Wert in Klammern an. Das Format „X(6)“ kann beispielsweise den Wert „Boston“ darstellen. Hinweis: Das Musterzeichen X unterscheidet nicht zwischen Groß-/Kleinschreibung und kann für Groß- oder Kleinbuchstaben aus der Datenquelle stehen.
'P' oder 'p'	Stellt „(“ dar, die öffnende Klammer.
'Q' oder 'q'	Stellt „)“ dar, die schließende Klammer.

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

Werte

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

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

Werte in der Zusammenfassungsansicht

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

Beispiel

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

Werte in der Detailansicht

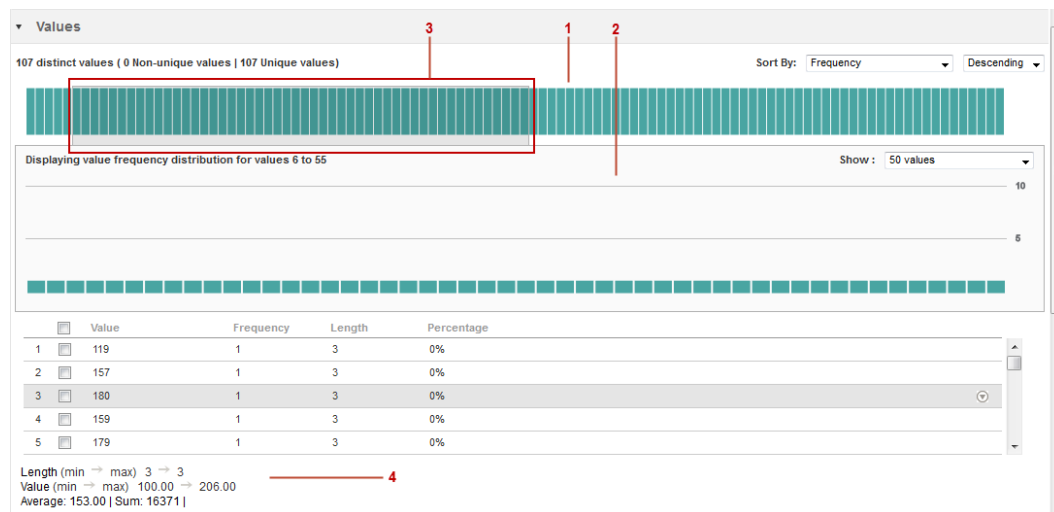
Zu den Spaltenwerten in der Detailansicht gehören Werte für eine Spalte und die Häufigkeit, mit der die Werte in der Spalte angezeigt werden.

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

Der Bereich „Werte“ enthält das grafische Layout und Wertauswahlen.

Das grafische Layout ist in zwei Bereiche unterteilt.

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



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

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

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

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

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

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

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

Werte in der Detailansicht für den Profilergebnisvergleich

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

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

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

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

Typen von Profilausführungen

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

Aktuelle Profilausführung

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

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

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

Historische Profilausführung

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

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

Zusammengefasste Profilausführung

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

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

Beispiel

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

Auswählen einer Profilausführung

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

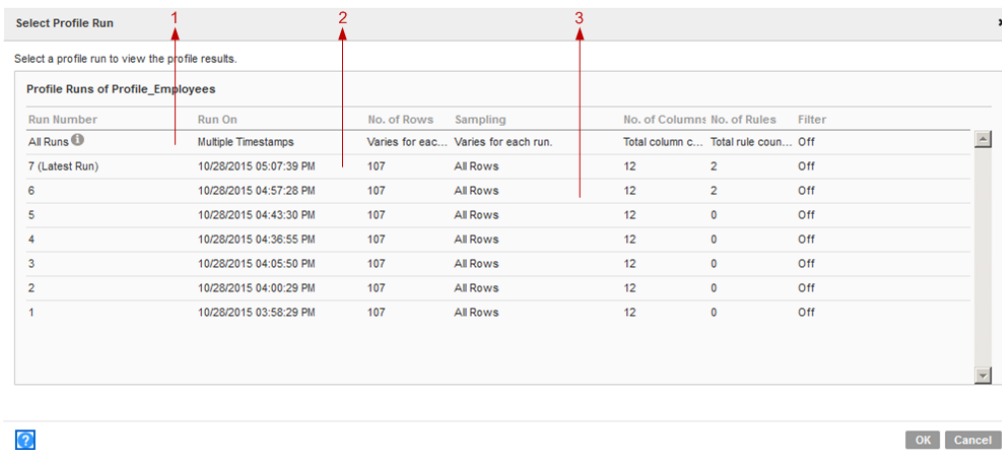
1. Wählen Sie im Arbeitsbereich **Bibliothek** das Projekt bzw. den Ordner aus, in dem das Profil enthalten ist, oder wählen Sie im Bereich **Objekte** das Profil aus.
2. Klicken Sie auf **Aktionen** > **Öffnen**, um das Profil zu öffnen.

Die Zusammenfassungsansicht wird im Arbeitsbereich **Erkennung** angezeigt.

3. Klicken Sie in der Zusammenfassungsansicht auf **Aktionen** > **Profilausführung auswählen**.

Das Dialogfeld **Profilausführung auswählen** wird angezeigt.

Die folgende Abbildung zeigt das Dialogfeld **Profilausführung auswählen**.



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

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

Das Analyst Tool führt eine Profilausführung durch und zeigt die Profilergebnisse in der Zusammenfassungsansicht an.
5. Klicken Sie in der Zusammenfassungsansicht auf eine Spalte, um die Spaltenergebnisse anzuzeigen. Die Detailansicht wird angezeigt.

Vergleichen mehrerer Profilergebnisse - Übersicht

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

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

Vergleichen mehrerer Profilergebnisse

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

1. Klicken Sie in der Zusammenfassungsansicht auf **Aktionen > Profilausführung vergleichen**.
Die folgende Abbildung zeigt das Dialogfeld **Profilausführungen vergleichen**.

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

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

Das Dialogfeld **Profilausführungen vergleichen** wird angezeigt.

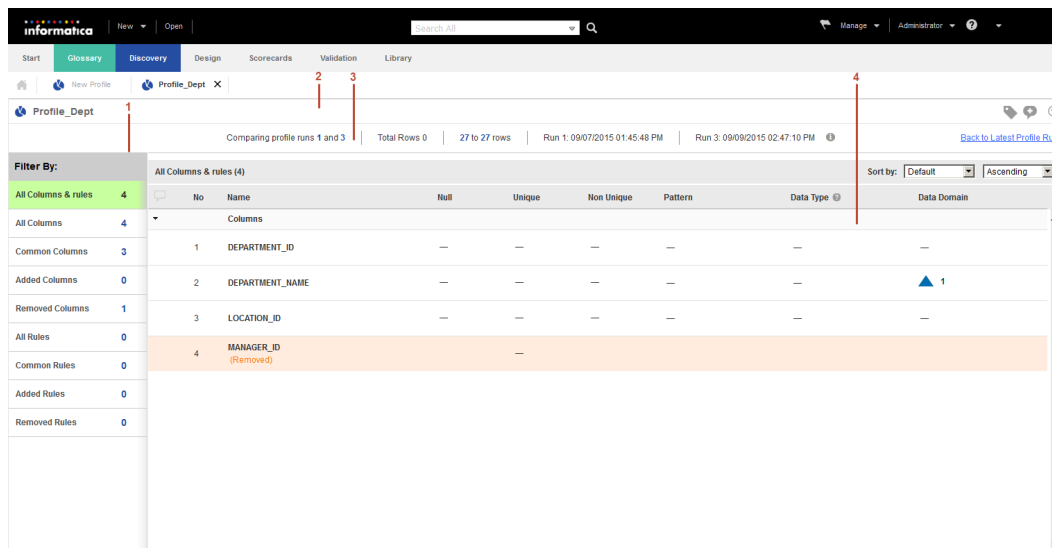
2. Wählen Sie ein Profil aus dem Bereich **Ausführung A** und ein anderes Profil aus dem Bereich **Ausführung B** aus.
3. Klicken Sie auf **OK**.

In der Zusammenfassungsansicht wird eine zusammengefasste Anzeige der Profilergebnisse angezeigt.

Zusammenfassungsansicht für den Profilergebnisvergleich

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

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



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

Eigenschaften der Zusammenfassungsansicht für den Profilergebnisvergleich

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

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

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

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

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

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

Standardfilter für den Profilergebnisvergleich in der Zusammenfassungsansicht

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

In der Zusammenfassungsansicht können Sie Quell- und virtuelle Spalten anzeigen. Die Ausgabe für eine Regel wird als virtuelle Spalte in der Zusammenfassungsansicht angezeigt. Wenn Sie den Ausgabebport für eine Regel ändern und die Profilausführung mit einer historischen Ausführung vergleichen, wird die Ausgabespalte der historischen Regel im Filter **Entfernte Regeln** und die Ausgabespalte der neuen Regel im Filter **Hinzugefügte Regeln** angezeigt. Wenn Sie die Regellogik für eine einzelne Ausgaberegeln oder die Eingaben für eine Mehrfachregelausgabe in einer Profilausführung ändern und mit einer historischen Ausführung vergleichen, bleibt die Filterausgabe unter **Hinzugefügte Regeln** und **Entfernte Regeln** unverändert. Die Filterausgabe ändert sich nicht, da die Filter nur Namensänderungen an den Spalten als gültige Eingaben in den Filter betrachten.

Sie können die folgenden Standardfilteroptionen zum Anzeigen der Profilergebnisse verwenden, die bestimmten Bedingungen entsprechen:

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

Standardfilteroption	Beschreibung
Hinzugefügte Regeln	Zeigt die Regeln an, die in der aktuellen Profilausführung zur Verfügung stehen. Wenn Sie beispielsweise Ausführung 5 mit Ausführung 3 vergleichen, werden unter „Hinzugefügte Regeln“ die Regeln angezeigt, die in Ausführung 5, nicht aber in Ausführung 3 zur Verfügung stehen.
Entfernte Regeln	Zeigt die Regeln an, die in der historischen Profilausführung zur Verfügung stehen. Wenn Sie beispielsweise Ausführung 5 mit Ausführung 3 vergleichen, werden unter „Entfernte Regeln“ die Regeln angezeigt, die in Ausführung 3, nicht aber in Ausführung 5 zur Verfügung stehen.

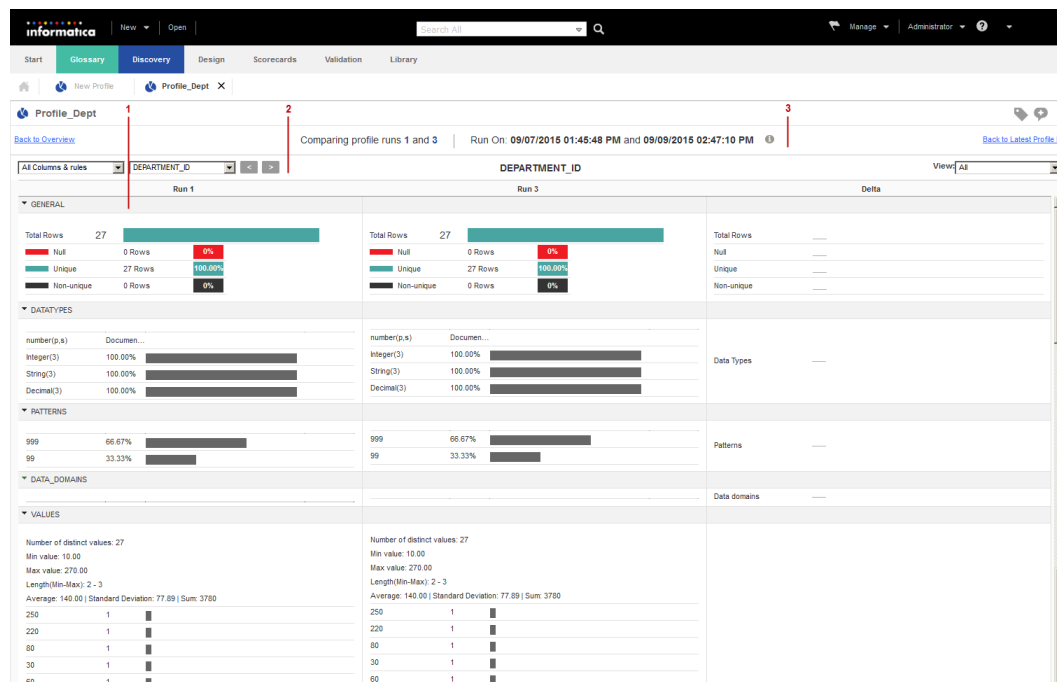
In der Zusammenfassungsansicht werden die Profilergebnisse für alle Quell- und virtuellen Spalten standardmäßig angezeigt.

Detailansicht des Profilergebnisvergleichs

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

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

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



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

3. Kopfzeile der Zusammenfassungsansicht. Sie können profilspezifische Informationen in der Kopfzeile der Zusammenfassungsansicht anzeigen. Sie können die Profilausführungen, die verglichen werden, sowie das Datum und die Uhrzeit der Profilausführungen anzeigen.

Bereiche der Detailansicht für den Profilergebnisvergleich

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

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

Spaltenprofil-Drilldown

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

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

Drilldown an Zeilendaten

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

1. Führen Sie ein Profil aus.
Die Profilergebnisse werden in der Zusammenfassungsansicht angezeigt.
2. Klicken Sie in der Zusammenfassungsansicht auf einen Spaltennamen.
Die Spaltenergebnisse werden in der Detailansicht angezeigt.
3. Klicken Sie in der Detailansicht im Bereich **Werte** mit der rechten Maustaste auf einen Wert und wählen Sie **Drilldown** aus.
Die Drilldown-Daten werden im Bereich **Datenvorschau** angezeigt.

Filter auf Drilldown-Daten anwenden

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

1. Wählen Sie einen Spaltenwert auf der Registerkarte **Werte** aus.

2. Klicken Sie mit der rechten Maustaste und wählen Sie **Drilldown** aus.
Die Drilldown-Ergebnisse werden im Bereich **Datenvorschau** angezeigt.
3. Wenn Sie eine Filterbedingung hinzufügen möchten, klicken Sie im Bereich **Datenvorschau** mit der rechten Maustaste auf einen Spaltenwert und wählen **Zu Filter hinzufügen** aus.
Das Dialogfeld **Drilldown-Filter** wird mit der Filterbedingung geöffnet.
4. Fügen Sie die erforderlichen Filterbedingungen hinzu und klicken Sie auf **OK**.
Drilldown-Filter können auf abgeleitete Datentypen nicht angewendet werden.
5. Klicken Sie auf **Aktionen > Filter speichern**, um den Filter zu speichern.
6. Klicken Sie auf **Aktionen > Aktualisieren**, um die Drilldown-Filter zu löschen.
7. Klicken Sie auf **Aktionen > Daten exportieren**, um die Drilldown-Daten in ein Microsoft Excel-Arbeitsblatt zu exportieren.

Wiederherstellung im Analyst-Tool

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

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

Genehmigen von Datentypen und Datendomänen

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

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

Ablehnen von Datentypen und Datendomänen

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

1. Führen Sie ein Profil aus.
Die Profilergebnisse werden in der Zusammenfassungsansicht angezeigt.
2. Klicken Sie in der Zusammenfassungsansicht auf einen Spaltennamen.
Die Spaltenergebnisse werden in der Detailansicht angezeigt.
3. Wählen Sie in der Detailansicht einen Datentyp im Bereich **Datentyp** oder eine Datendomäne im Bereich **Datendomäne** aus.
4. Klicken Sie auf **Aktionen > Ablehnen**.
Das Analyst Tool entfernt den abgelehnten Datentyp aus der Liste der Datentypen.
5. Klicken Sie zum Anzeigen der abgelehnten Datensätze auf **Aktionen > Abgelehnte einblenden**.

Exportdateien des Spaltenprofils in Informatica Analyst

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

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

Profilieren von Exportergebnissen in eine CSV-Datei

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

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

Profilieren von Exportergebnissen in Microsoft Excel

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

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

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

Exportieren von Profilergebnissen aus Informatica Analyst

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

1. Wählen Sie im Arbeitsbereich **Bibliothek** das Projekt bzw. den Ordner aus, in dem das Profil enthalten ist.
2. Klicken Sie auf das Profil, um es zu öffnen.
Die Profilergebnisse werden in der Zusammenfassungsansicht angezeigt.
3. Klicken Sie in der Zusammenfassungsansicht auf **Aktionen > Daten exportieren**.
Das Dialogfeld **Daten in Datei exportieren** wird angezeigt.
4. Geben Sie im Dialogfeld **Daten in eine Datei exportieren** einen Dateinamen ein. Optional können Sie den Standarddateinamen verwenden.
5. Wählen Sie **Alle (Zusammenfassung, Werte, Muster, Datentypen, Statistiken, Eigenschaften)** oder **Ergebnisse der Datendomänenerkennung** und anschließend eine **Codepage** aus. Klicken Sie auf **OK**.
Die Daten werden in ein Microsoft Excel-Arbeitsblatt exportiert.
6. Klicken Sie in der Zusammenfassungsansicht auf eine Spalte.
Die Spaltenergebnisse werden in der Detailansicht angezeigt.
7. Klicken Sie in der Detailansicht auf **Aktionen > Daten exportieren**.
Das Dialogfeld **Daten in Datei exportieren** wird angezeigt.

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

Die Daten werden in die Datei exportiert.

KAPITEL 10

Geschäftsbegriffe, Kommentare und Tags in Informatica Analyst

Dieses Kapitel umfasst die folgenden Themen:

- [Geschäftsbegriffe, Kommentare und Tags in Informatica Analyst - Übersicht, 82](#)
- [Geschäftsbegriffe, 82](#)
- [Kommentare, 83](#)
- [Tags, 84](#)

Geschäftsbegriffe, Kommentare und Tags in Informatica Analyst - Übersicht

Sie können einem Profil oder Spalten im Profil Geschäftsbegriffe, Kommentare und Tags hinzufügen. Sie können die Geschäftsbegriffe, Kommentare und Tags in der Zusammenfassungs- und Detailansicht zuweisen, anzeigen und bearbeiten.

Geschäftsbegriffe

Im Analyst Tool können Sie Spalten in einem Profil Geschäftsbegriffe zuweisen. Sie können eine Objektverknüpfung bearbeiten oder einen Geschäftsbegriff für eine Spalte löschen. Bei einem Business Glossary handelt es sich um eine Reihe von Bedingungen, die Geschäftssprache verwenden, um Konzepte für Geschäftsbenutzer zu definieren. Ein Geschäftsbegriff stellt die Geschäftsdefinition sowie die Verwendung eines Konzepts bereit.

Sie können in einer Spalte enthaltene Geschäftsbegriffe in der Zusammenfassungs- und Detailansicht zuweisen, anzeigen oder löschen. Zum Anzeigen des Geschäftsbegriffs im Arbeitsbereich **Glossar** klicken Sie in der Detailansicht auf den Geschäftsbegriff.

Sie können die Eigenschaften einer Objektverknüpfung für einen Geschäftsbegriff bearbeiten. Sie können Geschäftsbegriffe gemeinsam mit den Profilergebnissen als virtuelle Spalte in eine CSV-Datei exportieren.

Zuweisen von Geschäftsbegriffen zu Spalten

Weisen Sie Spalten in einem Profil in der Zusammenfassungs- und Detailansicht Geschäftsbegriffe zu. Sie können einen Geschäftsbegriff für eine Spalte im Bereich **Geschäftsbegriffe** löschen. Sie können die Eigenschaften einer Objektverknüpfung für einen Geschäftsbegriff im Dialogfeld **Geschäftsverknüpfung bearbeiten** bearbeiten.

1. Klicken Sie in der Zusammenfassungsansicht mit der rechten Maustaste auf den Spaltennamen und wählen Sie **Geschäftsbegriffe verwalten** aus. Wählen Sie in der Detailansicht die Option **Geschäftsbegriffe verwalten** im Menü **Aktionen** auf der Registerkarte **Geschäftsbegriff** aus.
Der Bereich **Geschäftsbegriffe** wird angezeigt.
2. Klicken Sie auf das Pluszeichen (+) oder auf die Verknüpfung **Geschäftsbegriff zuweisen**, um einen Geschäftsbegriff hinzuzufügen.
Der Bereich **Geschäftsbegriff zuweisen** wird angezeigt.
3. Wählen Sie einen Geschäftsbegriff aus der Liste der Geschäftsbegriffe im Bereich **Geschäftsbegriff zuweisen** aus. Geben Sie einen Objektnamen ein. Optional können Sie den Kontext und die Beschreibung für das Objekt hinzufügen. Klicken Sie auf **OK**.
Der Geschäftsbegriff wird im Bereich **Geschäftsbegriffe** angezeigt.

Kommentare

Sie können einem Profil und den Spalten im Profil Kommentare hinzufügen, um zusätzliche Informationen für die weitere Zusammenarbeit und Analyse bereitzustellen.

Auf der Profilebene können Sie Kommentare über das Profil, die Profildefinition oder die Metadaten des Profils hinzufügen. Sie können die Kommentare für ein Profil in der Zusammenfassungsansicht anzeigen.

Sie können Spaltenkommentare in der Zusammenfassungsansicht und in der Detailansicht hinzufügen und anzeigen.

Sie können folgende Aufgaben mit Kommentaren durchführen:

- Exportieren von Kommentaren als virtuelle Spalte in eine CSV-Datei zusammen mit den Profilergebnissen. Die CSV-Datei enthält alle Kommentare für ein Profil sowie für Spalten im Profil.
- Suchen Sie mithilfe der Schlüsselwörter in der Kommentarspalte nach Profilergebnissen.
- Fügen Sie Kommentare zu den Quell- und den virtuellen Spalten in einem Profil hinzu.

Hinweis: Wenn Sie keine Spalte auswählen bzw. keinen Spaltenkommentar hinzufügen, werden im Kommentarbereich in der Zusammenfassungsansicht Profilkommentare angezeigt.

Hinzufügen von Kommentaren zu einem Profil oder Spalten

Sie können im Bereich **Kommentare** einen Kommentar hinzufügen oder anzeigen.

1. Sie können Kommentare in der Zusammenfassungs- und Detailansicht hinzufügen.
 - Klicken Sie in der Zusammenfassungsansicht zum Hinzufügen von Profilkomentaren auf **Aktionen** > **Kommentare anzeigen**.
 - Klicken Sie in der Zusammenfassungsansicht zum Hinzufügen von Spaltenkommentaren mit der rechten Maustaste auf eine Spalte und wählen Sie **Kommentare anzeigen** aus.

- Klicken Sie in der Detailansicht im Bereich **Allgemein** auf **Kommentar hinzufügen**.
Der Bereich **Kommentare** wird angezeigt.
- 2. Klicken Sie auf **Kommentar hinzufügen**.
Ein Textfeld wird im Bereich **Kommentare** angezeigt.
- 3. Geben Sie einen beschreibenden Kommentartext ein und klicken Sie auf **Speichern**.
Der Kommentar wird im Bereich **Kommentare** mit dem aktuellen Benutzernamen sowie dem Datum und der Uhrzeit der Erstellung angezeigt.

Tags

Sie können einem Profil oder Spalten im Profil Tags zuweisen, um Objekte gemäß ihrer Nutzung im Unternehmen zu gruppieren.

Sie können Profil-Tags in der Zusammenfassungsansicht anzeigen oder zuweisen. Sie können Spalten-Tags in der Zusammenfassungs- und Detailansicht anzeigen oder zuweisen.

Sie können folgende Aufgaben mit Spalten-Tags durchführen:

- Exportieren von Tags gemeinsam mit den Profilergebnissen als virtuelle Spalte in eine CSV-Datei. Die CSV-Datei enthält alle Tags für ein Profil und die Spalten im Profil.
- Zuweisen von Tags zu Quell- und virtuellen Spalten im Profil.

Hinweis: Wenn Sie keine Spalte auswählen oder kein Spalten-Tag hinzufügen, werden im Bereich „Tags“ Profil-Tags in der Zusammenfassungsansicht angezeigt.

Zuweisen von Tags zu einem Profil oder Spalten

Fügen Sie einem Profil in der Zusammenfassungsansicht Tags hinzu. Fügen Sie einer Spalte in der Zusammenfassungs- und Detailansicht Tags hinzu.

1. Sie können Tags in der Zusammenfassungs- und Detailansicht hinzufügen.
 - Klicken Sie in der Zusammenfassungsansicht zum Zuweisen von Tags zu einem Profil auf **Aktionen > Tags anzeigen**.
 - Klicken Sie in der Zusammenfassungsansicht zum Hinzufügen von Tags zu einer Spalte mit der rechten Maustaste auf eine Spalte und wählen Sie **Tags anzeigen** aus.
 - Klicken Sie in der Detailansicht im Bereich **Allgemein** auf **Tag hinzufügen**.
Der Bereich **Tags** wird angezeigt.
2. Klicken Sie auf das Pluszeichen (+) oder auf die Verknüpfung **Tags zuweisen**, um ein Tag zuzuweisen.
Das Dialogfeld **Tags zuweisen** wird geöffnet.
3. Wählen Sie mindestens ein Tag zum Zuweisen zu einem Profil oder einer Spalte aus. Klicken Sie auf **OK**, um den Bereich **Tags** zu öffnen.

Hinweis: Klicken Sie zum Erstellen eines Tags auf **Neues Tag hinzufügen** im Bereich **Tags zuweisen**.

KAPITEL 11

Scorecards in Informatica Analyst

Dieses Kapitel umfasst die folgenden Themen:

- [Scorecards in Informatica Analyst - Übersicht, 85](#)
- [Informatica Analyst Scorecard-Prozess, 86](#)
- [Erstellen einer Scorecard in Informatica Analyst, 87](#)
- [Hinzufügen von Spalten zu vorhandenen Scorecards, 88](#)
- [Hinzufügen von Spalten zu einer vorhandenen Scorecard, 89](#)
- [Ausführen einer Scorecard, 90](#)
- [Anzeigen einer Scorecard, 90](#)
- [Bearbeiten einer Scorecard, 91](#)
- [Maße, 92](#)
- [Metrikgruppen, 93](#)
- [Spalten-Drilldown, 95](#)
- [Trenddiagramme, 95](#)
- [Scorecard-Dashboard in Informatica Analyst, 99](#)
- [Scorecard-Exportdateien in Informatica Analyst, 104](#)
- [Scorecard-Benachrichtigungen, 106](#)
- [Scorecard-Herkunft, 109](#)

Scorecards in Informatica Analyst - Übersicht

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

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

Scorecards zeigen die Werthäufigkeit von Spalten als Scores (Punkte). Die Scores spiegeln den Prozentsatz der gültigen Werte in den Spalten wider. Nachdem Sie ein Profil ausgeführt haben, können Sie Spalten aus dem Profil als Metriken zu einer Scorecard hinzufügen. Sie können Metrikgruppen erstellen, so dass Sie in Verbindung stehende Metriken zu einer einzigen Entität gruppieren können. Sie können Schwellenwerte definieren, die den Bereich von ungültigen Daten, welche für Spalten akzeptabel sind, in einem Datensatz angeben, und Metrikgewichtungen für jede Metrik zuweisen. Wenn Sie eine Scorecard ausführen, generiert

das Analyst Tool gewichtete Durchschnittswerte für jede Metrikgruppe. Zur weiteren Bewertung der Datenqualität können Sie jeder Metrik auch feste oder variable Kosten zuweisen. Beim Ausführen der Scorecard berechnet das Analyst Tool die Summe der Kosten für fehlerhafte Daten für jede Metrik und zeigt die Gesamtkosten an.

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

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

Sie können das Scorecard-Dashboard im Arbeitsbereich **Scorecards** anzeigen. Im Scorecard-Dashboard können Sie die Datenobjekte mit Scorecards, Scorecards in einem Projekt, den Trend bei der Scorecard-Ausführung für die letzten sechs Monate und die Zusammenfassung guter, akzeptabler und inakzeptabler Metriken für alle Scorecard-Ausführungen in einem Monat anzeigen.

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

Informatica Analyst Scorecard-Prozess

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

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

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

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

11. Erstellen Sie eine Gruppe, um der Scorecard verwandte Metriken hinzuzufügen oder diese in die Scorecard zu verschieben.
12. Bearbeiten oder Löschen einer Gruppe nach Bedarf.
13. Zeigen Sie das score-basierte Trenddiagramm für jeden Score an, um die Änderung des Scores während eines bestimmten Zeitraums zu überwachen.
14. Zeigen Sie optional das kostenbasierte Trenddiagramm für jede Metrik an, um den Wert der Datenqualität zu überwachen.
15. Die Scorecard-Herkunft kann für jede Metrik oder Metrikgruppe angezeigt werden.
16. Zeigen Sie konsolidierte Informationen zu den Scorecards an, für die Sie Lesezugriff haben.

Erstellen einer Scorecard in Informatica Analyst

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

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

Die Profilergebnisse werden in der Zusammenfassungsansicht im Arbeitsbereich **Erkennung** angezeigt.

3. Klicken Sie auf **Aktionen > Zu Scorecard hinzufügen**.

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

4. Im Bildschirm **Zu Scorecard hinzufügen** können Sie eine neue Scorecard erstellen oder eine vorhandene bearbeiten, um die Spalten einer vordefinierten Scorecard hinzuzufügen. Standardmäßig ist die Option **Neue Scorecard** ausgewählt. Klicken Sie auf **Weiter**.

5. Geben Sie im Bildschirm **Schritt 2 von 8** einen Namen für die Scorecard ein. Sie können optional eine Beschreibung für die Scorecard eingeben. Wählen Sie das Projekt und den Ordner aus, in dem die Scorecard gespeichert werden soll. Klicken Sie auf **Weiter**.

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

6. Wählen Sie im Bildschirm **Schritt 3 von 8** die Spalten und Regeln aus, die Sie der Scorecard als Metriken hinzufügen möchten. Aktivieren Sie optional das Kontrollkästchen in der linken Spaltenüberschrift, um alle Spalten auszuwählen. Optional können Sie auch **Spaltenname** auswählen, um die Spaltennamen zu sortieren. Klicken Sie auf **Weiter**.

7. Im Bildschirm **Schritt 4 von 8** können Sie der Metrik einen Filter hinzufügen.

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

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

Optional können Sie die ausgewählten Filter auf alle Metriken in der Scorecard anwenden.

Der Filter wird im Bereich **Metrikfilter** angezeigt.

8. Klicken Sie im Bildschirm **Schritt 4 von 8** auf **Weiter**.
 9. Wählen Sie im Bildschirm **Schritt 5 von 8** die einzelnen Metriken im Bereich **Metriken** aus, um die folgenden Aufgaben durchzuführen:
 - Konfigurieren gültiger Werte. Wählen Sie im Bereich **Score verwendet: Werte** unter **Verfügbare Werte** einen oder mehrere Werte aus und klicken Sie auf den nach rechts weisenden Pfeil, um diese in den Bereich **Gültige Werte** zu verschieben. Die Gesamtzahl der gültigen Werte für eine Metrik wird oben im Bereich **Verfügbare Werte** angezeigt.
 - Konfigurieren von Metrikschwellenwerten. Legen Sie im Bereich **Metrikschwellenwerte** die Schwellenwerte für die Scores **Gut**, **Akzeptabel** und **Inakzeptabel** fest.
 - Konfigurieren der Kosten ungültiger Daten. Um den Kosten für die Metrik einen konstanten Wert zuzuweisen, wählen Sie die Option **Feste Kosten** aus. Um eine numerische Spalte für variable Kosten an die Metrik anzuhängen, wählen Sie **Variable Kosten** aus und klicken zum Auswählen einer numerischen Spalte auf **Spalte auswählen**. Klicken Sie optional auf **Kosteneinheit ändern**, um die Kosteneinheit zu ändern. Wenn Sie die Kosten ungültiger Daten für die Metrik nicht konfigurieren möchten, wählen Sie **Keine** aus.
 10. Klicken Sie auf **Weiter**.
 11. Im Bildschirm **Schritt 6 von 8** können Sie eine Metrikgruppe auswählen, der Sie die Metriken hinzufügen können, oder eine neue Metrikgruppe erstellen. Zum Erstellen einer neuen Metrikgruppe klicken Sie auf das Gruppensymbol. Klicken Sie auf **Weiter**.
 12. Geben Sie im Bildschirm **Schritt 7 von 8** die Gewichtungen für die Metriken in der Gruppe sowie Schwellenwerte für die Gruppe an.
 13. Wählen Sie im Bildschirm **Schritt 8 von 8** die Option **Nativ** oder **Hadoop** als Laufzeitumgebung zur Ausführung der Scorecard aus. Sie können in der Hadoop-Laufzeitumgebung die **Blaze-Engine** oder die **Spark-Engine** auswählen. Wenn Sie für die Engine die Option **Blaze** oder **Spark** auswählen, klicken Sie auf **Durchsuchen**, um eine Hadoop-Verbindung zum Ausführen des Profils auszuwählen.
 14. Klicken Sie auf **Speichern**, um die Scorecard zu speichern, oder auf **Speichern und ausführen**, um die Scorecard zu speichern und auszuführen.
- Die Scorecard wird im Arbeitsbereich **Scorecard** angezeigt.

Hinzufügen von Spalten zu vorhandenen Scorecards

Nachdem Sie ein Profil ausgeführt haben, können Sie Spalten in den Profilergebnissen zu einer vorhandenen Scorecard hinzufügen. Sie können Metriken oder Metrikgruppen hinzufügen, gültige Werte für die Spalten konfigurieren und die Kosten ungültiger Daten für jede Metrik hinzufügen. Wenn Sie einer Scorecard eine Spalte aus einem Profil mit einer anderen Stichprobenoption als **Alle Zeilen** hinzufügen, werden auf der Scorecard unter Umständen nicht die Profilergebnisse angezeigt.

Wenn Sie einer vorhandenen Scorecard Spalten hinzufügen, können Sie die vorhandenen Metriken oder Metrikgruppen der Scorecard im Assistenten **Zur Scorecard hinzufügen** nicht bearbeiten. Zum Ändern der vorhandenen Metriken in der Scorecard navigieren Sie zum Arbeitsbereich „Scorecard“, bearbeiten die Scorecard und aktualisieren die Metriken oder Metrikgruppen nach Bedarf.

Hinzufügen von Spalten zu einer vorhandenen Scorecard

Nach der Ausführung eines Profils können Sie einer vorhandenen Scorecard Spalten hinzufügen.

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

Die Profilergebnisse werden in der Zusammenfassungsansicht angezeigt.

2. Wählen Sie eine Spalte aus. Klicken Sie auf **Aktionen > Zu Scorecard hinzufügen**.

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

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

- Sie können einer Scorecard keine Spalte hinzufügen, wenn Spaltenname und Scorecard-Name übereinstimmen.
- Das zweimalige Hinzufügen ein- und derselben Spalte zu einer Scorecard ist auch dann nicht möglich, wenn Sie den Spaltennamen ändern.

3. Wählen Sie **Vorhandene Scorecard** aus, um die Spalten einer vordefinierten Scorecard hinzuzufügen. Klicken Sie auf **Weiter**.

4. Wählen Sie im Bildschirm **Schritt 2 von 7** die Scorecard aus, der Sie die Spalten hinzufügen möchten. Klicken Sie auf **Weiter**.

Sie können die vorhandenen Metriken und Metrikgruppen anzeigen, die der Scorecard zugeordnet sind.

5. Wählen Sie im Bildschirm **Schritt 3 von 7** die Spalten und Regeln aus, die Sie der Scorecard als Metriken hinzufügen möchten. Aktivieren Sie optional das Kontrollkästchen in der linken Spaltenüberschrift, um alle Spalten auszuwählen. Klicken Sie auf **Spaltenname**, um die Spaltennamen zu sortieren. Klicken Sie auf **Weiter**.

6. Im Bildschirm **Schritt 4 von 7** können Sie Filter für die Metriken erstellen. Sie können den Filter, den Sie für das Profil erstellt haben, auch auf Metriken anwenden.

7. Im Bildschirm **Schritt 5 von 7** können Sie die folgenden Aufgaben durchführen:

- Wählen Sie im Bereich **Metriken** die einzelnen Metriken aus und konfigurieren Sie Metrikwerte in den anderen Bereichen.
- Wählen Sie im Bereich **Score verwendet: Werte** unter **Verfügbare Werte** mehrere Werte aus und klicken Sie auf den Pfeil nach rechts, um diese in den Bereich **Gültige Werte** zu verschieben. Die Gesamtzahl der gültigen Werte für eine Metrik wird oben im Bereich **Verfügbare Werte** angezeigt.
- Im Bereich **Metrikschwellenwerte**: können Sie Schwellenwerte für die Scores **Gut**, **Akzeptabel** und **Inakzeptabel** festlegen.
- Im Bereich **Kosten für ungültige Daten** können Sie die folgenden Aufgaben durchführen:
 - Wählen Sie alle Metriken aus und konfigurieren Sie die Kosten für ungültige Daten für die jeweilige Metrik.
 - Wählen Sie die Option **Feste Kosten** aus, um den Kosten für die Metrik einen konstanten Wert zuzuweisen. Sie können auf **Kosteneinheit ändern** klicken, um die Kosteneinheit zu ändern.
 - Wählen Sie die Option **Variable Kosten** aus, um eine numerische Spalte für variable Kosten an die Metrik anzuhängen. Sie können auf **Spalte auswählen** klicken, um eine numerische Spalte auszuwählen.

8. Klicken Sie auf **Weiter**.

9. Im Bildschirm **Schritt 6 von 7** können Sie die folgenden Aufgaben durchführen:

- Wählen Sie die Metrikgruppe aus, der Sie die Metriken hinzufügen möchten.
 - Im Bereich **Standard - Metriken** können Sie auf die standardmäßige Metrikgewichtung 0 doppelklicken, um den Wert zu ändern.
 - Im Bereich **Metrikschwellenwerte**: können Sie Schwellenwerte für die Scores **Gut**, **Akzeptabel** und **Inakzeptabel** festlegen.
10. Klicken Sie auf **Weiter**.
 11. Wählen Sie im Bildschirm **Schritt 7 von 7** eine Laufzeitumgebung aus.
 12. Klicken Sie auf **Speichern**, um die Scorecard zu speichern, oder auf **Speichern und ausführen**, um die Scorecard zu speichern und auszuführen.

Ausführen einer Scorecard

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

1. Wählen Sie im Bereich **Objekte** die Scorecard aus, die Sie ausführen möchten.
2. Klicken Sie auf die Scorecard, um sie zu öffnen.
Die Scorecard wird im **Scorecards**-Arbeitsbereich angezeigt.
3. Klicken Sie auf **Aktionen > Scorecard ausführen**.
4. Wählen Sie einen Score aus dem Bereich **Metriken** und wählen Sie die Spalten für den Drilldown im Bereich **Spalten** aus.
5. Mit der Option **Drilldown** können Sie wählen, ob der Drilldown an Live-Daten oder zwischengespeicherten Daten durchgeführt werden soll.
Optimale Leistung erzielen Sie mit einem Drilldown für Live-Daten.
6. Klicken Sie auf **Ausführen**.

Anzeigen einer Scorecard

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

1. Führen Sie eine Scorecard aus, um die Scores anzuzeigen.
2. Wählen Sie eine Metrik aus, die den anzuzeigenden Score enthält.
3. Klicken Sie auf **Aktionen > Drilldown**, um die Zeilen mit gültigen Daten oder Zeilen mit Daten, die für die Spalte nicht gültig sind, anzuzeigen.
Im Analyst-Tool werden die Zeilen der ungültigen Daten standardmäßig im Abschnitt **Drilldown** angezeigt.

Bearbeiten einer Scorecard

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

1. Klicken Sie im Arbeitsbereich **Bibliothek** auf die Scorecard, die Sie im Bereich **Objekte** bearbeiten möchten.
Die Scorecard wird im Arbeitsbereich **Scorecards** angezeigt.
2. Falls das Versionsverwaltungssystem aktiviert ist, klicken Sie auf **Aktionen** > **Auschecken**.
3. Klicken Sie auf **Aktionen** > **Bearbeiten** > **Allgemein**.
Das Dialogfeld **Scorecard bearbeiten** wird eingeblendet.
4. Auf der Registerkarte **Allgemein** können Sie gegebenenfalls den Namen und die Beschreibung der Scorecard bearbeiten.
5. Klicken Sie auf die Registerkarte **Metriken**.
6. Wählen Sie einen Score im Bereich **Metriken** aus und konfigurieren Sie die gültigen Werte in der Liste aller Werte im Bereich **Score verwendet: Werte**.
7. Nehmen Sie im Bereich **Metrikschwellenwerte** gegebenenfalls Änderungen an den Schwellenwerten des Scores vor.
8. Überprüfen Sie die Kosten für ungültige Daten für jede Metrik und nehmen Sie notwendige Änderungen vor.
9. Klicken Sie auf die Registerkarte **Scorecard-Filter**.
10. Sie können Filter hinzufügen, bearbeiten oder löschen.
11. Klicken Sie auf die Registerkarte **Metrikgruppen**.
12. Sie können die Metrikgruppen erstellen, bearbeiten oder entfernen.
Sie können darüber hinaus die Metrikgewichtungen und Metrikschwellenwerte auf der Registerkarte **Metrikgruppen** bearbeiten.
13. Klicken Sie auf die Registerkarte **Benachrichtigungen**.
14. Sie können gegebenenfalls Änderungen an den Einstellungen für Scorecard-Benachrichtigungen vornehmen.
Sie können globale und benutzerdefinierte Einstellungen für Metriken und Metrikgruppen festlegen.
15. Wählen Sie **Nativ**, **Blaze** oder **Spark** als Laufzeitumgebung aus. Wenn Sie **Blaze** oder **Spark** als Laufzeitumgebung auswählen, klicken Sie auf **Durchsuchen**, um die Hadoop-Verbindung auszuwählen.
16. Klicken Sie auf **Speichern**, um Änderungen an der Scorecard zu speichern, oder auf **Speichern und Ausführen**, um die Änderungen zu speichern und die Scorecard auszuführen.
17. Klicken Sie auf **Einchecken**.

Maße

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

Metrikgewichtungen

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

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

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

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

Wert der Datenqualität

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

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

Feste Kosten

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

Variable Kosten

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

Beispiel

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

Definieren von Schwellenwerten

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

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

- Öffnen Sie ein Profil und fügen Sie im Dialogfeld **Zu Scorecard hinzufügen** Spalten aus dem Profil zur Scorecard hinzu.
 - Sie können auch im Arbeitsbereich **Bibliothek** auf eine Scorecard klicken und **Aktionen > Bearbeiten** auswählen, um die Scorecard im Dialogfeld **Scorecard bearbeiten** zu bearbeiten.
1. Wählen Sie im Dialogfeld **Zu Scorecard hinzufügen** oder im Dialogfeld **Scorecard bearbeiten** jede Metrik im Bereich **Metriken** aus.
 2. Geben Sie im Bereich **Metrikschwellenwerte** die Schwellenwerte ein, mit denen die obere Grenze des inakzeptablen Bereichs und die untere Grenze des guten Bereichs wiedergegeben werden.
Sie können Schwellenwerte für maximal zwei Dezimalstellen festlegen.
 3. Klicken Sie auf **Weiter** oder **Speichern**.

Metrikgruppen

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

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

Erstellen einer Metrikgruppe

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

1. Klicken Sie im Arbeitsbereich **Bibliothek** auf die Scorecard, die Sie im Bereich **Objekte** bearbeiten möchten.
Die Scorecard wird im Arbeitsbereich **Scorecards** angezeigt.
2. Klicken Sie auf **Aktionen > Bearbeiten**.
Das Fenster **Scorecard bearbeiten** wird angezeigt.
3. Klicken Sie auf die Registerkarte **Metrikgruppen**.
Die Standardgruppe wird im Bereich **Metrikgruppen** angezeigt, und die Scores in der Standardgruppe werden im Bereich **Metriken** angezeigt.
4. Klicken Sie auf das Symbol **Neue Gruppe**, um eine Metrikgruppe zu erstellen.
Das Dialogfeld **Metrikgruppen** wird eingeblendet.
5. Geben Sie einen Namen und eine optionale Beschreibung ein.

6. Klicken Sie auf **OK**.
7. Klicken Sie auf **Speichern**, um die Änderungen an der Scorecard zu speichern.

Verschieben von Scores in eine Metrikgruppe

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

1. Klicken Sie im Arbeitsbereich **Bibliothek** auf die Scorecard, die Sie im Bereich **Objekte** bearbeiten möchten.
Die Scorecard wird im Arbeitsbereich **Scorecards** angezeigt.
2. Klicken Sie auf **Aktionen > Bearbeiten**.
Das Fenster **Scorecard bearbeiten** wird angezeigt.
3. Klicken Sie auf die Registerkarte **Metrikgruppen**.
Die Standardgruppe wird im Bereich **Metrikgruppen** angezeigt, und die Scores in der Standardgruppe werden im Bereich **Metriken** angezeigt.
4. Wählen Sie eine Metrik aus dem Bereich **Metriken** aus und klicken Sie auf das Symbol **Metriken verschieben**.
Das Dialogfeld **Metriken verschieben** wird eingeblendet.
Hinweis: Um mehrere Scores auszuwählen, halten Sie die Umschalttaste gedrückt.
5. Wählen Sie die Metrikgruppe aus, in die die Scores verschoben werden sollen.
6. Klicken Sie auf **OK**.

Bearbeiten einer Metrikgruppe

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

1. Klicken Sie im Arbeitsbereich **Bibliothek** auf die Scorecard, die Sie im Bereich **Objekte** bearbeiten möchten.
Die Scorecard wird im Arbeitsbereich **Scorecards** angezeigt.
2. Klicken Sie auf **Aktionen > Bearbeiten**.
Das Fenster **Scorecard bearbeiten** wird angezeigt.
3. Klicken Sie auf die Registerkarte **Metrikgruppen**.
Die Standardmetrikgruppe wird im Bereich **Metrikgruppen** und die Metriken in der Standardmetrikgruppe im Bereich **Metriken** angezeigt.
4. Klicken Sie im Bereich **Metrikgruppen** auf das Symbol **Gruppe bearbeiten**.
Das Dialogfeld **Bearbeiten** wird eingeblendet.
5. Geben Sie den Namen und optional eine Beschreibung ein.
6. Klicken Sie auf **OK**.

Löschen einer Metrikgruppe

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

1. Klicken Sie im Arbeitsbereich **Bibliothek** auf die Scorecard, die Sie im Bereich **Objekte** bearbeiten möchten.
Die Scorecard wird im Arbeitsbereich **Scorecards** angezeigt.
2. Klicken Sie auf **Aktionen > Bearbeiten**.
Das Fenster **Scorecard bearbeiten** wird angezeigt.
3. Klicken Sie auf die Registerkarte **Metrikgruppen**.
Die Standardmetrikgruppe wird im Bereich **Metrikgruppen** und die Metriken in der Standardmetrikgruppe im Bereich **Metriken** angezeigt.
4. Wählen Sie eine Metrikgruppe im Bereich **Metrikgruppen** und klicken Sie auf das Symbol **Gruppe löschen**.
Das Dialogfeld **Gruppen löschen** wird eingeblendet.
5. Wählen Sie die Option zum Löschen der Metriken in der Metrikgruppe oder die Option zum Verschieben der Metriken in die Standardmetrikgruppe, bevor die Metrikgruppe gelöscht wird.
6. Klicken Sie auf **OK**.

Spalten-Drilldown

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

1. Führen Sie eine Scorecard aus, um die Scores anzuzeigen.
2. Wählen Sie eine Spalte, die den anzuzeigenden Score enthält.
3. Klicken Sie auf **Aktionen > Zeilen anzeigen**, um die Zeilen der gültigen bzw. ungültigen Daten der Spalte anzuzeigen.
4. Klicken Sie auf **Aktionen > Drilldown-Spalten**
Die Spalten erscheinen im Fenster **Drilldown** für den ausgewählten Score. Im Analyst Tool werden standardmäßig die Zeilen mit gültigen Daten für die Spalten angezeigt. Optional können Sie auf **Ungültig** klicken, um die ungültigen Datenzeilen anzuzeigen.

Trenddiagramme

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

Die Trenddiagramme enthalten score- und kostenbasierte Diagramme, die die Score- und Kostenwerte auf der vertikalen Achse und alle Scorecard-Ausführungen auf der horizontalen Achse abbilden. Standardmäßig

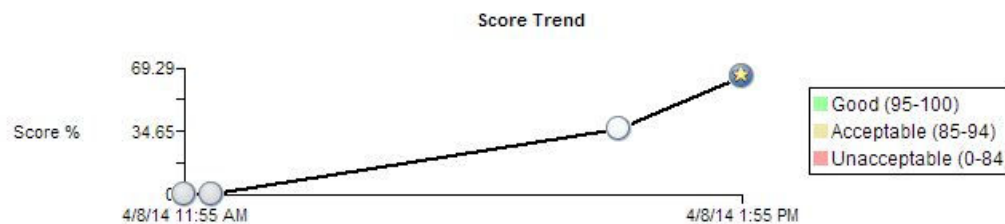
werden in einem Trenddiagramm die Daten aus den letzten 10 Scorecard-Ausführungen angezeigt. Sie können die Gesamtzahl der Zeilen sowie der ungültigen Zeilen für die Metrik im Trenddiagramm anzeigen. Das Trenddiagramm zeigt darüber hinaus, ob die Score- und Kostentrends konstant geblieben sind oder sich basierend auf der letzten Scorecard-Ausführung noch oben oder unten bewegt haben.

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

Score-basiertes Trenddiagramm

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

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



Beispiel

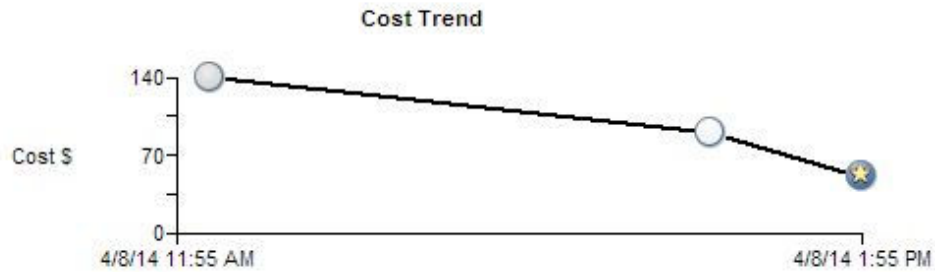
Als Datenanalyst können Sie die Datenqualität überwachen und analysieren, ob die Änderungen an Mappings und anderen Prozessen zu einer Erhöhung des Datenqualitäts-Scores führen. Nachdem Sie die Änderung in der Datenqualität gemessen haben, können Sie die Datenqualitätsänderung zurückmelden, damit sie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

Kostenbasiertes Trenddiagramm

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

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

Die folgende Abbildung zeigt ein Beispiel für ein kostenbasiertes Trenddiagramm:



Beispiel

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

Anzeigen von Trenddiagrammen

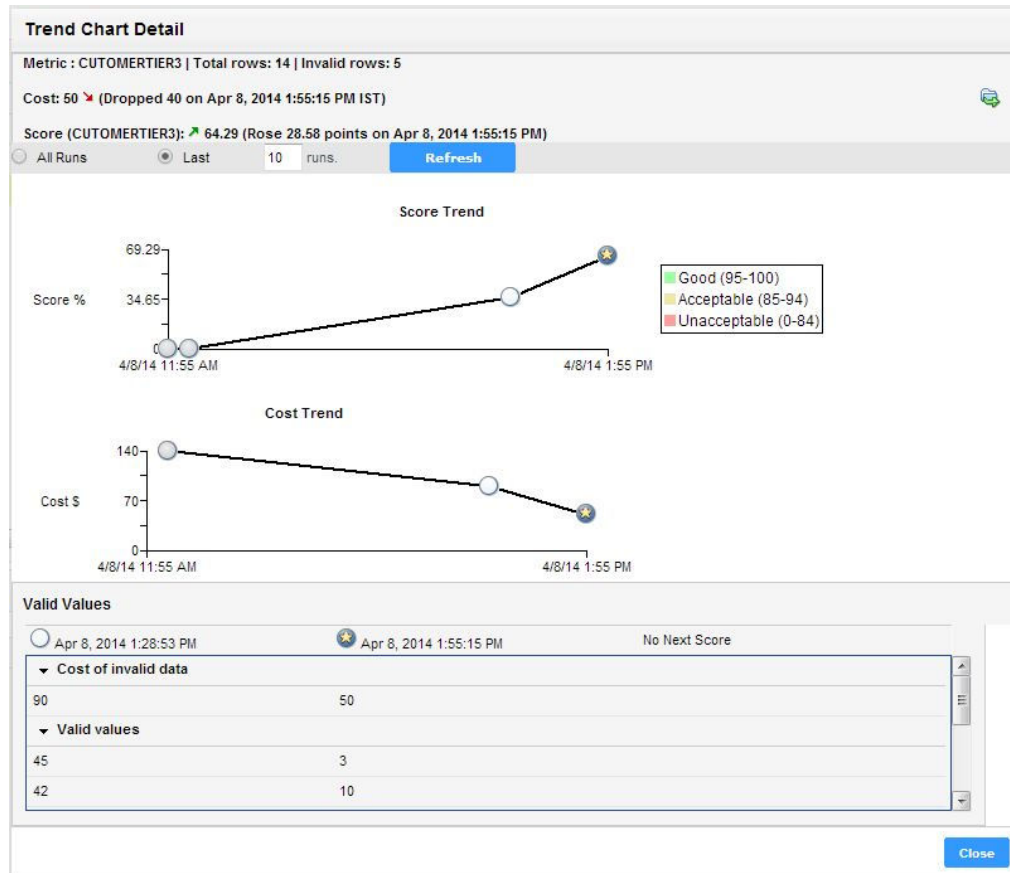
Sie können Trenddiagramme für jede Metrik anzeigen, um die Änderung des Scores oder der Kosten für ungültige Daten über einen bestimmten Zeitraum zu überwachen.

1. Wählen Sie im Arbeitsbereich **Bibliothek** das Projekt bzw. den Ordner aus, in dem die Scorecard enthalten ist.
2. Klicken Sie auf die Scorecard, um sie zu öffnen.
Die Scorecard wird im **Scorecards**-Arbeitsbereich angezeigt.
3. Wählen Sie in der Ansicht **Scorecard** eine Metrik aus.

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

Das Dialogfeld **Trenddiagramm-Detail** wird angezeigt.

Die folgende Abbildung zeigt das Dialogfeld **Trenddiagramm-Detail**:



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

Exportieren von Trenddiagrammen

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

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

Das Dialogfeld **Trenddiagramm-Details** wird angezeigt.

3. Klicken Sie auf das Symbol **Daten exportieren**.

Das Dialogfenster **Daten in Datei exportieren** wird geöffnet.

4. Geben Sie im Feld **Dateiname** einen Dateinamen ein. Sie können optional den Standarddateinamen verwenden.

Das Standarddateiformat ist Microsoft Excel.

5. Geben Sie im Feld **Ungültige Zeilen** die Anzahl der ungültigen Zeilen ein, die Sie exportieren möchten. Sie können einen Wert von maximal 100.000 im Feld eingeben.
6. Im Feld **Codepage** können Sie die Codepage der Datei auswählen.
7. Klicken Sie auf **OK**.

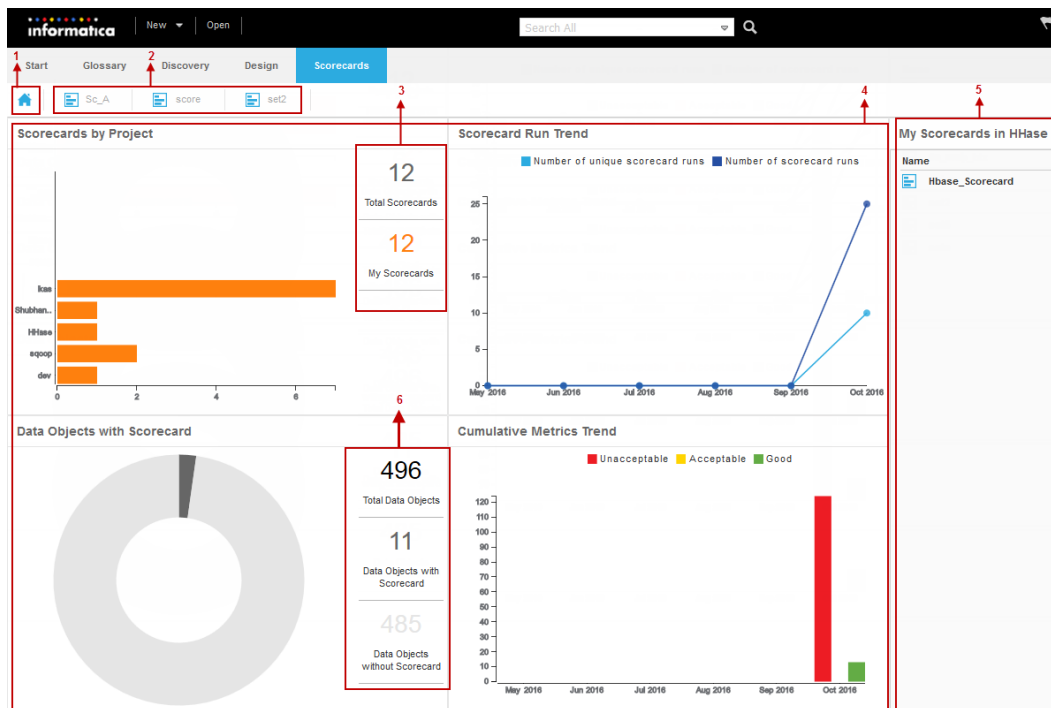
Scorecard-Dashboard in Informatica Analyst

Im Arbeitsbereich **Scorecards** in Informatica Analyst wird das Scorecard-Dashboard angezeigt. Im Scorecard-Dashboard können Sie die Datenobjekte mit Scorecards, den Trend bei der Scorecard-Ausführung für die letzten sechs Monate, Scorecards in einem Projekt, eine Zusammenfassung guter, akzeptabler und inakzeptabler Metriken für alle Scorecard-Ausführungen in einem Monat sowie den Objektlistenbereich anzeigen.

Das Scorecard-Dashboard auf Ihrem Computer wird nicht automatisch aktualisiert, wenn die Scorecards von anderen Benutzern geändert werden. Verwenden Sie die Funktionstaste F5 bzw. wechseln Sie zwischen Arbeitsbereichen oder Scorecard-Ergebnisregisterkarten, um das Scorecard-Dashboard zu aktualisieren.

Sie können die Daten als Datenreihe oder Datenpunkte in den Bereichen anzeigen. Die Datenpunkte werden als kleine undurchsichtige Kreise angezeigt. Die Datenreihe wird als horizontale oder vertikale Balken bzw. Segmente in den Diagrammen angezeigt.

Die folgende Abbildung zeigt das Scorecard-Dashboard und den Objektbereich im Arbeitsbereich **Scorecards**:



1. Symbol des Scorecard-Dashboards. Zeigt das Scorecard-Dashboard an.
2. Scorecard-Ergebnisregisterkarten. Zeigt die Scorecard-Ergebnisse für die geöffneten Scorecards an.
3. Legende im Bereich „Scorecards nach Projekt“. Zeigt die Gesamtzahl der Scorecards in allen Projekten und die Gesamtzahl der Scorecards mit Lesezugriff auf alle Projekte an.
4. Scorecard-Dashboard. Zeigt die Bereiche „Scorecards nach Projekt“, „Trend bei Scorecard-Ausführung“, „Datenobjekte mit Scorecards“ und „Trend kumulativer Metriken“ im Dashboard an.
5. Objektlistenbereich. Zeigt die Liste der Scorecards oder Datenobjekte an, die einer Legende, einer Datenreihe oder einem Datenpunkt im Diagramm zugeordnet sind.
6. Legende im Bereich „Datenobjekte mit Scorecard“. Zeigt die Gesamtanzahl der Datenobjekte sowie die Anzahl der Datenobjekte mit und ohne Scorecards an.

Nachdem Sie im Scorecard-Dashboard auf einen Datenpunkt oder eine Datenreihe geklickt haben, werden die Scorecards, die dem Datenpunkt oder der Datenreihe zugeordnet sind, im Objektlistenbereich angezeigt. Nachdem Sie im Objektlistenbereich auf eine Scorecard geklickt haben, werden die Scorecard-Ergebnisse auf einer Registerkarte im Arbeitsbereich **Scorecards** angezeigt. Im Objektlistenbereich werden die Scorecards angezeigt, für die Sie Lesezugriff haben.

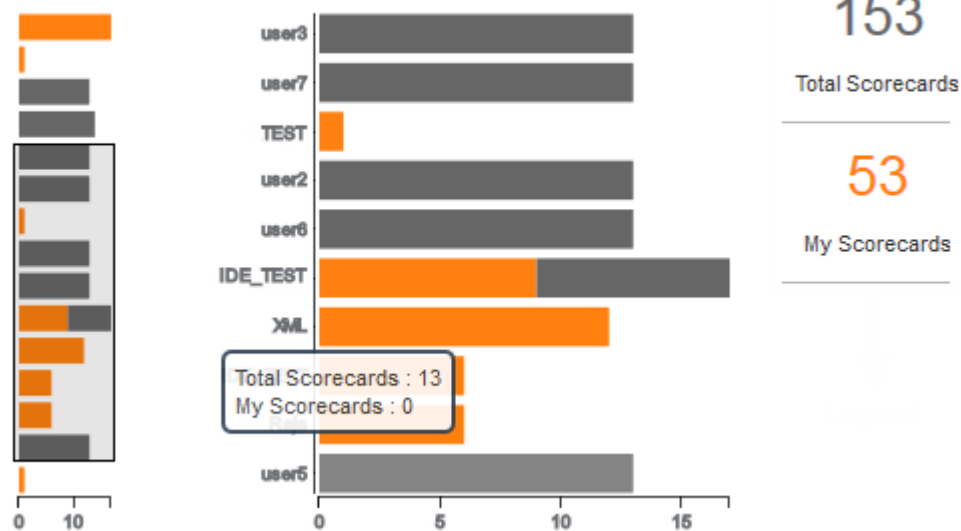
Scorecards nach Projekt

Im Bereich **Scorecards nach Projekt** werden die Projekte im Modellrepository gemeinsam mit der Anzahl der Scorecards pro Projekt als Balkendiagramm angezeigt. Ein Balken im Balkendiagramm entspricht einem Projekt. Auf der X-Achse des Diagramms wird die Anzahl der Scorecards angezeigt. Auf der Y-Achse befinden sich die Projekte mit Scorecards.

Die Scorecards in einem Projekt werden im Balkendiagramm in Grau und die Scorecards, für die Sie Lesezugriff haben, in Orange angezeigt. Im Abschnitt **Alle Scorecards** in der Legende wird die Gesamtzahl der Scorecards im Modellrepository angezeigt. Im Abschnitt **Eigene Scorecards** in der Legende wird die Anzahl der Scorecards, für die Sie Lesezugriff haben, im Modellrepository angezeigt.

Die folgende Abbildung zeigt den Bereich **Scorecards nach Projekt** im Scorecard-Dashboard:

Scorecards by Project



Sie können die folgenden Diagramme im Bereich anzeigen:

- **Detailliertes Diagramm.** Zeigt alle Projekte im Modellrepository mit Scorecards sowie die Anzahl der Scorecards in jedem Projekt an. Wenn die Anzahl der Projekte größer als 10 ist, wird im Bereich **Scorecards nach Projekt** ein Schieberegler angezeigt.
- **Miniaturdiagramm.** Zeigt im detaillierten Diagramm alle Projekte sowie die Anzahl der Scorecards pro Projekt innerhalb des Schiebereglers an.

Wenn Sie den Mauszeiger über das Miniaturdiagramm bewegen, wird die Gesamtzahl der Scorecards und die Anzahl der Scorecards, auf die Sie in einem Projekt zugreifen können, in einer Datenbeschriftung angezeigt.

Klicken Sie zum Anzeigen der Scorecards in einem Projekt, für die Sie Lesezugriff haben, auf den orangenen Bereich des horizontalen Balkens. Zur Anzeige aller Scorecards im Modellrepository, für die Sie Lesezugriff haben, klicken Sie im Balkendiagramm auf **Eigene Scorecards**. Die Scorecards werden im Objektlistenbereich angezeigt. Klicken Sie auf eine Scorecard im Objektbereich, um die Scorecard-Ergebnisse anzuzeigen.

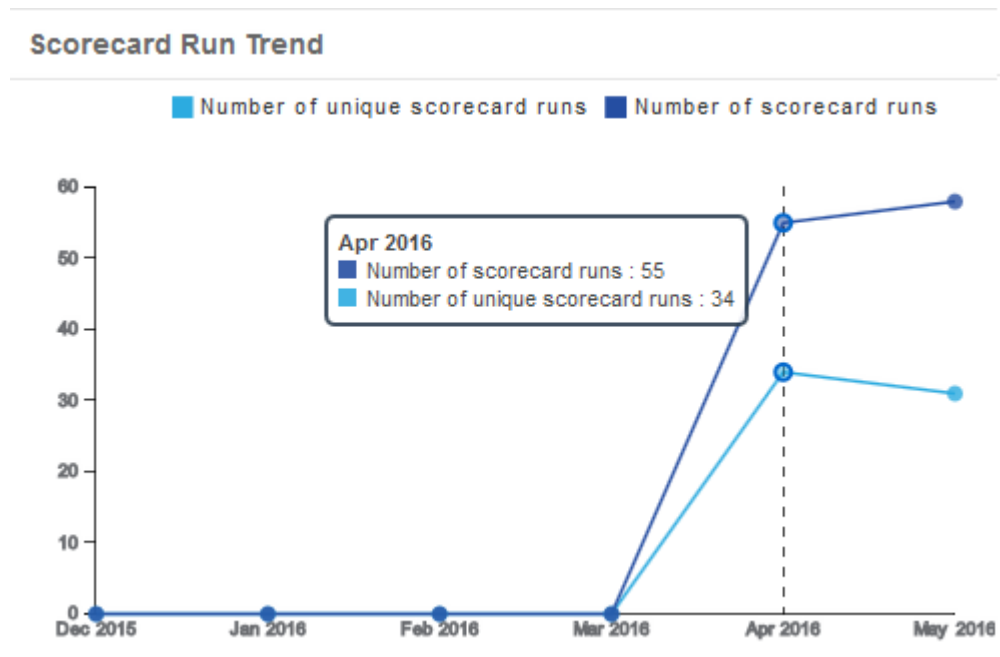
Trend bei der Scorecard-Ausführung

Im Bereich **Trend bei der Scorecard-Ausführung** wird der Scorecard-Ausführungstrend für den aktuellen Monat und die vergangenen fünf Monate als Liniendiagramm mit Markierungen angezeigt. Auf der X-Achse des Diagramms werden der aktuelle Monat und die letzten fünf Monate und auf der Y-Achse die Anzahl der Scorecards angezeigt. Bei einer Markierung handelt es sich um einen Datenpunkt im Liniendiagramm. Wenn Sie den Mauszeiger über eine Markierung im Diagramm bewegen, wird eine Zusammenfassung der Scorecard-Ausführungen für den entsprechenden Monat in einer Datenbeschriftung angezeigt.

Sie können die folgenden Markierungen im Bereich anzeigen:

- **Anzahl der Scorecard-Ausführungen.** Die Markierung zeigt die Gesamtzahl der Scorecard-Ausführungen im Monat an.
- **Anzahl der eindeutigen Scorecard-Ausführungen.** Die Markierung zeigt die Gesamtzahl der eindeutigen Scorecard-Ausführungen im Monat an.

Die folgende Abbildung zeigt den Bereich **Trend bei der Scorecard-Ausführung** im Scorecard-Dashboard:



Wenn Sie im Bereich auf eine Markierung klicken, werden die der Markierung zugeordneten Scorecards im Objektlistenbereich angezeigt. Sie können die Scorecards anzeigen, für die Sie Lesezugriff haben. Klicken Sie im Objektlistenbereich auf eine Scorecard, um die Scorecard-Ergebnisse im Arbeitsbereich **Scorecards** anzuzeigen.

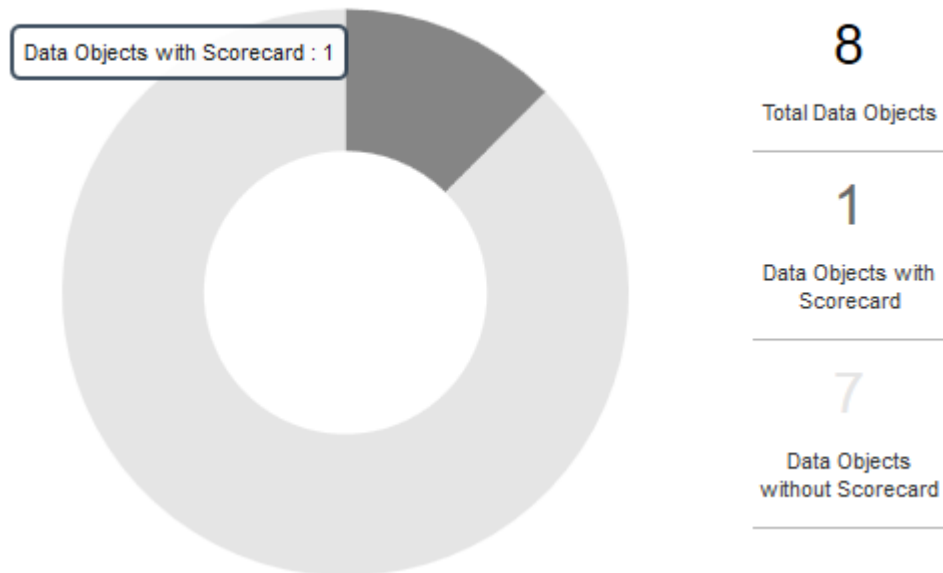
Datenobjekte mit Scorecards

Im Bereich **Datenobjekte mit Scorecards** wird ein Ringdiagramm angezeigt. Sie können die Anzahl der Datenobjekte mit und ohne Scorecards als Segmente im Diagramm anzeigen.

Wenn Sie den Mauszeiger über das Ringdiagramm bewegen, werden die dem jeweiligen Segment zugeordneten Daten in einer Datenbeschriftung angezeigt.

Die folgende Abbildung zeigt den Bereich **Datenobjekte mit Scorecards** im Scorecard-Dashboard:

Data Objects with Scorecard



In der Legende im Bereich **Datenobjekte mit Scorecards** werden die folgenden Datenstatistiken angezeigt:

- Gesamtzahl der Datenobjekte. Zeigt die Gesamtzahl der Datenobjekte im Ordner **Datenobjekte** im Bereich **Objekte** des Arbeitsbereichs **Bibliothek** an. Zu den Datenobjekten gehören logische und benutzerdefinierte Datenobjekte.
- Datenobjekte mit Scorecard. Zeigt die Anzahl der Datenobjekte mit Scorecards an.
- Datenobjekte ohne Scorecard. Zeigt die Anzahl der Datenobjekte ohne Scorecards an.

Nachdem Sie auf die Segmente im Ringdiagramm oder auf die Legenden **Datenobjekte mit Scorecard** und **Datenobjekte ohne Scorecard** geklickt haben, werden die Scorecards, die dem Segment im Ringdiagramm oder der Legende zugeordnet sind, im Objektlistenbereich angezeigt.

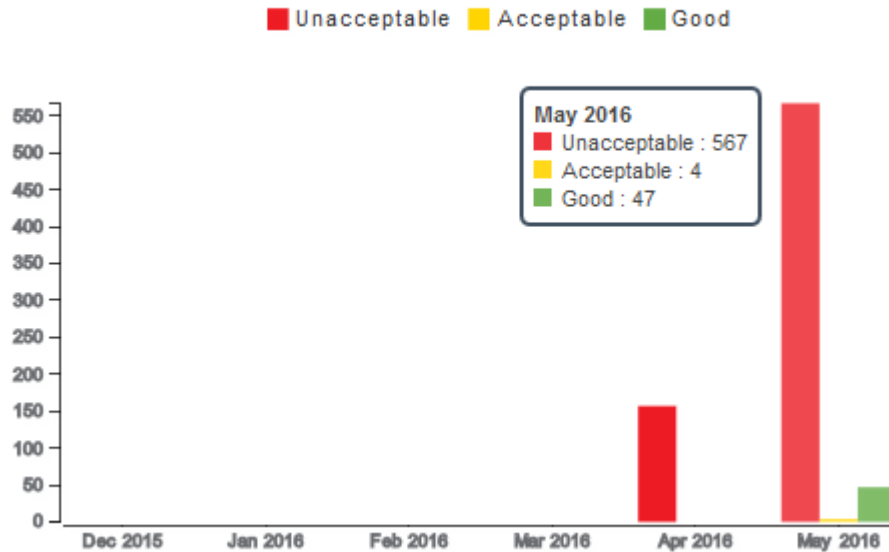
Trend kumulativer Metriken

Im Bereich **Trend kumulativer Metriken** werden Säulendiagramme angezeigt. Sie können die Zusammenfassung von guten, akzeptablen und inakzeptablen Metriken für alle Scorecard-Ausführungen in einem Monat als vertikale Balken im Diagramm anzeigen. Sie können den Bereich **Trend kumulativer Metriken** verwenden, um den Metriktrend für den aktuellen Monat und die vergangenen fünf Monate anzuzeigen und zu analysieren.

Wenn Sie den Mauszeiger über einen vertikalen Balken bewegen, wird die Metrikzusammenfassung für den Monat in einer Datenbeschriftung angezeigt. Wenn Sie im Bereich auf einen vertikalen Balken klicken, werden die zugehörigen Scorecards im Objektlistenbereich angezeigt. Sie können die Scorecards anzeigen, für die Sie Lesezugriff haben. Klicken Sie auf eine Scorecard im Objektlistenbereich, um die Scorecard-Ergebnisse anzuzeigen.

Die folgende Abbildung zeigt den Bereich **Trend kumulativer Metriken** im Scorecard-Dashboard:

Cumulative Metrics Trend



Im Objektlistenbereich werden unter Umständen einige Scorecards in der Liste der inakzeptablen Metriken und in der Liste der guten Metriken angezeigt, wenn sich der Scorecard-Trend innerhalb eines Monats ändert. Öffnen Sie zum Analysieren der Metrik die Scorecard und zeigen Sie die Scorecard-Ergebnisse an.

Beispiel

Sie arbeiten als Gebietsleiter für eine Reihe von Einzelhandelsgeschäften im US-Bundesstaat Kalifornien. Sie erstellen die Scorecard „Sales_SC“ in der Tabelle „Umsatz“.

Sie richten die folgenden Grenzwerte für die Metrik „Sales_amt“ in der Scorecard „Sales_SC“ ein:

- Inakzeptabel = 0 % bis 40 %
- Akzeptabel = 41 % bis 89 %
- Gut = 90 % bis 100 %

Verwenden Sie für die tägliche Erfassung der Verkaufsdaten den Scheduler-Dienst, um die Scorecards jeden Abend in der Tabelle „Umsatz“ auszuführen. Sie bereiten für das Management einen Bericht für den Monat Mai vor und verwenden das Scorecard-Dashboard zum Überprüfen des Berichts. Wenn Sie im Diagramm auf die horizontalen Balken für inakzeptable und gute Metriken klicken, wird die Scorecard „Sales_SC“ für den Monat Mai im Objektlistenbereich angezeigt.

Wenn Sie die Umsätze des Monats analysieren, bildet sich folgender Trend heraus:

1. Vom 1. bis zum 25. Mai liegt die Metrik „Sales_amt“ unter 40 % und wird als inakzeptable Metrik markiert.
2. In der letzten Maiwoche liegt die Metrik „Sales_amt“ aufgrund einer Umsatzsteigerung bei 98 % und wird als gute Metrik markiert.

Scorecard-Exportdateien in Informatica Analyst

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

Zusammenfassung, Trenddiagramme, ungültige Zeilen und Scorecard-Eigenschaften in mehreren Arbeitsblättern angezeigt.

Wenn Sie eine Scorecard exportieren, können Sie die folgenden Optionen konfigurieren:

Data

Wählen Sie eine der folgenden Optionen zum Exportieren der Daten aus:

- Alle. Exportiert Scorecard-Zusammenfassungen, Trenddiagramme, ungültige Zeilen und Scorecard-Eigenschaften in die Microsoft Excel-Datei.
- Zusammenfassungsansicht. Exportiert Scorecard-Zusammenfassungen, Trenddiagramme und Scorecard-Eigenschaften in die Microsoft Excel-Datei.

Ungültige Zeilen

Geben Sie die Anzahl der ungültigen Zeilen ein, die exportiert werden sollen. Sie können einen Wert von maximal 100.000 im Feld eingeben. Standardmäßig wird 100 im Feld angezeigt. Das Analyst Tool exportiert maximal 100 ungültige Zeilen pro Metrik in das Arbeitsblatt.

Wenn Sie mehr als 100 ungültige Zeilen pro Metrik exportieren möchten, führt der Datenintegrationsdienst die folgenden Schritte zum Exportieren der verbleibenden Daten durch:

1. Erstellt einen Ordner für die Scorecard im Verzeichnis `<INFA_HOME>/tomcat/bin/reject`. Der Datenintegrationsdienst verwendet für den Ordner basierend auf dem Namen der Scorecard und der Uhrzeit der Dateierstellung das Format `<scorecard_name>_HH_MM_SS`. Lauten der Name der Scorecard beispielsweise SD1 und die Uhrzeit der Dateierstellung 2:23:15, weist der Datenintegrationsdienst dem Ordner den Namen `SD1_02_23_15` zu.
2. Erstellt einen Unterordner für jede Metrik in der Scorecard. Weist die Scorecard SD1 beispielsweise Metriken mit den Bezeichnungen M1, M2 und M3 auf, erstellt der Datenintegrationsdienst drei Unterordner mit den Namen M1, M2 und M3.
3. Erstellt Microsoft Excel-Dateien im Unterordner der Metriken. Die Dateien werden basierend auf dem Namen der Metrik mit einem inkrementellen Wert im Format `<metric_name>_<IncrementalNumberStartingWith0>` benannt. Die zuletzt erstellte Datei verwendet das Format `<metric_name>_Remaining`. Jede Excel-Datei kann maximal 10.000 ungültige Zeilen enthalten. Weist Metrik M1 beispielsweise 30.000 ungültige Zeilen auf, erstellt der Datenintegrationsdienst drei Microsoft Excel-Dateien mit den Bezeichnungen M1_0, M1_1 und M1_Remaining und speichert 10.000 ungültige Zeilen pro Datei.
4. Der Datenintegrationsdienst wiederholt Schritt 3 für alle weiteren Metriken in der Scorecard.

Codepage

Wählen Sie eine Codepage aus.

Exportieren von Scorecard-Ergebnissen aus Informatica Analyst

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

1. Öffnen Sie eine Scorecard.
2. Klicken Sie auf **Aktionen > Daten exportieren**.
Das Dialogfeld **Daten in Datei exportieren** wird geöffnet.
3. Geben Sie einen Dateinamen ein. Optional können Sie den Standarddateinamen verwenden.
Das Standarddateiformat ist Microsoft Excel.
4. Wählen Sie die Codepage der Datei aus.

5. Klicken Sie auf **OK**.

Scorecard-Exportergebnisse in Microsoft Excel

Beim Exportieren der Scorecard-Ergebnisse speichert das Analyst Tool die Informationen in mehreren Arbeitsblättern in einer Microsoft Excel-Datei. Die Scorecard-Zusammenfassung, Trenddiagramme, ungültige Zeilen und Scorecard-Eigenschaften werden in der Datei als Arbeitsblätter angezeigt. Das Analyst Tool speichert die Datei im XLSX-Format.

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

Tab	Beschreibung
Scorecard-Zusammenfassung	Zusammenfassung der exportierten Scorecard-Ergebnisse. Zu den Informationen gehören der Scorecard-Name, die Gesamtanzahl der Zeilen für jede Spalte, die Anzahl der ungültigen Zeilen, der Score und die Metrikgewichtung.
Trenddiagramm	Trenddiagramme für Scores.
Ungültige Zeilen	Die Details der ungültigen Zeilen für jede Spalte. Das Analyst Tool exportiert maximal 100 Zeilen in das Arbeitsblatt. Das Arbeitsblatt für ungültige Zeilen wird angezeigt, wenn Sie die Option Daten > Alle im Dialogfeld Daten in eine Datei exportieren auswählen.
Eigenschaften	Scorecard-Eigenschaften, wie z. B. Name, Typ, Beschreibung und Speicherort.

Scorecard-Benachrichtigungen

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

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

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

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

Vorlage für Benachrichtigungs-E-Mail

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

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

Tag	Beschreibung
ScorecardName	Name der Scorecard
ObjectURL	Ein Hyperlink zur Scorecard. Sie müssen Benutzernamen und Kennwort angeben.
MetricGroupName	Name der Metrikgruppe, zu der die Metrik gehört
CurrentWeightedAverage	Gewichteter Durchschnittswert für die Metrikgruppe im aktuellen Scorecard-Lauf
CurrentRange	Der Score-Bereich, z. B. Inakzeptabel, Akzeptabel und Gut, für die Metrikgruppe im aktuellen Scorecard-Lauf
PreviousWeightedAverage	Gewichteter Durchschnittswert für die Metrikgruppe im vorherigen Scorecard-Lauf
PreviousRange	Der Score-Bereich, z. B. Inakzeptabel, Akzeptabel und Gut, für die Metrikgruppe im vorherigen Scorecard-Lauf
MetricName	Name der Metrik.
MetricGroupName	Name der Metrikgruppe.
CurrentScore	Score basierend auf der aktuellen Scorecard-Ausführung.
CurrentRange	Score-Bereich, in dem der aktuelle Score basierend auf der aktuellen Scorecard-Ausführung verbleibt.
PreviousScore	Score basierend auf der vorherigen Scorecard-Ausführung.
PreviousRange	Score-Bereich basierend auf der vorherigen Scorecard-Ausführung.
CurrentCost	Kosten für ungültige Daten in der Metrik basierend auf der aktuellen Scorecard-Ausführung.
PreviousCost	Kosten für ungültige Daten in der Metrik basierend auf der vorherigen Scorecard-Ausführung.
ColumnName	Name der Quellspalte, die der Metrik zugewiesen ist
ColumnType	Typ der Quellspalte

Tag	Beschreibung
RuleName	Name der Regel
RuleType	Typ der Regel
DataObjectName	Name des Quelldatenobjekts

Einrichten der Scorecard-Benachrichtigungen

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

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

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

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

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

Konfigurieren von globalen Einstellungen für Scorecard-Benachrichtigungen

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

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

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

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

Die Standard-E-Mail-ID des Absenders ist die in den SMTP-Eigenschaften der Domäne konfigurierte **E-Mail-Adresse des Absenders**.

8. Geben Sie den Text für die Betreffzeile der E-Mail ein.
9. Geben Sie im Feld **Text** den Einleitungs- und Schlusstext der E-Mail-Nachricht ein.
10. Um die globalen Einstellungen anzuwenden, wählen Sie **Einstellungen auf alle Metriken und Metrikgruppen anwenden** aus.
11. Klicken Sie auf **OK**.

Scorecard-Herkunft

Das Scorecard-Herkunftsdiagramm für eine Metrik oder Metrikgruppe zeigt den Ursprung der Daten, beschreibt den Pfad und verdeutlicht den Datenfluss. Mithilfe der Scorecard-Herkunft können Sie die Ursache einer inakzeptablen Score-Varianz bei Metriken oder Metrikgruppen analysieren. Die Scorecard-Herkunft können Sie im Analyst-Tool anzeigen lassen.

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

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

Hinweis: Der Name der Ressourcen-Datei, die Sie in Metadata Manager erstellen und laden, muss der folgenden Benennungskonvention folgen: <MRS-Name>_<Projektname>. Weitere Informationen zum Erstellen und Laden einer Ressourcendatei finden Sie im *Benutzerhandbuch für Informatica PowerCenter Metadata Manager*.

4. Öffnen Sie die Scorecard im Analyst-Tool und wählen Sie eine Metrik oder Metrikgruppe.
5. Lassen Sie sich die Scorecard-Herkunft anzeigen.

Anzeigen der Scorecard-Herkunft in Informatica Analyst

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

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

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

2. Wählen Sie in der Ansicht **Scorecard** eine Metrik oder Metrikgruppe.
3. Wählen Sie per Rechtsklick die Option **Herkunft anzeigen**.

Das Scorecard-Herkunftsdiagramm erscheint in einem neuen Fenster.

Wichtig: Wenn Sie eine Ressource nicht in Metadata Manager mit einer exportierten XML-Datei der Scorecard-Objekte erstellen und laden, erscheint möglicherweise eine Fehlermeldung, die Sie darauf hinweist, dass die Ressource im Katalog nicht verfügbar ist. Weitere Informationen zum Exportieren einer XML-Datei für die Scorecard-Herkunft finden Sie unter ["Exportieren einer Ressourcendatei zur Scorecard-Herkunftskontrolle" auf Seite 182](#).

KAPITEL 12

Datenqualitätsfunktionen in Informatica Analyst

Dieses Kapitel umfasst die folgenden Themen:

- [Datendomänenenerkennung in Informatica Analyst - Übersicht, 111](#)
- [Datendomänenglossar in Informatica Analyst, 111](#)
- [Optionen der Datendomänenenerkennung in Informatica Analyst, 114](#)
- [Erstellen eines Spaltenprofils zur Ausführung der Datendomänenenerkennung in Informatica Analyst, 117](#)
- [Bearbeiten eines Spaltenprofils und einer Datendomänenenerkennung in Informatica Analyst, 118](#)
- [Ausführen eines Profils zur Durchführung von Datendomänenenerkennung, 119](#)
- [Datendomänenenerkennungsergebnisse in Informatica Analyst, 119](#)
- [Datendomänenenerkennungs-Exportdateien in Informatica Analyst, 121](#)

Datendomänenenerkennung in Informatica Analyst - Übersicht

Bei der Erstellung eines Profils zur Ausführung von Datendomänenenerkennung können Sie die Quellspalten und Datendomänen auswählen, mit denen Sie die Spaltendaten und Spaltennamen sowie die Stichprobenoptionen abgleichen möchten. Sie können ein Konformitätskriterium für die Datendomänenenerkennung auswählen und Nullwerte während der Datendomänenenerkennung ausschließen.

Sie können ein Profil mit einer Stichprobenoption und Filtern erstellen, um Datendomänenenerkennung durchzuführen. Beim Ausführen des Profils wenden Sie die Stichprobenoption und die Filter auf die Datenquelle an und erzeugen einen Datensatz. Der Datendomänenenerkennungs-Prozess verwendet den Datensatz zur Erkennung von Datendomänen.

Datendomänenglossar in Informatica Analyst

Im Datendomänenglossar werden alle Datendomänen und Datendomänengruppen aufgelistet. Sie können die Liste nach Datendomänen oder Datendomänengruppen sortieren. Verwenden Sie das Datendomänenglossar zum Suchen, Hinzufügen, Bearbeiten und Entfernen von Datendomänen und Datendomänengruppen. Sie können die mit den Datendomänen verknüpfte Regellogik im Developer Tool anzeigen und ändern.

Erstellen einer Datendomänengruppe in Informatica Analyst

In einer Datendomänengruppe werden Datendomänen in bestimmten Gruppen angeordnet, wie beispielsweise Personal Health Information (PHI), Personally Identifiable Information (PII) oder jede andere konzeptionelle Gruppe, die für das Projekt relevant ist.

1. Klicken Sie auf **Verwalten > Datendomänenglossar**.
Das Datendomänenglossar wird in einer Registerkarte geöffnet, in der die aktuellen Datendomäne und Datendomänengruppen aufgelistet werden.
2. Klicken Sie im Navigator auf **Aktionen > Neu > Datendomänengruppe**.
Das Dialogfeld **Datendomänengruppe** wird angezeigt.
3. Geben Sie einen Namen und eine Beschreibung für die Datendomänengruppe ein.
4. Klicken Sie auf **Weiter**.
5. Wählen Sie im Bereich **Verfügbare Datendomänen** die Datendomänen aus, die Sie der Datendomänengruppe hinzufügen möchten, und klicken Sie auf **Hinzufügen**.
Das Analyst Tool verschiebt die ausgewählten Datendomänen in den Bereich **Ausgewählte Datendomänen**.
6. Klicken Sie auf **Fertig stellen**.
Das Analyst Tool fügt die Datendomänengruppe dem Datendomänenglossar hinzu.

Erstellen einer Datendomäne in Informatica Analyst

Sie können Datendomänen erstellen, sie zum Datendomänenglossar hinzufügen und Datendomänen in einer oder mehreren Datendomänengruppen gruppieren. Zum Erstellen einer Datendomäne können Sie vordefinierte Datenregeln und Spaltennamenregeln verwenden oder Datendomänen aus den Werten und Mustern in den Spaltenprofilergenernissen generieren.

Wenn Sie eine Datendomäne erstellen, kopiert das Analyst Tool Regeln und andere abhängige Objekte, die mit der Datendomäne verknüpft sind, in das Datendomänenglossar. Um eine mit der Datendomäne verknüpfte Regel zu bearbeiten, müssen Sie Änderungen an der ursprünglichen Regel vornehmen. Dann können Sie die geänderte Regel wieder mit der Datendomäne verknüpfen.

1. Klicken Sie auf **Verwalten > Datendomänenglossar**.
Das Datendomänenglossar wird in einer Registerkarte geöffnet, in der die aktuellen Datendomäne und Datendomänengruppen aufgelistet werden.
2. Klicken Sie im Navigator auf **Aktionen > Neu > Datendomäne**.
Das Dialogfeld **Datendomäne** wird angezeigt.
3. Geben Sie einen Namen und eine Beschreibung für die Datendomäne ein.
4. Klicken Sie auf das Kontrollkästchen **Datenregel**, um Datendomänen anhand der Spaltendaten zu erkennen. Sie können auch das Kontrollkästchen **Spaltennamenregel** aktivieren, um Datendomänen anhand der Spaltenüberschriften in der Datenquelle zu erkennen.
Die Schaltfläche **Auswählen** wird aktiviert.
5. Klicken Sie auf **Auswählen**, um das Dialogfeld **Regel auswählen** zu öffnen.
6. Wählen Sie eine passende Regel aus und klicken Sie auf **OK**.
Die ausgewählten Regeln erscheinen in den Feldern **Datenregel** und **Spaltennamenregel**.
7. Klicken Sie auf **Weiter**.

8. Wählen Sie im Bereich **Verfügbare Datendomänengruppen** die Datendomänengruppen aus, in die Sie die Datendomäne aufnehmen möchten, und klicken Sie auf **Hinzufügen**.

Das Analyst Tool verschiebt die ausgewählten Datendomänengruppen in den Bereich **Ausgewählte Datendomänengruppen**.

9. Klicken Sie auf **Fertig stellen**.

Das Analyst Tool fügt die Datendomäne dem Datendomänenglossar hinzu.

Erstellen einer Datendomäne aus Profilergebnissen in Informatica Analyst

Führen Sie ein Spaltenprofil aus, um die Werte und Muster der Datenquelle anzuzeigen. Sie können dann die Profilergebnisse überprüfen und eine Datendomäne davon erstellen.

1. Führen Sie ein Spaltenprofil aus, um deren Ergebnisse anzuzeigen.
Die Profilergebnisse werden in der Zusammenfassungsansicht angezeigt.
2. Klicken Sie in der Zusammenfassungsansicht auf eine Spalte, um die Spaltenergebnisse in der Detailansicht anzuzeigen.
3. Wählen Sie im Bereich **Werte** oder **Muster** den Wert bzw. das Muster aus, auf dessen Grundlage Sie eine Datendomäne erstellen möchten.
4. Klicken Sie mit der rechten Maustaste auf den Wert bzw. das Muster und wählen Sie dann **Datendomäne erstellen** aus.
Das Dialogfeld **Datendomäne erstellen** wird angezeigt.
5. Geben Sie den Namen der Datendomäne und eine optionale Beschreibung ein.
6. Klicken Sie auf **Erstellen**.
Die Datendomäne wird dem Datendomänenglossar hinzugefügt.

Suchen von Datendomänengruppen in Informatica Analyst

Sie können im Datendomänenglossar nach bestimmten Datendomänen und Datendomänengruppen suchen. Sie können zwischen den Ansichten **Datendomäne** und **Datendomänengruppe** wählen, um die Liste der Datendomänen im Datendomänenglossar anzuzeigen.

Nehmen wir an, Sie haben eine Datendomäne **Zipcode**, die Sie der Datendomänengruppe **PII** hinzugefügt haben. Sie können auf folgende Weise weitere Informationen zu Zipcode und seiner Datendomänengruppe PII finden:

Suchen nach Datendomänen.

Geben Sie in das Textfeld oben im Navigator einen Teil des Datendomänennamens ein, wie z. B. **zip** oder **code**. Wenn Sie sich in der Ansicht **Datendomänengruppe** befinden, listet das Analyst Tool **PII** auf, wobei es sich um die Datendomänengruppe handelt, die **Zipcode** enthält. Wenn Sie sich in der Ansicht **Datendomäne** befinden, listet das Analyst-Tool alle Datendomänen auf, in denen die Suchzeichenfolge **Zip** oder **code** enthalten ist.

Hinweis: Die Groß-/Kleinschreibung spielt bei der Suche keine Rolle.

Anzeigen aller Datendomänengruppen und der darin enthaltenen Datendomänen.

Klicken Sie im Navigator auf **Datendomänengruppenansicht anzeigen**.

Anzeigen aller Datendomänen.

Klicken Sie im Navigator auf **Datedomänennansicht anzeigen**.

Anzeigen der Eigenschaften einer Datendomäne.

Stellen Sie sicher, dass Sie sich in **Datendomänenansicht anzeigen** befinden. Klicken Sie im Navigator auf **Zipcode**, um dessen Eigenschaften im rechten Fensterbereich anzuzeigen. Sie können Folgendes einsehen: Name, Typ, Beschreibung, zugehörige Regeln und Domänenengruppen, zu denen es gehört - in diesem Fall **PII**.

Anzeigen der Eigenschaften einer Datendomänenengruppe.

Stellen Sie sicher, dass Sie sich in **Datendomänenengruppenansicht anzeigen** befinden. Klicken Sie im Navigator auf **PII**, um dessen Eigenschaften im rechten Fensterbereich anzuzeigen. Sie können Folgendes einsehen: Name, Typ, Beschreibung und die Liste der Datendomänen, darunter **Zipcode**, in **PII**.

Aktualisieren des Datendomänenglossars.

Klicken Sie im Navigator auf **Aktionen > Löschen**. Das Datendomänenglossar zeigt die aktuelle Liste der Datendomänen oder Datendomänenengruppen basierend auf der Ansicht, in der Sie sich befinden.

Optionen der Datendomänenenerkennung in Informatica Analyst

Verwenden Sie die Optionen der Datendomänenenerkennung zur Auswahl der Spalten-, Datendomänen- und Rückschlussoptionen für die Datendomänenenerkennung. Mit den Rückschlussoptionen wird ausgewählt, ob die Datendomänenenerkennung auf einer Regel nach Spaltendaten, Spaltenname oder beiden ausgeführt werden soll.

Datendomänen-Spaltenauswahl in Informatica Analyst

Sie können im Bildschirm **Einstellungen angeben** auf **Bearbeiten** klicken, um die Spalten auszuwählen, die als Teil der Datendomänenenerkennung ausgeführt werden sollen. Im Bildschirm **Quelle auswählen** des Profilassistenten können Sie alle Spalten in der Datenquelle anzeigen. Sie können unterschiedliche Spalten für das Spaltenprofil und die Datendomänenenerkennung auswählen.

In der folgenden Tabelle werden die Eigenschaften für die Datendomänenenerkennung im Dialogfeld **Bearbeiten** beschrieben:

Option	Beschreibung
Name	Zeigt den Spaltennamen an.
Typ	Zeigt den dokumentierten Datentyp der Spalte an.
Genauigkeit	Zeigt die maximale Genauigkeit für die Spalte an.
Skalierung	Zeigt die Skalierung der Spalte an.
Auf Null setzbar	Gibt eine Spalte an, die Nullwerte enthalten kann.
Schlüssel	Gibt an, ob die Spalte als Primär- oder Fremdschlüssel dokumentiert ist.

Datendomänenauswahl in Informatica Analyst

Im Bereich **Datendomäne** des Bildschirms **Einstellungen angeben** werden alle Datendomänen aus dem Datendomänen glossar aufgelistet. Sie können die Datendomänen wählen, die im Rahmen der Datendomänenerkennung ausgeführt werden sollen.

In der folgenden Tabelle werden die Eigenschaften der **Datendomäne** für die Datendomänenerkennung beschrieben:

Option	Beschreibung
Name	Zeigt den Datendomännennamen an. Sie können eine oder mehrere Datendomänen bzw. eine Datendomänengruppe auswählen.
Beschreibung	Zeigt die Beschreibung für die Datendomäne an.
DomainGroups	Zeigt den Namen der Datendomänengruppe an, zu der die Datendomäne gehört.

Ableitungsoptionen für Datendomänen in Informatica Analyst

Durch Rückschlussoptionen wird bestimmt, ob die Datendomänenerkennung für Spaltendaten, Spaltenname oder beide ausgeführt werden muss. Sie können die maximale Anzahl an Quellzeilen festlegen, die vom Profil analysiert werden können. Sie können ein Konformitätskriterium für die Datendomänenerkennung auswählen. Sie können Nullwerte aus der Datendomänenerkennung ausschließen. Im Bildschirm **Einstellungen angeben** des Profilassistenten können Sie die Optionen für die Datendomäneninterferenz festlegen.

In der folgenden Tabelle werden die Rückschlussoptionen für die Datendomänenerkennung beschrieben:

Option	Beschreibung
Daten	Führt das Profil für Spaltendaten aus.
Spalten	Führt das Profil für Spaltentitel aus.
Daten und Spalten	Führt das Profil für Spaltendaten und Spaltentitel aus.
Mindestprozentsatz der Zeilen	Der minimale Konformitätsprozentsatz der Zeilen im Datensatz, die für einen Datendomänenabgleich notwendig sind.
Mindestanzahl an Zeilen	Die Mindestanzahl der Zeilen im Datensatz, die für einen Datendomänenabgleich notwendig sind.
Nullwerte aus der Datendomänenerkennung ausschließen	Schließt die Nullwerte aus dem Datensatz für die Datendomänenerkennung aus.
Bearbeiten	Wählen Sie die Spalten für die Datendomänenerkennung aus..
Alle Zeilen	Führt das Profil für alle Zeilen aus der Quelle aus.
Erste Stichprobe	Wählen Sie die maximale Anzahl an Zeilen aus, für die das Profil ausgeführt werden kann. Das Analyst Tool wählt die Zeilen ab der ersten Zeile in der Quelle aus. Sie können maximal 2.147.483.647 Zeilen auswählen.

Option	Beschreibung
Zufallsstichprobe	Wählen Sie eine zufällige Stichprobe von Zeilen aus der Datenquelle aus. Sie können maximal 2.147.483.647 Zeilen auswählen.
Zufallsstichprobe (auto)	Das Analyst Tool wählt eine zufällige Stichprobe von Zeilen basierend auf der Größe der Datenquelle aus.
Genehmigte Datentypen und Datendomänen in den nachfolgenden Profilausführungen aus der Datentyp- und Datendomänenableitung ausschließen	Schließt den genehmigten Datentyp oder die genehmigte Datendomäne bei der nächsten Profilausführung aus der Datentyp- und Datendomänenableitung aus.

Minimaler Konformitätsprozentsatz

Sie können einen Mindestprozentsatz an Zeilen im Datensatz als Konformitätskriterium für die Datendomänenerkennung auswählen.

Der Konformitätsprozentsatz ist der Quotient aus der Anzahl übereinstimmender Zeilen geteilt durch die Gesamtanzahl der Zeilen.

Hinweis: Im Analyst Tool werden Nullwerte als nicht übereinstimmende Zeilen angesehen. Für Spalten mit einer großen Anzahl an Nullwerten wird unter Umständen keine Datendomänenableitung durchgeführt, es sei denn, Sie legen einen niedrigen Wert für den Mindestkonformitätsprozentsatz fest.

Beispiel

Sie haben eine Datenquelle mit 10.000 Zeilen, wobei 2.500 Zeilen in der Kommentarspalte Sozialversicherungsnummern aufweisen. Sie erstellen ein Spaltenprofil und eine Datendomänenerkennung und legen für eine Mindestanzahl an Zeilen als Konformitätskriterium 30 % fest. Wenn Sie das Profil ausführen, werden die Sozialversicherungsnummern in den Profilergebnissen nicht als abgeleitete Datendomäne angezeigt, da das minimale Konformitätskriterium bei 30 % der Zeilen oder 3.000 Zeilen in der Datenquelle liegt.

Mindestanzahl konformer Zeilen

Sie können eine Mindestanzahl an Zeilen im Datensatz als Konformitätskriterium für die Datendomänenerkennung auswählen.

Beispiel

Sie haben eine Datenquelle mit 10.000 Zeilen, wobei die Kommentarspalte E-Mail-Adressen in drei Zeilen aufweist. Sie erstellen ein Spaltenprofil und ein Datendomänenerkennungs-Profil und legen als Konformitätskriterium die Mindestanzahl der Zeilen auf 1 fest. Beim Ausführen des Profils zeigen die Profilergebnisse die E-Mail-Adresse als abgeleitete Datendomäne mit drei konformen Zeilen sowie die anderen abgeleiteten Datendomänen an.

Ausschließen von Nullwerten bei der Datendomänenerkennung

Sie können Nullwerte ausschließen, wenn Sie Datendomänenerkennung für eine Datenquelle durchführen. Wenn Sie den Mindestprozentsatz der Zeilen mithilfe der Option zum Ausschließen von Nullwerten auswählen, stellt der Konformitätsprozentsatz den Quotienten aus der Anzahl übereinstimmender Zeilen geteilt durch die Gesamtanzahl der Zeilen abzüglich der Nullwerte in der Spalte dar.

Der Datendomänenerkennungsprozess verläuft anders, wenn Sie die Option **Nullwerte aus der Datendomänenerkennung ausschließen** und die zahlreichen Stichprobenoptionen oder Filter auswählen.

In den folgenden Szenarien werden die Ergebnisse der Datendomänenenerkennung erläutert, wenn Sie die Option zum Ausschließen von Nullwerten mit einer Stichprobenoption und Filtern verwenden:

- Mit **Alle Zeilen** als Stichprobenoption, aber ohne Filter. Bei der Datendomänenenerkennung werden alle Nullwerte in der Spalte ignoriert.
- Mit einer Stichprobenoption, aber ohne Filter. Bei der Datendomänenenerkennung werden alle Nullwerte in den Stichprobendaten ignoriert und die verbleibenden Stichprobendaten berücksichtigt.
- Mit **Alle Zeilen** als Stichprobenoption und Filtern. Bei der Datendomänenenerkennung werden alle Nullwerte in den gefilterten Daten ignoriert und die verbleibenden gefilterten Daten berücksichtigt.
- Mit einer Stichprobenoption und Filtern. Bei der Datendomänenenerkennung werden alle Nullwerte in den gefilterten Daten der Stichprobe ignoriert und die verbleibenden gefilterten Daten berücksichtigt.

Beispiel

Sie haben eine Datenquelle mit 10.000 Zeilen, von denen 3.000 Zeilen Sozialversicherungsnummern in der Kommentarspalte aufweisen. Erstellen Sie ein Spaltenprofil und eine Datendomänenenerkennung und wählen Sie folgende Optionen aus:

- Wählen Sie die Option **Nullwerte aus der Datendomänenenerkennung ausschließen** aus.
- Wählen Sie **Alle Zeilen** als Stichprobenoption aus.
- Wählen Sie die Option **Mindestprozensatz der Zeilen** aus und setzen Sie die Option auf 12 %.

Bei Ausführung des Profils wird das Profil für den Datensatz ausgeführt und die Nullwerte werden bei der Datendomänenenerkennung ignoriert.

Erstellen eines Spaltenprofils zur Ausführung der Datendomänenenerkennung in Informatica Analyst

Sie müssen mindestens eine Datendomäne erstellen, bevor Sie ein Spaltenprofil zur Ausführung der Datendomänenenerkennung im Analyst Tool erstellen können. Das Profil kann sowohl Spaltennamen als auch Spaltendaten erkennen, die vordefinierte Datendomänen entsprechen.

1. Klicken Sie im Arbeitsbereich **Erkennung** auf **Profil** oder wählen Sie **Neu > Profil** im Analyst Tool aus. Der Assistent **Neues Profil** wird eingeblendet.
2. Die Option **Einzelquelle** ist standardmäßig ausgewählt. Klicken Sie auf **Weiter**.
3. Geben Sie im Bildschirm **Allgemeine Eigenschaften angeben** einen Namen und eine optionale Beschreibung für das Profil ein. Wählen Sie im Feld „Speicherort“ das Projekt oder den Ordner aus, in dem das Profil erstellt werden soll. Klicken Sie auf **Weiter**.
4. Klicken Sie im Bildschirm **Quelle auswählen** auf **Auswählen**, um ein Datenobjekt auszuwählen, oder auf **Neu**, um ein Datenobjekt zu importieren. Klicken Sie auf **Weiter**.
5. Legen Sie im Bildschirm **Einstellungen angeben** die Ausführung eines Spaltenprofils, einer Datendomänenenerkennung oder eines Spaltenprofils und einer Datendomänenenerkennung fest. Standardmäßig ist die Option für das Spaltenprofil ausgewählt.
 - Wählen Sie **Datendomänenenerkennung ausführen** zur Durchführung der Datendomänenenerkennung aus. Wählen Sie die Optionen für Datendomänen im Bereich **Datendomäne** aus.
 - Wählen Sie **Spaltenprofil ausführen** und **Datendomänenenerkennung ausführen** aus, um das Spaltenprofil und die Datendomänenenerkennung auszuführen. Wählen Sie die Optionen für Datendomänen im Bereich **Datendomäne** aus.

Hinweis: Standardmäßig sind die für das Spaltenprofil ausgewählten Spalten auch auf die Datendomänenenerkennung anwendbar. Klicken Sie auf **Bearbeiten**, um Spalten für die Datendomänenenerkennung ungeachtet der für das Spaltenprofil ausgewählten Spalten zu aktivieren oder zu deaktivieren.

- Wählen Sie Daten, Spalten oder Daten und Spalten für die Datendomänenenerkennung aus.
 - Wählen Sie eine Stichprobenoption im Bereich **Profil ausführen für** aus.
 - Wählen Sie eine Drilldown-Option im Bereich **Drilldown** aus. Optional können Sie auf **Spalten auswählen** klicken, um Spalten für den Drilldown auszuwählen. Sie können Datentyp- und Datendomäneninferenz für Spalten mit genehmigtem Datentyp oder genehmigter Datendomäne weglassen.
 - Wählen Sie ein Konformitätskriterium aus. Danach steht die Option **Nullwerte aus der Datendomänenenerkennung ausschließen** zur Verfügung.
 - Wählen Sie **Nativ** oder **Hadoop** als Laufzeitumgebung aus. Sie können in der Hadoop-Laufzeitumgebung die Option „Blaze“ oder „Spark“ wählen. Klicken Sie bei Verwendung der Option „Blaze“ auf **Auswählen**, um eine Hadoop-Verbindung im Dialogfeld **Hadoop-Verbindung auswählen** auszuwählen. Klicken Sie bei Verwendung der Option „Spark“ auf **Auswählen**, um eine Hadoop-Verbindung im Dialogfeld **Hadoop-Verbindung auswählen** auszuwählen.
6. Im Bildschirm **Regeln und Filter angeben** können Sie Regeln und Filter für das Profil hinzufügen, bearbeiten oder löschen.
 7. Klicken Sie auf **Speichern und beenden**, um das Profil zu erstellen, oder auf **Speichern und ausführen**, um das Profil zu erstellen und auszuführen.

Bearbeiten eines Spaltenprofils und einer Datendomänenenerkennung in Informatica Analyst

Sie können die Eigenschaften eines Profils nach dessen Ausführung ändern. Wenn Sie ein Spaltenprofil im Rahmen der Datendomänenenerkennung ausgeführt haben, können Sie die Spaltenprofileinstellungen ändern.

1. Wählen Sie im Arbeitsbereich **Bibliothek** das Projekt mit dem Profil aus oder wählen Sie das Profil im Bereich **Objekte** aus.
2. Klicken Sie auf den Profilnamen.
Die Zusammenfassungsansicht wird im Arbeitsbereich **Erkennung** angezeigt.
3. Falls das Versionsverwaltungssystem aktiviert ist, klicken Sie auf **Aktionen > Auschecken**, um das Profil auszuchecken.
4. Klicken Sie auf **Aktionen > Profil bearbeiten**.
Der **Profilassistent** wird angezeigt.
5. Wählen Sie je nach den von Ihnen gewünschten Änderungen eine der folgenden Seitenoptionen aus:
 - **Allgemeine Eigenschaften angeben.** Ändern Sie die allgemeinen Eigenschaften, wie z. B. Name, Beschreibung und Speicherort.
 - **Quelle auswählen.** Wählen Sie eine andere passende Datenquelle und Spalten aus, für die Sie das Profil ausführen.

- **Einstellungen angeben.** Führen Sie das Spaltenprofil oder das Spaltenprofil und die Datendomänenenerkennung aus. Bearbeiten Sie die Datendomänen-, Stichproben- und Drilldown-Optionen.
 - **Regeln und Filter angeben.** Erstellen, bearbeiten oder löschen Sie Regeln und Filter.
6. Klicken Sie auf **Speichern und beenden**, um das Profil zu bearbeiten, oder auf **Speichern und ausführen**, um das Profil zu bearbeiten und auszuführen.
 7. Falls das Versionsverwaltungssystem aktiviert ist, müssen Sie die folgenden Aufgaben durchführen:
 - Klicken Sie auf **Speichern und beenden**, um die Bearbeitung des Profils abzuschließen.
 - Klicken Sie in der Zusammenfassungsansicht auf **Einchecken**, um das Profil einzuchecken.
 - Klicken Sie auf **Aktionen > Profil ausführen**, um das Profil auszuführen.

Ausführen eines Profils zur Durchführung von Datendomänenenerkennung

Führen Sie ein Profil im Rahmen der Datendomänenenerkennung aus, um die Spalten anzuzeigen, die den Mustern der Datendomänenenerkennungsregel entsprechen.

1. Wählen Sie im **Bibliotheksnavigator** das Projekt oder den Ordner mit dem Profil im Bereich „Projekte“ oder das Profil im Bereich „Objekte“ aus.
2. Klicken Sie auf **Aktionen > Öffnen**.
Die Zusammenfassungsansicht wird im Arbeitsbereich **Erkennung** angezeigt.
3. Klicken Sie auf **Aktionen > Profil ausführen**.
Das Analyst Tool führt eine Profilausführung durch und zeigt die Profilergebnisse in der Zusammenfassungsansicht an.
4. Klicken Sie in der Zusammenfassungsansicht auf eine Spalte, um die Spaltenergebnisse anzuzeigen.
Die Detailansicht wird angezeigt.

Datendomänenenerkennungsergebnisse in Informatica Analyst

Sie können die Ergebnisse der Datendomänenenerkennung in der Zusammenfassungs- und in der Detailansicht anzeigen.

Im Datendomänenfeld werden Statistiken über Spalten angezeigt, die mit Datendomänen übereinstimmen. In der Zusammenfassungsansicht können Sie die abgeleiteten Datendomänen mit dem Prozentsatz und der Anzahl der konformen Zeilen anzeigen.

In der Detailansicht können Sie die folgenden Aufgaben durchführen:

- Zeigen Sie die abgeleiteten Datendomänen mit dem Prozentsatz und der Anzahl der konformen Zeilen in einem horizontalen Balkendiagramm an.

- Führen Sie einen Drilldown der Ergebnisse nach konformen Zeilen, nicht konformen Zeilen und Nullwerten durch.
- Genehmigen Sie die Datendomäne, lehnen Sie sie ab oder setzen Sie sie zurück.
- Zeigen Sie abgelehnte Datendomänen an oder blenden Sie sie aus.
- Führen Sie eine Datendomänenerkennung für alle Zeilen in der Datenquelle aus, um abgeleitete Datendomänen zu ermitteln.

Genehmigen von Datendomänen

Sie können im Analyst-Tool mehrere Datendomänen genehmigen.

1. Wählen Sie im Arbeitsbereich **Bibliothek** das Projekt bzw. den Ordner aus, in dem das Profil enthalten ist.
2. Klicken Sie auf das Profil, um es zu öffnen.
Die Profilergebnisse werden in der Zusammenfassungsansicht angezeigt.
3. Klicken Sie auf die Spalte, für die Sie die Datendomäne genehmigen möchten.
Die Spaltenergebnisse werden in der Detailansicht angezeigt.
4. Wählen Sie in der Detailansicht im Bereich **Datendomäne** die Datendomäne aus. Klicken Sie auf **Aktionen** > **Genehmigen**.
Der Status der Spalte oder Datendomäne ändert sich auf „Genehmigt“.
5. Zum Wiederherstellen des abgeleiteten Status der Spalte oder Datendomäne wählen Sie die Datendomäne aus und klicken auf **Aktionen** > **Zurücksetzen**.

Ablehnen von Datendomänen

Beim Öffnen der Profilergebnisse zeigt das Analyst-Tool standardmäßig die genehmigten Datendomänen. Sie können die abgelehnten Datendomänen ein- oder ausblenden.

1. Wählen Sie im Bibliotheksnavigator das Projekt bzw. den Ordner aus, in dem das Profil enthalten ist.
2. Klicken Sie auf das Profil, um es zu öffnen.
Die Profilergebnisse werden in der Zusammenfassungsansicht angezeigt.
3. Klicken Sie auf die Spalte, für die Sie die Datendomäne ablehnen möchten.
Die Spaltenergebnisse werden in der Detailansicht angezeigt.
4. Für die Ablehnung abgeleiteter Datendomänen klicken Sie auf **Aktionen** > **Ablehnen**.
Das Analyst-Tool entfernt die abgelehnte Datendomäne aus den Datendomänenerkennungsergebnissen.
5. Zum Anzeigen der abgelehnten Datendomänen klicken Sie auf **Aktionen** > **Abgelehnte einblenden**.
6. Zum Ausblenden der abgelehnten Datendomänen klicken Sie auf **Aktionen** > **Abgelehnte ausblenden**.

Datendomänenenerkennungs-Exportdateien in Informatica Analyst

Beim Exportieren der Ergebnisse der Datendomänenenerkennung aus dem Analyst Tool können Sie den Dateinamen und den Codepage-Wert angeben. Sie können Datendomänenenerkennungsergebnisse in eine Microsoft Excel-Datei exportieren.

Die Microsoft Excel-Datei enthält mehrere Arbeitsblätter, die Erkennungsergebnisse basierend auf Spalten, Datendomänen und Datendomänenengruppen getrennt darstellen. Das Arbeitsblatt "Eigenschaften" zeigt Profileigenschaften wie Namen, Beschreibung, Typ, Speicherort, Datum und Uhrzeit der letzten Änderungen am Profil und eine Verknüpfung zum Profil an.

Datendomänenenerkennungsergebnisse in Microsoft Excel

Beim Exportieren der Ergebnisse der Datendomänenenerkennung in Microsoft Excel speichert das Analyst Tool die Spaltennamen, Namen passender Datendomänen, Konformitätskriterien und Nullwerte. Die Excel-Datei enthält auch die Gruppennamen der Datendomäne für jede Datendomäne und dokumentierte Datentypen für die Spalten.

Die folgende Tabelle beschreibt alle Arbeitsblätter in der Exportdatei:

Register	Beschreibung
Nach Spalten anzeigen	Datendomänenenerkennungsergebnisse, sortiert nach Datenquellenspalte.
Nach Datendomänen anzeigen	Datendomänenenerkennungsergebnisse sortiert nach Datendomäne.
Nach Datendomänenengruppen anzeigen	Datendomänenenerkennungsergebnisse, sortiert nach Datendomänenengruppe.
Eigenschaften	Basis-Profileigenschaften wie Name, Beschreibung, Typ, Speicherort, Datum und Uhrzeit der letzten Änderungen am Profil sowie eine Verknüpfung zum Profil.

Exportieren von Datendomänenenerkennungsergebnissen aus Informatica Analyst

Sie können die Datendomänenenerkennungsergebnisse in eine **XLSX**-Datei exportieren, so dass Sie die Daten in einer Datei anzeigen können und sie innerhalb des Unternehmens zur weiteren Verwendung verteilen können.

1. Führen Sie ein Profil zur Durchführung der Datendomänenenerkennung aus.
2. Klicken Sie in der Zusammenfassungsansicht bzw. Detailansicht auf **Aktionen > Daten exportieren**.
Das Dialogfenster **Daten in Datei exportieren** wird geöffnet.
3. Geben Sie den Dateinamen ein. Optional können Sie den Standarddateinamen verwenden.
4. Wählen Sie die Codeseite der Datei.
5. Klicken Sie auf **OK**.

KAPITEL 13

Enterprise-Erkennung in Informatica Analyst

Dieses Kapitel umfasst die folgenden Themen:

- [Enterprise-Erkennung in Informatica Analyst - Übersicht, 122](#)
- [Enterprise-Erkennungsprozess in Informatica Analyst, 123](#)
- [Konfigurationsoptionen für Enterprise-Erkennung, 123](#)
- [Erstellen eines Enterprise-Erkennungsprofils in Informatica Analyst, 125](#)
- [Bearbeiten der Enterprise-Erkennungsoptionen, 127](#)

Enterprise-Erkennung in Informatica Analyst - Übersicht

Die Enterprise-Erkennung ist der Prozess, Spaltenmetadaten und Datendomänen in mehreren Datenquellen über mehrere Schemas und externe relationale Verbindungen zu finden. Sie können die Enterprise-Erkennung sowohl für Datenquellen, die Sie in das Modellrepository importiert haben, und für Datenquellen von externen relationalen Verbindungen ausführen.

Als Datenanalytiker können Sie die Enterprise-Erkennung im Analyst Tool ausführen, um Rückschlüsse auf bestimmte Metadatenmerkmale in einer großen Anzahl von Datenquellen zu ziehen. Sie wollen vielleicht auch die Quelldaten anzeigen, die mit vordefinierten Datendomänen übereinstimmen. Sie können dann die abgeleiteten Enterprise-Erkennungsergebnisse wiederherstellen und die Daten für die Datenerkennungssuche und Datenqualitätsinitiativen vorbereiten. Im Analyst Tool generiert die Enterprise-Erkennung eine konsolidierte Ergebniszusammenfassung der Profilergebnisse.

Enterprise-Erkennungsergebnisse umfassen Spaltenprofilstatistiken wie Muster, eindeutige Werte und Spalten mit Datentypkonflikten. Die Datendomänenerkennung identifiziert die Quellspalten, die vordefinierten Datendomänen entsprechen.

Sie können ein Betriebssystemprofil in Informatica Analyst auswählen. Nach der Auswahl eines Betriebssystemprofils erstellt der Datenintegrationsdienst die Enterprise-Erkennungsprofile und führt sie basierend auf den Berechtigungen des Betriebssystembenutzers aus, den Sie im Betriebssystemprofil definiert haben.

Enterprise-Erkennungsprozess in Informatica Analyst

Sie können Enterprise-Erkennungsprofile erstellen, bearbeiten und löschen. Sie können ein Enterprise-Erkennungsprofil im Erkennungs-Arbeitsbereich ausführen. Sie müssen die Inferenzoptionen für Spaltenprofil- und Datendomänenenerkennung konfigurieren, bevor Sie das Enterprise-Erkennungsprofil ausführen.

Schließen Sie die folgenden Schritte ab, um die Enterprise-Erkennung im Analyst-Tool auszuführen:

1. Konfigurieren Sie die allgemeinen Eigenschaften für das Enterprise-Erkennungsprofil.
2. Wählen Sie die Datenobjekte, die Sie ins Enterprise-Erkennungsprofil aufnehmen möchten, aus dem Modellrepository aus.
3. Importieren Sie relationale Datenquellen aus externen Datenbankverbindungen.
4. Konfigurieren Sie die Dateninferenz- und Erkennungsoptionen für das Enterprise-Erkennungsprofil.
5. Speichern Sie die Änderungen und führen Sie das Enterprise-Erkennungsprofil aus.
6. Überwachen Sie die Profilausführung, und zeigen Sie, falls erforderlich, die Status der Profilaufgaben an, die das Analyst-Tool ausführt.
7. Überprüfen Sie die Zusammenfassung der Enterprise-Erkennungsergebnisse. Die Ergebnisse werden in den Bereichen **Zusammenfassung** und **Profile** angezeigt.

Konfigurationsoptionen für Enterprise-Erkennung

Zu den Konfigurationsoptionen für Enterprise-Erkennung gehören Datendomänenenerkennungsoptionen, Spaltenprofil-Stichprobenoptionen und allgemeine Profileigenschaften wie Name und Beschreibung.

Sie können auswählen, ob Sie ein Spaltenprofil oder ein Profil zur Ausführung der Datendomänenenerkennung ausführen möchten. Sie können auch sowohl ein Spaltenprofil als auch ein Profil zur Ausführung der Datendomänenenerkennung als Teil der Konfiguration auswählen.

Einstellungen für Datendomänenenerkennung

Die Einstellungen der Datendomänenenerkennung beinhalten die Auswahl, ob die Datendomänenenerkennung auf Spaltendaten, Spaltennamen oder auf Spaltendaten und Spaltennamen ausgeführt werden muss. Sie können die Datendomänen auswählen und festlegen, ob die Datendomänenenerkennung alle Zeilen in der Datenquelle verarbeiten muss. Sie können ein Konformitätskriterium für die Datendomänenenerkennung auswählen. Sie können Nullen aus der Datendomänenenerkennung ausschließen.

Die folgende Tabelle beschreibt die Einstellungen der Datendomänenenerkennung, die Sie für die Enterprise-Erkennung im Analyst-Tool konfigurieren können:

Option	Beschreibung
Datendomänenenerkennung aktivieren	Führt die Datendomänenenerkennung als Teil der Enterprise-Erkennung durch.
Datendomänenenerkennung für Daten ausführen	Führt Datendomänenenerkennung für Spaltendaten aus.
Datendomänenenerkennung für Spaltennamen ausführen	Führt Datendomänenenerkennung für die Namen aller Spalten aus.

Option	Beschreibung
Minimaler Konformitätsprozentsatz	Der minimale Konformitätsprozentsatz der Zeilen im Datensatz, die für einen Datendomänenabgleich notwendig sind. Der Konformitätsprozentsatz ist der Quotient aus der Anzahl übereinstimmender Zeilen geteilt durch die Gesamtanzahl der Zeilen. Hinweis: Im Analyst Tool werden Nullwerte als nicht übereinstimmende Zeilen angesehen.
Mindestanzahl konformer Zeilen	Die Mindestanzahl der Zeilen im Datensatz, die für einen Datendomänenabgleich notwendig sind.
Nullwerte aus der Datendomänenerkennung ausschließen	Schließt die Nullwerte aus dem Datensatz für die Datendomänenerkennung aus.
Spalten mit genehmigten Datendomänen ausschließen	Schließt Spalten mit genehmigten Datendomänen aus der Datendomäneninferenz der Profilausführung aus.
Alle Zeilen	Führt Datendomänenerkennung für alle Quellzeilen aus.
Erste	Die maximale Anzahl an Zeilen, für die ein Profil ausgeführt werden kann. Das Analyst Tool wählt Zeilen beginnend bei der ersten Zeile in der Quelle aus. Sie können maximal 2.147.483.647 Zeilen auswählen.

Spaltenprofileinstellungen

Durch die Stichprobenoptionen wird festgelegt, ob das Analyst Tool ein Spaltenprofil über allen Zeilen der Datenquellen oder über eine begrenzte Anzahl von Zeilen ausführt.

In der folgenden Tabelle werden die Spaltenprofileinstellungen beschrieben, die für ein Enterprise-Erkennungsprofil konfiguriert werden können:

Option	Beschreibung
Spalten-Profil aktivieren	Führt ein Spaltenprofil als Teil der Enterprise-Erkennung aus.
Genehmigte Datentypen und Datendomänen in den nachfolgenden Profilausführungen aus der Datentyp- und Datendomänenableitung ausschließen	Schließt den genehmigten Datentyp oder die genehmigte Datendomäne bei der nächsten Profilausführung aus der Datentyp- und Datendomänenableitung aus.

In der folgenden Tabelle wird die Laufzeitumgebungsoption beschrieben, die für ein Enterprise-Erkennungsprofil konfiguriert werden kann:

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

In der folgenden Tabelle werden die Stichprobenoptionen beschrieben, die für ein Enterprise-Erkennungsprofil konfiguriert werden können:

Option	Beschreibung
Alle Zeilen	Führt ein Spaltenprofil für alle Zeilen in der Datenquelle aus. Unterstützt in der nativen Laufzeitumgebung und in der Blaze- und Spark-Laufzeitumgebung.
Erste <Anzahl> Zeilen	Führt ein Profil auf Stichprobenzeilen aus, die sich am Anfang der Zeilen des Datenobjekts befinden. Sie können maximal 2.147.483.647 Zeilen auswählen. Unterstützt in der nativen Laufzeitumgebung und in der Blaze-Laufzeitumgebung.
Limit n <number> Zeilen	Führt ein Profil basierend auf der Anzahl der Zeilen im Datenobjekt aus. Wenn Sie ein Profil in der Hadoop-Validierungsumgebung ausführen, sammelt die Spark-Engine Stichproben von mehreren Partitionen des Datenobjekts und überträgt die Stichproben in einen einzigen Knoten, um den Stichprobenumfang zu berechnen. Die Stichprobenoption „Limit n“ unterstützt nur Oracle-, SQL Server- und DB2-Datenbanken. Sie können den erweiterten Filter nicht zusammen mit der Stichprobenoption „Limit n“ anwenden. Sie können maximal 2.147.483.647 Zeilen auswählen. Unterstützt in der Spark-Laufzeitumgebung.
Zufallsprozensatz	Führt ein Profil auf einem Prozentsatz von Zeilen im Datenobjekt aus. Unterstützt in der Spark-Laufzeitumgebung.

Erstellen eines Enterprise-Erkennungsprofils in Informatica Analyst

Sie können das Spaltenprofil und die Datendomänenerkennung als Teil der Enterprise-Erkennung in Informatica Analyst ausführen.

1. Klicken Sie im Arbeitsbereich **Erkennung** auf **Neu > Profil**.
Der Assistent **Neues Profil** wird eingeblendet.
2. Wählen Sie **Enterprise-Erkennung** aus. Klicken Sie auf **Weiter**.

Die Registerkarte **Allgemeine Eigenschaften angeben** wird angezeigt.

3. Geben Sie auf der Registerkarte **Allgemeine Eigenschaften angeben** einen Namen für das Enterprise-Erkennungsprofil und eine optionale Beschreibung ein. Wählen Sie im Feld „Speicherort“ das Projekt oder den Ordner aus, in dem das Profil erstellt werden soll. Klicken Sie auf **Weiter**.

Die Registerkarte **Datenobjekte auswählen** wird angezeigt.

4. Klicken Sie auf der Registerkarte **Datenobjekte auswählen** auf **Auswählen**.

Das Dialogfeld **Datenobjekte auswählen** wird angezeigt.

5. Wählen Sie im Dialogfeld **Datenobjekte auswählen** ein oder mehrere Datenobjekte aus, die dem Profil hinzugefügt werden sollen. Klicken Sie auf **Speichern**.

Die Datenobjekte werden im Bereich **Datenobjekte** angezeigt.

6. Klicken Sie auf **Weiter**.

Die Registerkarte **Ressourcen auswählen** wird geöffnet.

7. Klicken Sie auf der Registerkarte **Ressourcen auswählen** auf **Auswählen**, um die Registerkarte **Ressourcen auswählen** zu öffnen.

Sie können Daten aus mehreren relationalen Datenquellen importieren.

8. Wählen Sie auf der Registerkarte **Ressourcen auswählen** die Verbindungen, Schemas, Tabellen und Ansichten aus, die Sie in das Profil aufnehmen möchten. Klicken Sie auf **Speichern**.

Der linke Bereich im Dialogfeld listet alle internen und externen Verbindungen, Schemas und Ansichten in der Informatica-Domäne auf.

Die Ressourcen werden im Bereich **Ressource** angezeigt.

9. Klicken Sie auf **Weiter**.

Die Registerkarte **Einstellungen angeben** wird angezeigt.

10. Auf der Registerkarte **Einstellungen angeben** können Sie die Spaltenprofil- und Datendomänenerkennungsoptionen konfigurieren. Klicken Sie auf **Speichern und beenden**, um das Enterprise-Erkennungsprofil zu speichern, oder auf **Speichern und ausführen**, um das Profil auszuführen.

Sie können auf der Registerkarte **Einstellungen angeben** die folgenden Aufgaben durchführen:

- Aktivieren Sie die Datendomänenerkennung. Klicken Sie auf **Auswählen**, um die zu erkennenden Datendomänen im Dialogfeld **Datendomänen auswählen** auszuwählen. Die ausgewählten Datendomänen werden im Bereich **Datendomänen für Datendomänenerkennung** angezeigt.
- Führen Sie Datendomänenerkennung für Daten oder Spaltennamen bzw. für Daten und Spaltennamen aus.
- Wählen Sie alle Zeilen in der Datenquelle oder eine maximale Anzahl an Zeilen aus, für die Domänenerkennung ausgeführt werden soll.
- Wählen Sie einen minimalen Konformitätsprozentsatz aus oder geben Sie die Mindestanzahl an konformen Zeilen für die Datendomänenerkennung an.
- Aktivieren Sie Spaltenprofileinstellungen und wählen Sie alle Zeilen bzw. die ersten Zeilen in der Datenquelle für das Spaltenprofil aus. Sie können die Datentypinferenz für Spalten mit genehmigten Datentypen im Spaltenprofil ausschließen.
- Wählen Sie **Nativ**, **Blaze** oder **Spark** als Laufzeitumgebung aus. Wählen Sie nach Auswahl der Option **Blaze** oder **Spark** eine Hadoop-Verbindung zum Ausführen der Profile aus.

Sie können die Enterprise-Erkennungsergebnisse auf den Registerkarten **Zusammenfassung** und **Profile** anzeigen.

Bearbeiten der Enterprise-Erkennungsoptionen

Sie können Änderungen an den Enterprise-Erkennungsoptionen vornehmen, nachdem Sie die Enterprise-Erkennung ausführen. Sie können Sie das Profil umbenennen und die Datenobjektauswahl, Datendomänenauswahl und Rückschlussoptionen ändern.

1. Öffnen Sie ein Profil, das Sie ausgeführt haben, um die Enterprise-Erkennung auszuführen.
Die Profilergebnisse erscheinen im **Discovery-Workspace**.
2. Falls das Versionsverwaltungssystem aktiviert ist, klicken Sie auf **Aktionen > Auschecken**, um das Profil auszuchecken.
3. Klicken Sie auf **Profil bearbeiten**.
4. Aktualisieren Sie auf der Registerkarte **Allgemeine Eigenschaften angeben** nach Bedarf die Profileigenschaften.
5. Um Änderungen am Datenobjekt vorzunehmen, klicken Sie auf die Registerkarte **Datenobjekte auswählen**.
6. Um Änderungen an den externen Datenquellen der Enterprise-Erkennung vorzunehmen, klicken Sie auf die Registerkarte **Ressourcen auswählen**.
7. Um die Ableitungsoptionen der Datendomäne und Spaltenprofileinstellungen zu ändern, klicken Sie auf die Registerkarte **Einstellungen angeben**.
8. Wählen Sie zum Anwenden der Konfigurationsänderungen auf alle Datendomänen- und Spaltenprofilaufgaben im Enterprise-Erkennungsprofil die Option **Globale Einstellungen für alle Profile** aus. Wenn Sie diese Option nicht auswählen, gelten die an den Profileinstellungen vorgenommenen Änderungen nur für die Datenobjekte oder Ressourcen, die dem Profil neu hinzugefügt wurden.
Standardmäßig gelten die vorgenommenen Änderungen für die im Enterprise-Erkennungsprofil neu hinzugefügten Datenobjekte.
9. Um die Änderungen rückgängig zu machen, klicken Sie auf **Abbrechen**.
10. Klicken Sie auf **Speichern und ausführen**, um die Änderungen zu speichern und das Profil erneut auszuführen.
11. Falls das Versionsverwaltungssystem aktiviert ist, müssen Sie die folgenden Aufgaben durchführen:
 - Klicken Sie auf **Speichern und beenden**, um die Bearbeitung des Profils abzuschließen.
 - Klicken Sie in der Zusammenfassungsansicht auf **Einchecken**, um das Profil einzuchecken.
 - Klicken Sie auf **Aktionen > Profil ausführen**, um das Profil auszuführen.

KAPITEL 14

Enterprise-Erkennungsergebnisse in Informatica Analyst

Dieses Kapitel umfasst die folgenden Themen:

- [Ergebnisse der Enterprise-Erkennung in Informatica Analyst - Übersicht, 128](#)
- [Zusammenfassungsansicht, 128](#)
- [Datentypkonflikt, 131](#)
- [Profilansicht, 131](#)

Ergebnisse der Enterprise-Erkennung in Informatica Analyst - Übersicht

Sie können die Ergebnisse der Enterprise-Erkennung in den Ansichten **Zusammenfassung** und **Profile** ansehen.

Die Ansicht **Zusammenfassung** zeigt die Profilergebnisse und Datendomänenenerkennungsergebnisse an. Der Abschnitt **Datendomänenenerkennung** listet die Datendomänen auf, die Sie in die Profilausführung eingeschlossen haben, und die Anzahl der Spalten, die eine Datendomänenübereinstimmung haben. Der Abschnitt **Spalten-Profiling** zeigt Statistiken zu den Quellspalten. Sie können jede Profilergebniszeile anklicken, um die Details im rechten Fensterbereich der Ansicht **Zusammenfassung** anzuzeigen.

Zusammenfassungsansicht

Die Ansicht **Zusammenfassung** zeigt eine Zusammenfassung der Profilergebnisse und Datendomänenenerkennungsergebnisse. Sie können Sie die Datendomänennamen, die Übereinstimmungen in Spalten haben, und die Anzahl der Spalten mit Übereinstimmungen der Datendomänen anzeigen. Die Spaltenstatistik umfasst die Anzahl der Spalten für die 10 besten Musterübereinstimmungen, alle eindeutigen

Werte und alle Nullwerte. Zudem umfasst die Spaltenstatistik die Anzahl der Spalten mit Datentypkonflikten zwischen abgeleiteten und dokumentierten Datentypen.

Profilergebnisse Zusammenfassungsansicht

Die Zusammenfassungsansicht zeigt Enterprise-Erkennungsergebnisse in den Bereichen Datendomänenerkennung und Spalten-Profiling an.

Datendomänenerkennung

Die folgende Tabelle beschreibt die Spalten in den Datendomänenerkennungsergebnissen:

Spaltenname	Beschreibung
Name	Name der Datendomäne.
Gefunden in Spalten	Die Gesamtzahl der Spalten mit einem Datendomänenabgleich.
Profil	Name des Profils, das die übereinstimmende Spalte enthält.
Spaltenname	Name der übereinstimmenden Spalte.
% Datenkonformitätsprozentsatz	Der minimale Konformitätsprozentsatz der Zeilen, die für einen Datendomänenabgleich notwendig sind.
Verbindungsname	Name der relationalen Datenbankverbindung.
Quellenname	Name der Datenquelle.
Ableitungsstatus	Status der Datendomänenableitung. Folgende Statusangaben sind möglich: Akzeptiert , Abgelehnt und Abgeleitet .
% Null	Der Prozentsatz der Nullwerte für die Spalte.
Gesamtzeilen	Die Gesamtanzahl an Zeilen.
Konforme Zeilen	Die Mindestanzahl an Zeilen, die für einen Datendomänenabgleich notwendig sind.
Spaltennamenvergleich	Gibt an, ob der Spaltenname einem Datendomännennamen entspricht.
Dokumentierter Datentyp	Datentyp, der für die Spalte im Profilobjekt angegeben wurde.
Überprüft	Gibt die Überprüfung des Datendomänenvergleichs für alle Zeilen der Datenquelle an.
Letzte Ausführung	Datum und Uhrzeit der letzten Profilausführung.

Spalten-Profiling

Die folgende Tabelle beschreibt die Spalten in den Spaltenprofilerggebnissen:

Spaltenname	Beschreibung
Name	Name des Profilergebnistyps, wie beispielsweise „Muster“, „100 % Nullen“ und „100 % eindeutig“.
Gefunden in Spalten	Gesamtzahl der Spalten mit einem übereinstimmenden Profilergebnistyp.
Profil	Name des Profils, das die übereinstimmende Spalte enthält.
Verbindung	Name der relationalen Datenbankverbindung.
Datenquelle	Datenquelle des Profils.
Anzahl der Spalten	Anzahl der Spalten im Profil mit einem übereinstimmenden Profilergebnistyp.

Anzeigen von Datendomänenenerkennungsergebnissen

Sie können auf einen Datendomänennamen klicken, um die Datendomänenenerkennungsergebnisse anzuzeigen. Sie können bestimmte Profile aus den Datendomänenenerkennungsergebnissen öffnen.

1. Führen Sie ein Profil aus, um die Enterprise-Erkennung durchzuführen.
2. Stellen Sie sicher, dass Sie in der Ansicht **Zusammenfassung** sind.
3. Klicken Sie auf eine Datendomäne im Abschnitt **Datendomänenenerkennung** zum Anzeigen der Erkennungsergebnisse.
Eine Liste von Profilen, die die Datendomäne beinhalten, erscheint im rechten Bereich.
4. Wählen Sie bei Bedarf eine Zeile im rechten Bereich.
Die Hyperlinks zu dem Profil werden in blauer Farbe angezeigt.
5. Klicken Sie zum Öffnen des Profils auf die Profilnamenverknüpfung oder die Spaltennamenverknüpfung.
Das Profil öffnet sich und zeigt die Datendomänenenerkennungsergebnisse. Das Analyst Tool markiert die Zeile mit der Datendomäne in den Ergebnissen. Falls erforderlich, können Sie die Profilergebnisse an anderer Stelle effektiv verwenden, wie bei der Erkennungssuche.
6. Klicken Sie auf **Zurück zur Enterprise-Erkennung**, um zur Ansicht **Zusammenfassung** zurückzukehren.

Anzeigen von Spaltenprofilerggebnissen

Sie können die Spaltenprofilerggebnisse der Enterprise-Erkennung in der Ansicht **Zusammenfassung** anzeigen. Sie können bestimmte Profile aus den Datendomänenenerkennungsergebnissen öffnen.

1. Führen Sie ein Profil aus, um die Enterprise-Erkennung durchzuführen.
2. Stellen Sie sicher, dass Sie in der Ansicht **Zusammenfassung** sind.
3. Zur Ansicht der abgeleiteten Muster klicken Sie auf eines der 10 besten Muster im Bereich **Spalten-Profiling**.
Eine Liste von Profilen, die abgeleitete Musterergebnisse enthalten, erscheinen im rechten Bereich.
4. Zur Ansicht von Informationen, beispielsweise aller Nullwerte, aller eindeutigen Werte oder Datentypkonflikte, klicken Sie auf **100% Nullen**, **100% eindeutig** oder **Konflikt zwischen abgeleiteten und dokumentierten Datentypen**.

Die Übereinstimmungsliste von Profilen wird im rechten Bereich angezeigt.

5. Klicken Sie zum Öffnen des Profils auf die Profilnamenverknüpfung oder die Spaltennamenverknüpfung. Das Profil wird geöffnet und zeigt die Spaltenprofilergebnisse an.
6. Klicken Sie auf **Zurück zur Enterprise-Erkennung**, um zur Ansicht **Zusammenfassung** zurückzukehren.

Datentypkonflikt

Die Enterprise-Erkennung identifiziert Datentypkonflikte in Spalten. Ein Datentypkonflikt ist eine Nichtübereinstimmung der abgeleiteten und dokumentierten Datentypen einer Spalte nach dem Ausführen der Enterprise-Erkennung. Ein abgeleiteter Datentyp ist der Datentyp, den das Analyst Tool für eine Datenquellenspalte basierend auf den Spaltendaten ableitet. Ein dokumentierter Datentyp ist der deklarierte Datentyp für eine Spalte in der Quelldatenbank.

Die Enterprise-Erkennung leitet möglicherweise einen anderen Datentyp für eine Spalte ab, und zwar basierend auf den Spaltendaten im Vergleich zum dokumentierten Datentyp einer Spalte. So kann die Enterprise-Erkennung beispielsweise eine Spalte mit einem dokumentierten Zeichenfolgendatentyp als Datumsdatentyp ableiten. Sie können den Datentypkonflikt überprüfen, das passendste Datentypdatum für die Spalte auswählen und es genehmigen.

Anzeigen von Datentypkonflikten

Wenn Sie ein Profil mit Datentypkonflikten in der Ansicht **Zusammenfassung** öffnen, werden die Datentypkonflikte vom Analyst Tool rot markiert.

1. Führen Sie ein Profil aus, um die Enterprise-Erkennung durchzuführen.
2. Stellen Sie sicher, dass Sie in der Ansicht **Zusammenfassung** sind.
3. Klicken Sie im Abschnitt **Spalten-Profiling** auf **Konflikt zwischen abgeleiteten und dokumentierten Datentypen**, um die Datentypkonflikte in den Spaltenprofilergebnissen anzuzeigen.
Im rechten Bereich wird eine Liste von Profilen angezeigt, die Spalten mit den Datentypkonflikten enthält.
4. Wählen Sie bei Bedarf eine Zeile im rechten Bereich.
Die Hyperlinks zu dem Profil werden in blauer Farbe angezeigt.
5. Klicken Sie zum Öffnen des Profils auf die Profilnamenverknüpfung oder die Spaltennamenverknüpfung. Das Profil wird geöffnet und die Datentypkonflikte werden in roter Farbe angezeigt. Sie können die abgeleiteten Datentypen wiederherstellen, um den Datentypkonflikt zu beheben.
6. Zum Wiederherstellen der Datentypen wählen Sie eine Zeile mit den in Konflikt stehenden Datentypen aus und klicken Sie in die Ansicht **Datentypen**.
7. Klicken Sie auf **Aktionen** und wählen Sie dann **Genehmigen** oder **Ablehnen** aus.
8. Klicken Sie auf **Zurück zur Enterprise-Erkennung**, um zur Ansicht **Zusammenfassung** zurückzukehren.

Profilansicht

Die Ansicht **Profile** zeigt eine Liste aller einzelnen Datenobjektprofile, die das Analyst-Tool als Teil der Enterprise-Erkennung ausführt. Die Profilliste zeigt außerdem den Ausführungsstatus jedes Profils an. Sie

können jedes Profil zum Anzeigen der Spaltenprofilerggebnisse und Datendomänenenerkennungsergebnisse öffnen.

Anzeigen von Profileigenschaften

Sie können die Liste der Profile anzeigen, die Teil der Enterprise-Erkennung in der Ansicht **Profile** sind. Sie können bei Bedarf jedes Profil öffnen und die Profilergebnisse wiederherstellen.

1. Führen Sie ein Profil aus, um die Enterprise-Erkennung durchzuführen.
2. Stellen Sie sicher, dass Sie in der Ansicht **Profile** sind.
3. Zum Anzeigen der Profileigenschaften eines Profils klicken Sie auf den Profilnamen.
Die Profileigenschaften erscheinen im rechten Bereich. Die Profileigenschaften beinhalten den Namen der Quelldatenobjekte, den Verbindungsnamen und die Zeilenanzahl.
4. Zum Anzeigen der Profilergebnisse klicken Sie auf **Profil öffnen**.
Das Profil zeigt die Spaltenprofilerggebnisse an.
5. Zum Zurückkehren zur Ansicht **Profile** klicken Sie auf den Ordner oder die Projektnamenverknüpfung in der oberen linken Ecke des Workspace **Discoveries**.

KAPITEL 15

Erkennungssuche in Informatica Analyst

Dieses Kapitel umfasst die folgenden Themen:

- [Erkennungssuche in Informatica Analyst – Übersicht, 133](#)
- [Voraussetzungen der Erkennungssuche, 134](#)
- [Prozess der Erkennungssuche in Informatica Analyst, 134](#)
- [Optionen für die Erkennungssuche, 135](#)
- [Ergebnisse der Erkennungssuche in Informatica Analyst, 136](#)
- [Match-Typen, 138](#)
- [Zugehörige Ressourcen, 140](#)
- [Häufig gestellte Fragen \(FAQ\), 141](#)

Erkennungssuche in Informatica Analyst – Übersicht

Bei der Erkennungssuche werden Objekte gesucht und Beziehungen zu anderen Objekten in den Datenbanken und Schemata des Unternehmens ermitteln. Enterprise-Benutzer können anhand der Erkennungssuche ermitteln, wo sich die Daten und Metadaten im Unternehmen befinden. Sie können nach bestimmten Objekten wie Datenobjekten, Regeln und Profilen suchen.

Bei einer globalen Suche führt das Analyst-Tool eine textbasierte Suche nach Datenobjekten, Datenquellen und Ordnern durch. Wenn Sie eine Erkennungssuche durchführen, enthalten die Suchergebnisse neben den Textübereinstimmungen Objekte mit Beziehungen zu den Objekten, die mit den Suchkriterien übereinstimmen. Die Erkennungssuche umfasst auch Übereinstimmungen basierend auf Profilmetadaten wie Datentypen und Datenmustern. Sie können beispielsweise nach Objekten suchen, die ein bestimmtes Datenmuster enthalten und Namen haben, die ein bestimmtes Schlüsselwort enthalten.

Die Erkennungssuche umfasst folgende Typen von Informationen in den Suchergebnissen:

Objekte im Modellrepository

Enthält Primärobjekte im Zusammenhang mit den Objekten, die mit den Kriterien der Erkennungssuche übereinstimmen. Wenn Sie z. B. nach einem Profil suchen, ist das Datenobjekt des Profils in den Profilergebnissen enthalten.

Profil-Warehouse-Ergebnisse

Enthält Inferenzergebnisse von Profilen, z. B. eine Datendomäne oder ein Datenmuster.

Begriffe im Geschäftsglossar

Berücksichtigt je nach Lizenz Metadaten in der Suche, z. B. einen Geschäftsbegriff in Verbindung mit einer Regel

Erkennungssuche – Beispiel

Sie sind ein Datenverwalter im Unternehmen und dafür verantwortlich, dass die sachgerechte Maskierung vertraulicher Unternehmensdaten sichergestellt ist. Sie möchten möglicherweise personenbezogene Daten (PII, Personally Identifiable Information) in Schemata und Datenbanken identifizieren, auf denen Sie oder der Datenarchitekt eine Enterprise-Erkennung durchführen. Sie haben möglicherweise Datendomänen zum Identifizieren von wichtigen Daten erstellt, die sonst in Datenquellen unentdeckt bleiben. Sie führen die Suche für die Zeichenfolge „SSN“ aus. Das Analyst-Tool zeigt die SSN-Datendomäne und alle übereinstimmenden Spalten aus den Datenquellen an. Darüber hinaus werden mit der Suchzeichenfolge möglicherweise noch weitere Spalten oder Tabellen gefunden, die „SSN“ in ihren Beschreibungen oder Namen haben. Um die Suche einzugrenzen, können Sie Mapping-Spezifikationen filtern, um die Mapping-Spezifikationen anzuzeigen, die auf übereinstimmende Datenobjekte verweisen. Sie können noch weitere Filter anwenden, um zusätzliche Mapping-Spezifikationen auf der Grundlage von Projekten oder Benutzern zu filtern. Möglicherweise möchten Sie anschließend die Mapping-Spezifikationen in den Ergebnissen öffnen, um zu überprüfen, ob die Mapping-Spezifikationen den Datenschutzrichtlinien des Unternehmens entsprechen.

Voraussetzungen der Erkennungssuche

Bevor Sie eine effektive Erkennungssuche in den Datenbanken im Unternehmen ausführen können, sollten Sie eine Enterprise-Erkennung für die Datenbanken und Schemas im Unternehmen durchführen.

Nachdem Sie die Enterprise-Erkennung durchgeführt haben, speichert das Analyst-Tool alle Profilergebnisse im Profiling Warehouse. Stellen Sie sicher, dass alle erforderlichen Datenquellen im Modellrepository vorhanden sind. Überprüfen Sie optional, ob die entsprechenden Objekte im Modellrepository Geschäftsbegriffen zugeordnet sind. Bei der Erkennungssuche ruft der Suchdienst Suchindexinformationen basierend auf den Modellrepository-Objekten und Profiling Warehouse-Ergebnissen ab. Der Suchdienst verwendet dann die indizierten Informationen, um Suchergebnisse auf der Grundlage der entsprechenden Objektmetadaten und -beziehungen anzuzeigen.

Prozess der Erkennungssuche in Informatica Analyst

Sie können Profilergebnisse anhand von Kriterien wie Text, Mustern und Datentypen nach Objekten durchsuchen. Die Suche gibt eine Liste der mit der Suchzeichenfolge verbundenen Objekte zurück.

Schließen Sie die folgenden Schritte ab, um die Erkennungssuche im Analyst-Tool auszuführen:

1. Führen Sie die Enterprise-Erkennung durch und führen Sie die erforderlichen einzelnen Datenobjektprofile auf Datenquellen im Unternehmen aus. Bei der Erkennungssuche sucht das Analyst-Tool nach Informationen in den Profilergebnissen und Modellrepository-Objekten.
2. Wählen Sie aus, nach welchem Typ von Informationen Sie suchen möchten. Sie möchten z. B. alle Objekte finden, die einer Datendomänendefinition für sensible Daten oder einem bestimmten Datenmuster zugeordnet sind.
3. Führen Sie die Suche aus.

4. Analysieren Sie die Suchergebnisse, um Objekte und deren Beziehungen zu anderen Objekten zu ermitteln.
5. Überprüfen Sie erforderlichenfalls die ermittelten Daten auf Konformität mit den Geschäftsanforderungen.

Optionen für die Erkennungssuche

Sie können eine globale Suche oder eine Erkennungssuche durchführen, um nach Ressourcen zu suchen oder Beziehungen zu anderen Ressourcen identifizieren. Bei einer globalen Suche werden Ergebnisse aus dem Modellrepository und optional aus dem Business Glossary abgerufen. Bei einer Erkennungssuche werden Ergebnisse aus dem Modellrepository und dem Business Glossary und sowie Profile abgerufen, die auf Profilergebnissen im Profiling Warehouse basieren.

Sie können nach Ressourcen wie Datenobjekten, Profilen und Mapping-Spezifikationen suchen. Geben Sie eine Suchzeichenfolge für Ressourcen ein, die der Suchzeichenfolge entsprechen und mit der Suchzeichenfolge in Verbindung stehen. Sie können Platzhalterzeichen verwenden, wenn Sie nach Ressourcen suchen.

Verwenden Sie die folgenden Platzhalterzeichen, wenn Sie nach Ressourcen suchen:

* (Sternchen)

Muss am Ende der Suchzeichenfolge eingefügt werden, um alle Namen von Ressourcen zu finden, die mit der Suchzeichenfolgen beginnen Beispiel: Um nach allen Ressourcennamen zu suchen, die mit der Zeichenfolge „emp“ beginnen, müssen Sie „emp*“ in das Suchfeld eingeben.

? (Fragezeichen)

Stellt in der Suchzeichenfolge ein alphanumerisches Zeichen dar

Hinweis: Wenn Sie nach Ressourcen suchen, können Sie die Suchzeichenfolge nicht mit einem Platzhalterzeichen beginnen. Bei der Suche wird die Groß- und Kleinschreibung nicht beachtet.

Wenn Sie nach zwei oder mehr Wörtern in einem Ausdruck suchen möchten, setzen Sie die Wörter in doppelte Anführungszeichen. Verwenden Sie das Zeichen + für den UND-Operator und suchen Sie nach einem Begriff, der in den Suchergebnissen angezeigt werden soll. Beispiel: Wenn die Suchzeichenfolge `+vertraulich +daten` lautet, findet der Suchdienst Metadaten, die beide Begriffe enthalten. Verwenden Sie ein Leerzeichen für die Angabe des ODER-Operators. Beispiel: Wenn die Suchzeichenfolge `vertrauliche Daten` lautet, findet der Suchdienst Metadaten, die nur einen der beiden Suchbegriffe enthalten.

Wenn die Suchzeichenfolgen einen Bindestrich (-), Unterstrich (_) oder Binnenmajuskel enthält, findet der Suchdienst das ganze Wort und durch Delimiter getrennte Wortteile. Wenn Sie beispielsweise nach `Profile_Customer` suchen, findet die Suchmaschine in den Repositories `Profil`, `Kunde` und `Profile_Customer`. Um Sonderzeichen wie z. B. * und ? in Ihrer Suchzeichenfolge zu verwenden, setzen Sie die Suchzeichenfolge mit den Sonderzeichen in doppelte Anführungszeichen.

Sie können eine Erkennungssuche durchführen, die aus einer Stichwortsuche und einem Erkennungsschlüssel besteht. Beispiel: Sie möchten nach Mitarbeiter-ID-Spalten suchen, die das Format `<VornameInitialen> <NachnameInitialen>-<SSN>` verwenden, um das Datensicherheitsrisiko zu bestimmen. Für die Suche nach den Mitarbeiter-ID-Spalten geben Sie `Mitarbeiter-ID` im Abschnitt **Suche** des Arbeitsbereichs „Bibliothek“ ein und setzen den Musterfilter auf `XX-999999999 <= 100%`.

Kriterien der Erkennungssuche

Verwenden Sie die Kriterien der Erkennungssuche, um anhand von Kriterien wie Mustern, Datentypen, eindeutigen Werten und Nullwerten nach Informationen zu suchen. Sie können in der Suche die Bedingungsoperatoren =, > = oder < = verwenden.

In der folgenden Tabelle werden die Kriterien beschrieben, die Sie bei der Erkennungssuche verwenden können:

Option	Beschreibung
Suchen	Textausdruck, den Sie suchen möchten.
Löschen	Löscht die Suchzeichenfolge und alle anderen Suchkriterien, die Sie zuvor ausgewählt haben.
Muster von	Spaltenmuster und Prozentsatz, den Sie in die Suche aufnehmen möchten. Hinweis: Die Option akzeptiert keine Steuerzeichen in einem Muster.
Datentyp von	Datentyp der Spalte und Prozentsatz, den Sie in die Suche aufnehmen möchten.
Eindeutige Werte	Prozentsatz der eindeutigen Werte in Spalten, die Sie in die Suche aufnehmen möchten.
Nullen	Prozentsatz der Nullwerte in Spalten, die Sie in die Suche aufnehmen möchten.

Suchen nach einem Objekt

Sie können im Arbeitsbereich **Bibliothek** nach einem Objekt suchen. Die Suchergebnisse enthalten Objekte, die sowohl im Developer-Tool als auch im Analyst-Tool erstellt wurden.

1. Öffnen Sie den Arbeitsbereich **Bibliothek**.
2. Stellen Sie sicher, dass Sie sich im Abschnitt **Erkennungssuche** befinden.
3. Geben Sie im **Suchfeld** die Zeichenfolge ein, nach der Sie suchen möchten.
4. Konfigurieren Sie die Suchfilter, um die Suche einzugrenzen.
Die Filter umfassen Muster, Datentypen, eindeutige Werte und Nullwerte.
5. Klicken Sie auf das Symbol **Suchen**.

Ergebnisse der Erkennungssuche in Informatica Analyst

Bei der Erkennungssuche werden Ressourcen in allen für Erkennungssuchvorgänge lizenzierten Repositorys ermittelt, z. B. Modellrepository und Profiling Warehouse.

Die Ergebnisse der Enterprise-Erkennungssuche enthalten die Gesamtanzahl der Übereinstimmungen sowie die Liste der Übereinstimmungen. Sie können jede Übereinstimmung erweitern, um die Übereinstimmungseigenschaften, direkte und indirekte Übereinstimmungsinformationen sowie die Gesamtanzahl der verwandten Ressourcen, soweit vorhanden, anzuzeigen. Eine direkte Übereinstimmung ist eine Übereinstimmung mit einigen oder allen Metadaten der Ressource, die der Suchanfrage entsprechen. Eine indirekte Übereinstimmung ist eine Ressourcenübereinstimmung, die mit der Ressource verbunden ist, die der Suchanfrage direkt entspricht.

In welcher Reihenfolge die Suchergebnisse angezeigt werden, hängt von den folgenden Faktoren ab:

- Objekteigenschaft, die den Suchkriterien entspricht Der Name des Objekts besitzt eine höhere Priorität als seine Beschreibung. Die Objektbeschreibung besitzt eine höhere Priorität als andere Objekteigenschaften.
- Objekttyp. Datendomänen und Datendomänengruppen haben eine niedrigere Priorität als andere Objekte.
- Wiederherstellung. Wiederhergestellte Profilergebnisse haben eine höhere Priorität als Profilergebnisse, die Sie nicht wiederhergestellt haben.
- Häufigkeit, mit der die Suchkriterien den Objekten, einschließlich direkter und indirekter Übereinstimmungen, entsprechen.
- Relative Häufigkeit des Stichworts

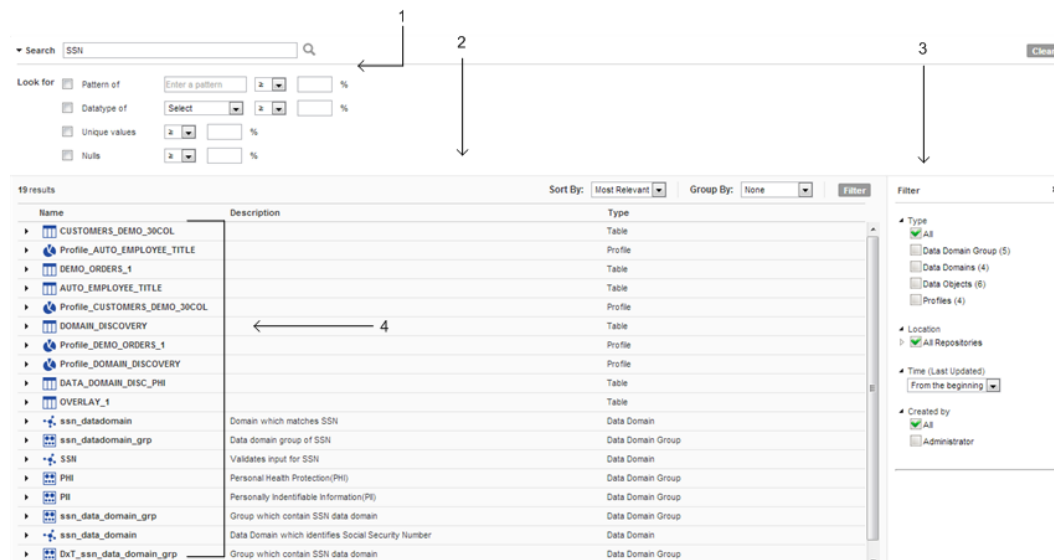
Zum Anzeigen der Suchergebnisse benötigen Sie die entsprechenden Berechtigungen für Projekte, die die direkten und indirekten Übereinstimmungen enthalten.

Bereich für die Erkennungssuchergebnisse

Zu den Ergebnissen der Erkennungssuche gehören der Name des Objekts, der Objekttyp und die Objektbeschreibung. Verwenden Sie Filter, um die Suchergebnisse einzuzugrenzen.

Die Suchergebnisse werden im Ergebnissgitter des Arbeitsbereichs **Bibliothek** angezeigt. Sie können die Ergebnisse nach Relevanz sortieren. Das Gruppieren der Ergebnisse anhand von Objekttyp, Repository-Speicherort, Zeit und Benutzer, von dem die Objekte erstellt wurden, ist ebenfalls möglich.

Die folgende Abbildung zeigt die Schnittstelle für die Ergebnisse der Erkennungssuche:



1. Suchkriterien
2. Ergebnissgitter
3. Filter
4. Suchergebnisse

Im Bereich für die Erkennungssuchergebnisse werden die folgenden Abschnitte standardmäßig angezeigt:

Suchkriterien

Zeigt die Suchfelder einschließlich der auf Profilmetadata basierenden Suchfilter an, die Sie festlegen können, um die Suche zu begrenzen. Die Suchfelder werden im oberen Bereich des Analyst-Tools angezeigt.

Ergebnisgitter

Zeigt die Gesamtanzahl der Matches und die Liste der Matches anhand der Kriterien an, die in den Suchkriterienfeldern ausgewählt werden. Das Ergebnisgitter enthält auch Objektbeschreibungen, Objekttypen sowie Dropdownlisten zum Sortieren und Gruppieren der Suchergebnisse.

Filter

Zeigt die Filter an, die zum Filtern der Suchergebnisse festgelegt werden können. Der Abschnitt **Filter** wird im rechten Bereich des Analyst-Tools angezeigt.

Suchergebnisse

Zeigt die übereinstimmenden Suchergebnisse anhand der Suchzeichenfolge an, einschließlich der Objekte, die den Suchkriterien entsprechen. Der Suchergebnisabschnitt enthält die Match-Eigenschaften, die Abschnitte **Direkte Matches** und **Indirekte Matches**, die beim Erweitern eines Matches angezeigt werden. Sie können auch die Gesamtanzahl der mit dem Match im Ergebnisgitter verbundenen Objekte anzeigen.

Filtern von Ergebnissen der Erkennungssuche

Sie können die Suchergebnisse anhand von Objekttyp, Repository-Speicherort, Zeit und Benutzer, von dem die Objekte erstellt wurden, filtern. Wenn Sie Business Glossary installiert haben, können Sie auch objektspezifische Filter für Geschäftsbegriffe, Kategorien und Richtlinien in Business Glossary verwenden.

1. Führen Sie die globale Suche oder Erkennungssuche im Abschnitt **Erkennungssuche** im Arbeitsbereich **Bibliothek** aus.
2. Klicken Sie im Ergebnisgitter auf **Filter**, um den Abschnitt **Filter** zu öffnen.
3. Wählen Sie im Abschnitt **Filter** die erforderlichen Filter und die zugehörigen Einstellungen aus.
4. Die überarbeiteten Suchergebnisse werden im Ergebnisgitter basierend auf den ausgewählten Filtereinstellungen angezeigt.
5. Klicken Sie zum Löschen aller Filtereinstellungen auf **Alle löschen** im oberen Bereich des Ergebnisgitters.

Match-Typen

Zu den Ergebnissen der Erkennungssuche gehören direkte und indirekte Matches. Ein direkter Match ist eine Übereinstimmung mit einigen oder allen Metadaten des Objekts, das mit der Suchanfrage übereinstimmt. Ein indirekter Match ist ein Objekt-Match, der mit einem anderen Objekt verknüpft ist, das mit der Suchanfrage direkt übereinstimmt.

Wenn eine Suchabfrage mehrere Suchkriterien enthält, entsprechen die Suchergebnisse den Suchkriterien direkt oder indirekt bzw. sowohl direkt als auch indirekt. Sie können einige direkte und indirekte Matches aus den Suchergebnissen im Schreibschutz- oder Bearbeitungsmodus öffnen.

Direkter Match

Ein direkter Match ist eine Übereinstimmung mit einigen oder allen Metadaten des Objekts, das mit der Suchanfrage übereinstimmt. Wenn Sie beispielsweise nach allen Objekten mit dem Namen „Kunde“ suchen, listet das Analyst-Tool Datenobjekte und Profile mit dem Namen „Kunde“ als direkte Matches auf. Nachdem

Sie die Erkennungssuche ausgeführt haben, enthält die angezeigte Match-Liste Verknüpfungen zu einigen der Objekte.

Sie können ein Objekt in den Suchergebnissen erweitern, um weitere Informationen zum direkten Match, wie z. B. Objekteigenschaften, anzuzeigen.

Indirekte Übereinstimmung

Eine indirekte Übereinstimmung ist eine Übereinstimmung, die mit der direkten Übereinstimmung verbunden ist. Beispiel: Eine Scorecard verwendet eine Regel mit dem Suchschlüsselwort. Die Erkennungssuche gibt die Regel als eine direkte Übereinstimmung zurück und die Scorecard als eine indirekte Übereinstimmung. Die Scorecard ist eine indirekte Übereinstimmung, da sie auf die Regel verweist.

Verwenden Sie die Informationen von indirekten Übereinstimmungen, um ausgeblendete Beziehungen zwischen Objekten zu identifizieren und Objektbeziehungen besser zu verstehen. Die Ergebnisse indirekter Übereinstimmungen machen außerdem deutlich, warum die Erkennungssuche ein Objekt zurückgibt.

Anzeigen der Match-Informationen

Nachdem Sie die Erkennungssuche ausgeführt haben, können Sie die Match-Informationen einschließlich direkter und indirekter Matches anzeigen. Sie können auch die Objekteigenschaften anzeigen, beispielsweise Objekttyp, Beschreibung und verbundene Objekte. Sie können einige Objekte aus den Suchergebnissen heraus öffnen und sie bei Bedarf ändern.

1. Führen Sie die globale Suche oder Erkennungssuche im **Suchbereich** im Arbeitsbereich **Bibliothek** aus.
2. Klicken Sie im Ergebnissgitter am Anfang eines Objektnamens auf das Symbol „Erweitern“.
Die Objekteigenschaften und Match-Informationen werden in einem Abschnitt unter dem Objektnamen angezeigt.
3. Überprüfen Sie die Informationen zu direkten und indirekten Matches.
Sie können Objektbeziehungen und andere Informationen anzeigen, beispielsweise die Gesamtanzahl der verbundenen Objekte. Die Objektbeziehungen beinhalten sowohl direkte Matches als auch indirekte Matches.
4. Wenn die Objektinformationen Hyperlinks enthalten, klicken Sie auf die Hyperlinks, um die Objekte in einem anderen Arbeitsbereich zu öffnen.
5. Klicken Sie erneut auf das Symbol „Erweitern“, um den Abschnitt mit den Match-Informationen zu schließen.

Öffnen von Objekten über die Ergebnisse der Erkennungssuche

Sie müssen über die erforderlichen Projekt-, Objekt- und Lizenzberechtigungen verfügen, um Objekte aus den Ergebnissen der Erkennungssuche heraus anzuzeigen.

1. Führen Sie die globale Suche oder Erkennungssuche im **Suchbereich** im Arbeitsbereich **Bibliothek** aus.
2. Klicken Sie mit der rechten Maustaste im Ergebnissgitter auf einen Objektnamen.
Ein Menü für Schnellzugriff wird geöffnet.
3. Wählen Sie **Öffnen** aus, um das Objekt im schreibgeschützten Modus in seinem Arbeitsbereich anzuzeigen.
4. Wählen Sie **Bearbeiten** aus, um das Objekt in seinem Arbeitsbereich zu ändern.
5. Wählen Sie **Löschen** aus, um das Objekt aus den Suchergebnissen zu löschen.

Wenn Sie ein Objekt aus den Suchergebnissen löschen, entfernt das Analyst-Tool das Objekt aus dem Modellrepository.

6. Klicken Sie auf den Arbeitsbereich **Bibliothek**, um zurück zur **Bibliothek** zu navigieren.

Zugehörige Ressourcen

Sie können die zugehörigen Ressourcen einer Ressource aus den Suchergebnissen anzeigen. Eine zugehörige Ressource ist eine Ressource im Modellrepository oder Geschäftsglossar, die mit einer ausgewählten Ressource in den Suchergebnissen verbunden ist. Die zugehörige Ressource teilt einige Metadaten mit der Ressource in den Suchergebnissen. Eine Datenquelle kann über Profile, abgeleitete Datendomänen und Mappings als zugehörige Ressourcen verfügen.

Beispiel: Ein Profil kann Teil der Suchergebnisse sein. Sie können die zugehörigen Ressourcen des Profils, wie Regeln und Datenquellen für das Profil, anzeigen. Sie können die zugehörigen Ressourcen im Arbeitsbereich **Verbundene Ressourcen** anzeigen. Die zugehörigen Ressourcen, die angezeigt werden, unterscheiden sich aufgrund ihres Ressourcentyps. Wenn Sie beispielsweise die zugehörigen Ressourcen einer Regel anzeigen, sehen Sie auch Ressourcen wie den zugeordneten Geschäftsbegriff, die Mapping-Spezifikation und das Profil.

Verbundene Objekte für jeden Objekttyp

Die für ein Objekt im Analyst-Tool angezeigten verbundenen Objekte hängen vom Typ des Objekts ab, nach dem Sie suchen.

In der folgenden Tabelle werden die verbundenen Objekte für jeden Objekttyp beschrieben:

Objekttyp	Verbundene Objekte
Geschäftsbegriff	Datendomäne, Datenobjekt, Mapplet und Regel.
Datendomäne	Geschäftsbegriff, Datendomänengruppe, Datenobjekt und Profil.
Datendomänengruppe	Datendomäne, Datenobjekt und Profil.
Datenobjekt	Geschäftsbegriff, Datendomäne, Datendomänengruppe, Mapping-Spezifikation, Profil, Scorecard, Mapping und Mapplet.
Enterprise-Erkennungsprofil	Datenobjekt und Profil.
Mapping Hinweis: Dieses Objekt wird im Developer-Tool geöffnet.	Datenobjekt, Mapping-Spezifikation, Mapplet und Regel.
Mapping-Spezifikation	Datenobjekt, Profil, Scorecard, Mapping, Mapplet und Regel.
Mapplet Hinweis: Dieses Objekt wird im Developer-Tool geöffnet.	Geschäftsbegriff, Datenobjekt, Mapping-Spezifikation, Mapping, Mapplet und Regel.

Objekttyp	Verbundene Objekte
Profil	Datendomäne, Datendomänengruppe, Datenobjekt, Regel und Mapping-Spezifikation. Hinweis: Eine Scorecard ist nicht im verbundenen Objekt für ein Profil enthalten.
Regel	Datenobjekt, Regel, Geschäftsbegriff, Mapping-Spezifikation, Profil, Scorecard und Mapping.

Anzeigen verbundener Objekte

Sie können die Gesamtanzahl der verbundenen Objekte einsehen, wenn Sie die Informationen zum Objekt-Match in den Suchergebnissen anzeigen.

1. Führen Sie die Erkennungssuche im Arbeitsbereich **Bibliothek** aus.
2. Klicken Sie im Ergebnissgitter auf das Symbol zum Erweitern und klicken Sie auf die Verknüpfung für die Anzahl verbundener Objekte. Klicken Sie alternativ mit der rechten Maustaste auf den Objektnamen und wählen Sie **Verbundene Objekte zeigen** aus.

Eine Liste aller verbundenen Objekte wird im Arbeitsbereich **Verbundene Objekte** angezeigt.

3. Klicken Sie zum Anzeigen der Details eines Objekts auf den Objektnamen oder klicken Sie mit der rechten Maustaste und wählen Sie **Öffnen** aus.
4. Um die verbundenen Objekte eines verbundenen Objekts anzuzeigen, klicken Sie mit der rechten Maustaste auf den Objektnamen und wählen Sie **Verbundene Objekte zeigen** aus.

Die Informationen zum verbundenen Objekt werden im Arbeitsbereich angezeigt.

5. Um zwischen mehreren Arbeitsbereichen für verbundene Objekte zu navigieren, wählen Sie im Arbeitsbereich **Verbundene Objekte** ein kürzlich geöffnetes Objekt aus.

Häufig gestellte Fragen (FAQ)

Warum kann ich einige der erwarteten Suchergebnisse nicht anzeigen?

Die Suchergebnisse können aufgrund verschiedener Ursachen nicht angezeigt werden. Stellen Sie sicher, dass die Suchkriterien die folgenden Richtlinien erfüllen:

- Die in den Suchergebnissen angezeigten Ressourcen sind von Projektberechtigungen abhängig.
- Die Ergebnisse der Erkennungssuche enthalten keine Werthäufigkeiten aus den Spaltenprofilergebnissen.
- Die Suchergebnisse enthalten keine Profilergebnisse, die Sie ablehnen, wenn Sie die Profilergebnisse wiederherstellen.
- Die angezeigten Suchergebnisse sind abhängig vom Extraktionsintervall des Suchindexes und der Verfügbarkeit der Ressourcen im Suchindex.

Kann ich die Ergebnisse der Erkennungssuche für zukünftige Verwendungen speichern bzw. die Ergebnisse für einen anderen Benutzer freigeben?

Nein. Sie können Ergebnisse der Erkennungssuche nicht speichern oder freigeben.

Warum werden einige der Ergebnisse der Erkennungssuche am Anfang oder Ende der Suchergebnisse angezeigt?

Die Reihenfolge, in das Analyst-Tool die Suchergebnisse anzeigt, hängt von mehreren Faktoren ab. Dazu zählen der Objekttyp, die wiederhergestellten Profilergebnisse, die Objekteigenschaft, die in erster Linie mit den Suchkriterien übereinstimmt, sowie der interne Suchrang jedes Objekts.

Kann ich die Ergebnisse der Erkennungssuche exportieren?

Nein. Sie können die Suchergebnisse nicht exportieren.

KAPITEL 16

Business Glossary Desktop in Informatica Analyst

Dieses Kapitel umfasst die folgenden Themen:

- [Geschäftsbegriffe, 143](#)
- [Geschäftsbegriffe verwalten im Unternehmensglossar des Metadata Manager, 144](#)
- [Suchen eines Geschäftsbegriffs im Unternehmensglossar-Desktop, 144](#)

Geschäftsbegriffe

Sie können Geschäftsbegriffe im Unternehmensglossar-Desktop suchen. Sie können Geschäftsbegriffe anzeigen und Geschäftsbegriffsaufgaben ausführen, die auf der Lizenz für Metadata Manager basieren.

Ein Unternehmensglossar ist ein Satz von Bedingungen, die Geschäftssprache benutzen, um Konzepte für Geschäftsbenutzer zu definieren. Ein Geschäftsbegriff enthält die Geschäftsdefinition und Verwendung eines Konzepts.

Der Unternehmensglossar-Desktop ist ein Client, der mit dem Metadata Manager-Service verbunden ist, welcher das Geschäftsglossar beinhaltet. Sie müssen den Unternehmensglossar-Desktop öffnen, bevor Sie einen Analyst-Tool-Objektnamen suchen können. Sie können die Bedeutung eines Analyst-Tool-Objektnamen als Geschäftsbegriff im Unternehmensglossar-Desktop nachsehen, um dessen Geschäftserfordernisse und aktuellen Gebrauch zu verstehen.

Der Metadata Manager beinhaltet Geschäftsglossare. Sie müssen einen Metadata Manager-Service mit dem Analyst Service assoziieren, um nach einem Unternehmensglossar des Metadaten-Managers aus dem Analyst-Tool zu suchen. Sie können die Geschäftsbegriffe in einem Unternehmensglossar oder in Kategorien zusammengeführt anzeigen zu lassen. Sie können die Geschäftsbegriffe von Metadata Manager bearbeiten.

Sie können im Metadata Manager-Repository anhand des Metadata Manager-Geschäftsbegriffs nach Metadata Manager-Objekten suchen. Sie können Metadata Manager-Objekte aus den Suchergebnissen auswählen und diese als Datenobjekte im Analysetool importieren. Sie können keinen Metadata Manager-Geschäftsbegriff zum Unternehmensglossar von Metadata Manager hinzufügen.

Geschäftsbegriffe verwalten im Unternehmensglossar des Metadata Manager

Sie können das Unternehmensglossar von Metadata Manager im Analysetool öffnen, um die Geschäftsbegriffe von Metadata Manager zu verwalten.

1. Klicken Sie in der Kopfzeile des Analysetools auf **Verwalten > Begriffe verwalten**.
Metadata Manager und das Unternehmensglossar von Metadata Manager öffnen sich in einer anderen Registerkarte. Die Geschäftsbegriffe von Metadata Manager erscheinen in der Ansicht **Glossar** von Metadata Manager.
2. Um ein Unternehmensglossar auszuwählen, wählen Sie ein Glossar aus der Anzeigen-Liste.
3. Um Geschäftsbegriffe anhand von Kategorien anzusehen, klicken Sie auf **Aktionen > Ansicht > Kategorien**.
4. Um alle Geschäftsbegriffe in einem Unternehmensglossar in alphabetischer Reihenfolge anzusehen, klicken Sie auf **Aktionen > Ansicht > Alphabet**.
5. Um alle Geschäftsbegriffe anzusehen, die mit einem speziellen Buchstaben beginnen, klicken Sie auf den Buchstaben.
6. Um einen Geschäftsbegriff zu bearbeiten, wählen Sie den Geschäftsbegriff und klicken Sie auf **Aktionen > Eigenschaften bearbeiten**.

Suchen eines Geschäftsbegriffs im Unternehmensglossar-Desktop

Suchen Sie einen Analyst-Tool-Objektnamen im Unternehmensglossar-Desktop als Geschäftsbegriff, um seine Geschäftserfordernisse und den aktuellen Gebrauch zu verstehen.

Der Unternehmensglossar-Desktop muss auf Ihrem Rechner installiert sein.

1. Markieren Sie den Namen eines Objekts.
2. Verwenden Sie die Abkürzungstasten-Kombination, um den Namen des Objekts als Geschäftsbegriff im Unternehmensglossar-Desktop zu suchen.

Die standardmäßige Abkürzungstasten-Kombination ist `STRING+ALT+Q`.

Teil III: Data Discovery mit Informatica Developer

Dieser Teil enthält die folgenden Kapitel:

- [Informatica Developer-Profil, 146](#)
- [Datenobjektprofil, 149](#)
- [Spaltenprofile in halbstrukturierten Datenquellen, 164](#)
- [Regeln in Informatica Developer, 171](#)
- [Mapplet- und Mapping-Profil, 173](#)
- [Spaltenprofilergebnisse in Informatica Developer, 176](#)
- [Scorecards in Informatica Developer, 181](#)
- [Datendomänenerkennung in Informatica Developer, 184](#)
- [Enterprise-Erkennung in Informatica Developer, 197](#)
- [Ergebnisse der Enterprise-Erkennung, 215](#)
- [Business Glossary Desktop in Informatica Developer, 226](#)

KAPITEL 17

Informatica Developer-Profile

Dieses Kapitel umfasst die folgenden Themen:

- [Informatica Developer-Profile - Übersicht, 146](#)
- [Informatica Developer-Profilansichten, 147](#)
- [Repository-Objektsperren und teambasierte Entwicklung mit versionierten Objekten, 148](#)

Informatica Developer-Profile - Übersicht

In Informatica Developer können Sie Profile erstellen und ausführen, um Datenqualitätsprobleme in einem Dataset zu erkennen und die Spaltenbeziehungen in einem Dataset zu verstehen.

Für die folgenden Typen von Datenanalysen können Sie Profile erstellen:

- Spalten-Profiling
- Spaltenprofile in halbstrukturierten Datenquellen
- Primärschlüsselerkennung
- Funktionale Abhängigkeitserkennung
- Fremdschlüsselerkennung
- Join-Analyse
- Überschneidungserkennung
- Datendomänenenerkennung
- Enterprise-Erkennung

Zum Erstellen der Profile im Developer Tool wird der Assistent verwendet. Der Profilerstellungsassistent bietet Ihnen die Optionen **Profil**, **Mehrere Profile** und **Enterprise-Erkennungsprofil** zum Erstellen von Profilen an.

Profil

Erstellen eines Profils für ein einzelnes Datenobjekt. Sie definieren für ein einzelnes Profil Filter, Regeln und Drilldown-Optionen zum Spalten-Profiling. Sie können auch erweiterte Optionen wählen, um ein Spaltenprofil, ein Primärschlüsselprofil, ein funktionales Abhängigkeitsprofil für die Datendomänenenerkennung zu erstellen. In den Ergebnissen werden das Spaltenprofil, die Primärschlüsselinferenz, die funktionale Abhängigkeit und die Datendomäneninferenz angezeigt. Sie können ein Spaltenprofil für ein Einfachdatei-Datenobjekt, ein relationales Datenobjekt und halbstrukturierte Datenobjekte erstellen.

Mehrfachprofile

Erstellen eines Satzes von Profilen für mehrere Objekte. Das Developer Tool erstellt ein Profil für jedes Objekt und führt die Profile gleichzeitig aus. Wenn Sie gleichzeitig mehrere Profile erstellen, können Sie Daten nicht objektübergreifend analysieren.

Enterprise-Erkennungsprofil

Aufbauen eines Datenmodells aus mehreren Datenobjekten und Erstellen eines Profils, das Daten objektübergreifend analysiert. Erstellen Sie ein Enterprise-Erkennungsprofil und fügen ihm physikalische Datenobjekte hinzu, die Sie zusammen profilieren möchten. Sie können ein Datenobjektprofil, ein Fremdschlüsselprofil und ein Join-Profil erstellen. Sie können für jedes Datenobjekt im Enterprise-Erkennungsprofil allgemeine Eigenschaften, Spalten zum Profilieren, Schlüssel und Beziehungen konfigurieren. Sie können sich überschneidende Daten in einer Datenquelle oder in mehreren Datenquellen erkennen.

Sie können auch eine Enterprise-Erkennung ausführen, die Datenerkennungsaufgaben wie Spaltenprofil, Datendomänenerkennung, Primärschlüsselprofil und Fremdschlüsselprofil erstellt und ausführt. Die Enterprise-Erkennung wird auf vielen Datenquellen über mehrere Verbindungen hinweg ausgeführt.

Welche Operationen Sie mit dem jeweiligen Profiltyp durchführen können, geht aus folgender Tabelle hervor:

Profiloption	Profilvorgänge
Profil	<ul style="list-style-type: none">- Ausführen eines Spaltenprofils.- Suchen nach Primärschlüsseln.- Suchen nach funktionalen Abhängigkeiten.- Identifizieren von Datendomänen
Mehrfachprofile	Erstellen und gleichzeitiges Ausführen von Spaltenprofilen für mehrere Objekte.
Enterprise-Erkennungsprofil	<ul style="list-style-type: none">- Ausführen eines Spaltenprofils für einen einzelnen Datensatz- Suchen nach Primärschlüsseln.- Suchen nach Fremdschlüsseln.- Suchen nach funktionalen Abhängigkeiten.- Durchführen von Join-Analysen.- Erkennen von Überschneidungen zwischen Spalten.- Ausführen der Enterprise-Erkennung.

Informatica Developer-Profilansichten

Sie können im Informatica Developer mit den Ansichten **Übersicht**, **Definition**, **Kommentare** und **Ergebnisse** Informationen zu einem Profil anzeigen und hinzufügen.

Wenn Sie ein Profil aus dem **Objekt-Explorer** öffnen, zeigt der Editor im rechten Bereich unter den folgenden Ansichten die Profilinformationen an:

Übersicht

Ansicht und allgemeine Informationen zu dem Profil, wie Name, Beschreibung und Ort.

Definition

Ansicht und Einrichtung der Profildefinition.

Zu diesen Informationen gehören Listen von Filtern und Regeln, die Sie dem Profil zuordnen können, Drilldown-Optionen und Profilfunktionen, die während der Profilausführung aktiviert werden.

Ergebnisse

Anzeigen der Ergebnisse der Profilausführung. Sie können die Ergebnisse exportieren, nachdem Sie ein Profil ausgeführt haben.

Kommentare

Ansicht und Hinzufügen von Kommentaren zu dem Profil.

Repository-Objektsperren und teambasierte Entwicklung mit versionierten Objekten

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

Das Modellrepository behält Objektsperren bei, wenn das Developer Tool unerwartet beendet wird. Wenn Sie erneut eine Verbindung zum Modellrepository herstellen, können Sie die von Ihnen gesperrten Objekte anzeigen. Sie können mit der Bearbeitung der Objekte fortfahren oder die Objekte entsperren. Sie können gesperrte Objekte über das Dialogfeld **Gesperrte Objekte** anzeigen und entsperren. Klicken Sie zum Anzeigen des Dialogfelds **Gesperrte Objekte** auf **Anzeigen > Gesperrte Objekte**.

Falls das Modellrepository in ein Versionsverwaltungssystem integriert ist, können Sie die Objektversionen im Developer Tool mit der Verwaltung von versionierten Objekten verwalten. Sie können unter anderem die folgenden Aktionen durchführen: Einchecken und Auschecken von Objekten, Anzeigen und Abrufen historischer Objektversionen oder Rückgängigmachen des Auscheckens.

Das Modellrepository verhindert, dass Objekte von anderen Mitgliedern des Entwicklungsteams überschrieben werden. Wenn Sie ein Objekt öffnen, das von einem anderen Benutzer ausgecheckt wurde, werden Sie in einer Benachrichtigung über den Benutzer informiert, der das Objekt ausgecheckt hat. Sie können ein ausgechecktes Objekt im schreibgeschützten Modus öffnen oder es unter einem anderen Namen speichern.

Wenn Objekte geändert werden, erstellt das Modellrepository neue Versionen der entsprechenden Objekte.

Nachdem Sie einen Datentyp, Primärschlüssel, Fremdschlüssel oder eine Datendomäne wiederhergestellt haben, erhöht das Modellrepository die Versionsnummer.

Bei der Wiederherstellung einer Version werden in der Ansicht „Ergebnisse“ die neuesten Profilergebnisse und nicht die Profilergebnisse der wiederhergestellten Version angezeigt. Der Grund dafür ist, dass im Versionsverwaltungssystem Profildefinitionen für alle Versionen im Modellrepository gespeichert sind und die Profilergebnisse aus dem Profiling-Warehouse extrahiert werden. Weitere Informationen zu Repository-Objektsperren und zur Verwaltung von versionierten Objekten finden Sie im *Developer Tool-Handbuch*.

KAPITEL 18

Datenobjektprofile

Dieses Kapitel umfasst die folgenden Themen:

- [Datenobjektprofile - Übersicht, 149](#)
- [Spaltenprofile in Informatica Developer, 150](#)
- [Laufzeitumgebung, 153](#)
- [Primärschlüssel-Erkennung, 154](#)
- [Funktionale Abhängigkeitserkennung, 156](#)
- [Betriebssystemprofile in Informatica Developer, 158](#)
- [Erstellen eines einzelnen Datenobjektprofils in Informatica Developer, 158](#)
- [Erstellen mehrerer Datenobjektprofile in Informatica Developer, 159](#)
- [Bearbeiten eines Profils, 160](#)
- [Synchronisierungsoption, 160](#)
- [Kommentare, 163](#)

Datenobjektprofile - Übersicht

Ein Datenobjektprofil erkennt Informationen über die Spaltendaten und Metadaten in einer Datenquelle. Sie können in Informatica Developer ein Profil für ein einzelnes Datenobjekt und mehrere Datenobjekte ausführen. Ein einzelnes Datenobjektprofil analysiert eine Datenquelle. Mehrere Datenobjektprofile analysieren mehr als eine Datenquelle. Wenn Sie mehrere Datenobjektprofile erstellen, kann für sie ein Spaltenprofil ausgeführt werden.

In der folgenden Tabelle werden die Datenerkennungsaufgaben beschrieben, die in einem einzelnen Datenobjektprofil durchgeführt werden können:

Aufgabe	Beschreibung
Spalten-Profiling	Erkennt die Eigenschaften von Daten, wie Häufigkeiten, Prozentsätze und Muster. Um zu bestimmen, welche Zeilen das Profil während der Laufzeit liest, können Sie Filter hinzufügen. Zeilen, die die Filterkriterien nicht erfüllen, werden vom Profil nicht bearbeitet.
Primärschlüsselerkennung	Erkennt Spalten mit Werten, die die Zeilen in einer Datenquelle eindeutig identifizieren können.

Aufgabe	Beschreibung
Funktionale Abhängigkeitserkennung	Erkennt Abhängigkeiten zwischen Spaltenpaaren in einer Datenquelle.
Datendomänenerkennung	Ermittelt alle Datendomänen für eine Spalte auf der Basis des Spaltenwerts oder Spaltennamens.

In der folgenden Tabelle werden die Datenerkennungsaufgaben beschrieben, die an mehreren Datenobjekten durchgeführt werden können, wenn Sie mit der Option **Enterprise-Erkennungsprofil** ein Datenmodell erstellen:

Aufgabe	Beschreibung
Fremdschlüsselerkennung	Erkennt Spalten mit Werten, die den Primärschlüsselwerten in einer anderen Datenquelle entsprechen.
Join-Analyse	Erkennt das Ausmaß potenzieller Joins zwischen den Daten in zwei Spalten einer Datenquelle oder zwischen zwei Datenquellen.
Überschneidungserkennung	Erkennt die prozentuale Überschneidung von Daten zwischen Spaltenpaaren einer oder mehreren Datenquellen.
Enterprise-Erkennung	Erkennt Spaltenprofilstatistiken, Datendomänen, Primärschlüssel und Fremdschlüssel in einer großen Anzahl von Datenquellen über mehrere Verbindungen oder Schemas hinweg.

Spaltenprofile in Informatica Developer

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

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

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

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

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

Filteroptionen

Sie können erweiterte Filter oder SQL-Filter zum Festlegen der Zeilen hinzufügen, die von einem Spaltenprofil bei der Ausführung des Profils verwendet werden. Zeilen, die die Filterkriterien nicht erfüllen, werden vom Profil nicht bearbeitet.

Erstellen eines erweiterten Filters

Sie können einen erweiterten Filter mit Ausdrücken, wie z. B. AND, OR und NOT, erstellen, um eine Teilmenge der ursprünglichen Datenquelle zu erhalten.

1. Erstellen oder öffnen Sie ein einzelnes Datenobjektprofil.
2. Wählen Sie die Ansicht **Filter** aus.
3. Klicken Sie auf **Hinzufügen**.
Das Dialogfeld **Assistenten auswählen** wird angezeigt.
4. Klicken Sie im Dialogfeld **Assistenten auswählen** auf **Erweiterter Filter**.
Das Dialogfeld **Filter** wird angezeigt.
5. Geben Sie einen Namen und eine optionale Beschreibung für den erweiterten Filter ein.
6. Wählen Sie **Als aktiv festlegen** aus, um den Filter auf das Profil anzuwenden. Klicken Sie auf **Weiter**.
7. Wählen Sie **Filterdefinition** aus, um einen Filter zu definieren.
8. Sie können einen erweiterten Filter im Bereich **Funktionen** oder im Bereich **Spalten** erstellen.
 - Wählen Sie im Bereich **Funktionen** eine Funktionskategorie aus und klicken Sie auf die Schaltfläche mit den nach rechts weisenden Pfeilen (>>).
Geben Sie im Dialogfeld die Parameter an und klicken Sie auf **OK**. Die Funktion wird zusammen mit den Spalten und Werten im Bereich **Ausdruck** angezeigt.
 - Wählen Sie im Bereich **Spalten** eine Spalte aus und klicken Sie auf die Schaltfläche mit den nach rechts weisenden Pfeilen (>>). Die Spalte wird im Bereich **Ausdruck** angezeigt.
Fügen Sie Funktionen, Ausdrücke und Werte hinzu, um einen erweiterten Filter zu erstellen.
9. Klicken Sie zum Überprüfen des erweiterten Filters auf **Validieren**.
10. Nach dem Erstellen oder Bearbeiten des Filters wählen Sie **Datenvorschau** aus, um die gefilterten Daten anzuzeigen. Sie können die Option **Max. Anzahl Zeilen in der Vorschau** festlegen.
11. Klicken Sie auf **Fertig stellen**.
Der Assistent **Neues Profil** wird mit dem Filter in der Ansicht **Filter** angezeigt.

Erstellen eines SQL-Filters

Sie können einen SQL-Filter mit SQL-Abfragen erstellen. Sie können einen SQL-Filter für relationale Datenquellen erstellen.

1. Erstellen oder öffnen Sie ein einzelnes Datenobjektprofil.
2. Wählen Sie die Ansicht **Filter** aus.
3. Klicken Sie auf **Hinzufügen**.
Das Dialogfeld **Assistenten auswählen** wird angezeigt.
4. Klicken Sie im Dialogfeld **Assistenten auswählen** auf **SQL-Filter**.
Das Dialogfeld **Filter** wird angezeigt.
5. Geben Sie einen Namen und eine optionale Beschreibung für den erweiterten Filter ein.
6. Wählen Sie **Als aktiv festlegen** aus, um den Filter auf das Profil anzuwenden. Klicken Sie auf **Weiter**.

7. Wählen Sie **Filterdefinition** aus, um einen Filter zu definieren.
8. Verwenden Sie die Spalten im Bereich **Spalten**, um einen SQL-Filter zu erstellen.
9. Klicken Sie zum Überprüfen des Filters auf **Validieren**.
10. Nach dem Erstellen oder Bearbeiten des Filters wählen Sie **Datenvorschau** aus, um die gefilterten Daten anzuzeigen. Sie können die Option **Max. Anzahl Zeilen in der Vorschau** festlegen.
11. Klicken Sie auf **Fertig stellen**.
Der Assistent **Neues Profil** wird mit dem Filter in der Ansicht **Filter** angezeigt.

Stichprobenoptionen

Die Stichprobenoptionen bestimmen die Anzahl der Zeilen, auf denen das Developer Tool ein Profil ausführt. Sie können Stichprobenoptionen konfigurieren, wenn Sie ein Profil definieren oder ausführen.

Die folgende Tabelle beschreibt die Stichprobenoptionen für ein Profil:

Eigenschaft	Beschreibung
Alle Zeilen	Führt ein Profil für alle Zeilen im Datenobjekt aus. Unterstützt in der nativen Laufzeitumgebung und in der Blaze- und Spark-Laufzeitumgebung.
Stichprobe erste <number> Zeilen	Führt ein Profil auf Stichprobenzeilen aus, die sich am Anfang der Zeilen des Datenobjekts befinden. Sie können maximal 2.147.483.647 Zeilen auswählen. Unterstützt in der nativen Laufzeitumgebung und in der Blaze-Laufzeitumgebung.
Zufällige Stichprobe <number> Zeilen	Führt ein Profil auf einer zufällig ausgewählten Anzahl von Zeilen im Datenobjekt aus. Sie können maximal 2.147.483.647 Zeilen auswählen. Unterstützt in der nativen Laufzeitumgebung und in der Blaze-Laufzeitumgebung.
Zufallsstichprobe (auto)	Führt ein Profil auf den Stichprobenzeilen aus, die auf der Grundlage der Anzahl der Zeilen im Datenobjekt berechnet werden. Unterstützt in der nativen Laufzeitumgebung und in der Blaze-Laufzeitumgebung.
Limit n <number> Zeilen	Führt ein Profil basierend auf der Anzahl der Zeilen im Datenobjekt aus. Wenn Sie ein Profil in der Hadoop-Validierungsumgebung ausführen, sammelt die Spark-Engine Stichproben von mehreren Partitionen des Datenobjekts und überträgt die Stichproben in einen einzigen Knoten, um den Stichprobenumfang zu berechnen. Die Stichprobenoption „Limit n“ unterstützt nur Oracle-, SQL Server- und DB2-Datenbanken. Sie können den erweiterten Filter nicht zusammen mit der Stichprobenoption „Limit n“ anwenden. Unterstützt in der Spark-Laufzeitumgebung.
Zufallsprozentsatz	Führt ein Profil auf einem Prozentsatz von Zeilen im Datenobjekt aus. Unterstützt in der Spark-Laufzeitumgebung.
Genehmigte Datentypen und Datendomänen in den nachfolgenden Profilausführungen aus der Datentyp- und Datendomänenableitung ausschließen	Schließt den genehmigten Datentyp oder die genehmigte Datendomäne bei der nächsten Profilausführung aus der Datentyp- und Datendomänenableitung aus.

Nachdem Sie das Profil für eine zufällige Stichprobe von Zeilen ausgeführt haben, wählt der zufällige Stichproben-Algorithmus die Zeilen zufällig in dem Datenobjekt aus, auf dem das Profil ausgeführt werden soll. Wenn Sie eine zufällige Stichproben-Option für Spaltenprofile auswählen, führt das Developer Tool ein Drilldown für die Staging-Daten durch. Dies kann sich auf die Drilldown-Leistung auswirken. Wenn Sie eine zufällige Stichproben-Option für Datendomänenenerkennungs-Profile auswählen, führt das Developer Tool ein Drilldown auf Live-Daten aus.

Laufzeitumgebung

Wählen Sie „Nativ“ oder „Hadoop“ als Laufzeitumgebung für ein Spaltenprofil aus. Sie können in der Hadoop-Laufzeitumgebung die Option „Blaze“ oder „Spark“ wählen. Informatica Developer legt die Laufzeitumgebung in der Profildefinition fest, nachdem Sie eine Laufzeitumgebung ausgewählt haben.

Native Umgebung

Wenn Sie ein Profil in der nativen Laufzeitumgebung ausführen, übermittelt das Developer Tool die Profiljobs an das Profilerstellungsdienst-Modul. Das Profilerstellungsdienst-Modul teilt die Profiljobs dann in mehrere Zuordnungen auf. Der Datenintegrationsdienst führt diese Zuordnungen auf demselben Computer aus, auf dem der Datenintegrationsdienst ausgeführt wird, und schreibt die Profilergebnisse in das Profiling-Warehouse. Standardmäßig werden alle Profile in der nativen Laufzeitumgebung ausgeführt.

Sie können native Quellen verwenden, um Profile in der nativen Umgebung zu erstellen und auszuführen. Bei einer nativen Datenquelle handelt es sich um eine Nicht-Hadoop-Quelle, wie z. B. eine Einfachdatei, eine relationale oder Mainframe-Quelle. Sie können ein Profil auch in einer Zuordnungsspezifikation oder einer logischen Datenquelle mit einer Hive- oder HDFS-Datenquelle in der nativen Umgebung ausführen.

Hadoop-Umgebung

Sie können die Option „Blaze“ oder „Spark“ auswählen, um die Profile in der Hadoop-Laufzeitumgebung auszuführen.

Nach Auswahl der Option „Blaze“ können Sie eine Hadoop-Verbindung auswählen. Der Datenintegrationsdienst übergibt die Profillogeik an die Blaze-Engine auf dem Hadoop-Cluster, um Profile auszuführen.

Wenn Sie ein Profil in der Hadoop-Umgebung ausführen, übermittelt das Developer Tool die Profiljobs an das Profilerstellungsdienst-Modul. Das Profilerstellungsdienst-Modul teilt die Profiljobs dann in mehrere Zuordnungen auf. Der Datenintegrationsdienst übergibt die Zuordnungen über die Hadoop-Verbindung an die Blaze-Engine. Die Blaze-Engine verarbeitet die Zuordnungen und der Datenintegrationsdienst schreibt die Profilergebnisse in das Profiling-Warehouse.

Nach Auswahl der Option „Spark“ können Sie eine Hadoop-Verbindung auswählen. Der Datenintegrationsdienst übergibt die Profillogeik an die Spark-Engine auf dem Hadoop-Cluster, um Profile auszuführen. Wenn Sie ein Profil in der Hadoop-Umgebung ausführen, übermittelt das Developer Tool die Profiljobs an das Profilerstellungsdienst-Modul. Das Profilerstellungsdienst-Modul teilt die Profiljobs dann in mehrere Zuordnungen auf. Der Datenintegrationsdienst übergibt die Zuordnungen über die Hadoop-Verbindung an die Spark-Engine. Die Spark-Engine verarbeitet die Zuordnungen und der Datenintegrationsdienst schreibt die Profilergebnisse in das Profiling-Warehouse.

Spaltenprofile für Sqoop-Datenquellen

Sie können ein Spaltenprofil für Datenobjekte ausführen, die Sqoop verwenden. Nachdem Sie Hadoop als Validierungsumgebung ausgewählt haben, können Sie die Blaze-Engine oder die Spark-Engine auf der Hadoop-Verbindung auswählen, um die Spaltenprofile auszuführen.

Beim Ausführen eines Spaltenprofils für ein logisches oder benutzerdefiniertes Datenobjekt können Sie das `num-mappers`-Argument konfigurieren, um Parallelismus zu erreichen und die Leistung zu optimieren. Sie müssen darüber hinaus das `split-by`-Argument konfigurieren, um die Spalte anzugeben, auf deren Grundlage Sqoop die Arbeitseinheiten aufteilen muss.

Verwenden Sie folgende Syntax:

```
--split-by <column_name>
```

Wenn der Primärschlüssel keine gleichmäßige Verteilung der Werte zwischen dem minimalen und maximalen Bereich aufweist, können Sie das `split-by`-Argument zur Angabe einer anderen Spalte mit gleichmäßiger Datenverteilung konfigurieren, um die Arbeitseinheiten aufzuteilen.

Wenn Sie die `split-by`-Spalte nicht definieren, teilt Sqoop die Arbeitseinheiten basierend auf folgenden Kriterien auf:

- Wenn das Datenobjekt einen einzigen Primärschlüssel enthält, verwendet Sqoop den Primärschlüssel als `split-by`-Spalte.
- Wenn das Datenobjekt einen zusammengesetzten Primärschlüssel enthält, verarbeitet Sqoop zusammengesetzte Primärschlüssel standardmäßig ohne das `split-by`-Argument. Weitere Informationen finden Sie in der Sqoop-Dokumentation.
- Wenn ein Datenobjekt zwei Tabellen mit einer identischen Spalte enthält, müssen Sie die `split-by`-Spalte mit einem für Tabellen qualifizierten Namen definieren. Wenn beispielsweise `CUSTOMER` als Tabellename und `FULL_NAME` als Spaltenname verwendet werden, definieren Sie die `split-by`-Spalte folgendermaßen:

```
--split-by CUSTOMER.FULL_NAME
```
- Wenn das Datenobjekt keinen Primärschlüssel enthält, wird für den Wert des `m`-Arguments und des `num-mappers`-Arguments standardmäßig 1 verwendet.

Wenn Sie Cloudera Connector Powered by Teradata oder Hortonworks Connector for Teradata verwenden und die Tabelle keinen Primärschlüssel enthält, wird das `split-by`-Argument benötigt.

Primärschlüssel-Erkennung

Bei der Analyse der Primärschlüssel werden aus den angegebenen Spalten Primärschlüsselkandidaten generiert.

Ein Primärschlüssel ist eine Spalte oder Kombination von Spalten mit einer einmaligen Kennung für eine Zeile in einer Datenquelle. Die Primärschlüsselanalyse erkennt die Spalten und Spaltenkombinationen einer bestimmten Sicherheitsstufe. Die Sicherheitsstufe sowie die maximale Anzahl der für Primärschlüsselkennungen zu kombinierenden Spalten können bearbeitet werden.

Mit Hilfe der Primärschlüsselerkennung können Sie potenzielle Datenqualitätsprobleme erkennen, indem die Zeilen, die in einem Primärschlüsselkandidaten nicht einmalig vorkommen, ermittelt werden. Diese Funktion ist insbesondere dann von Nutzen, wenn bei der Primärschlüsselerkennung viele Spalten kombiniert werden, da nicht konforme Datensätze wahrscheinlich doppelte Informationen enthalten.

Rückschlusseigenschaften von Primärschlüsseln

Wenn Sie ein einzelnes Datenobjektprofil erstellen, können Sie die Ansicht **Primärschlüssel-Profil** verwenden, um die Rückschlusseigenschaften von Primärschlüsseln zu konfigurieren.

Die folgende Tabelle beschreibt die Rückschlusseigenschaften von Primärschlüsseln in der Ansicht **Primärschlüssel-Profil**:

Eigenschaft	Beschreibung
Überschreiben der Standard-Rückschlussoptionen	Ermöglicht das Konfigurieren der benutzerdefinierten Rückschlusseigenschaften von Primärschlüsseln.
Max. Schlüsselspalten	Maximale Anzahl an Spalten, die einen Primärschlüssel bilden können.
Max. Zeilen	Anzahl der zu profilierenden Zeilen.
Konformitätskriterien	Der Mindestprozentsatz oder die Maximalanzahl der Zeilen für Schlüsselverstöße, die das Profil bei der Bestimmung der Primärschlüssel zulässt.
Datenobjekte mit dokumentierten, benutzerdefiniertem Schlüssel ausschließen	Schließt Datenobjekte mit dokumentierten Primärschlüsseln oder benutzerdefinierten Primärschlüsseln aus
Schließt Datenobjekte mit genehmigtem Schlüssel aus	Schließt Datenobjekte mit genehmigten Primärschlüsseln aus

Gefolgte Primärschlüsseigenschaften

Nachdem Sie ein einzelnes Datenobjektprofil ausgeführt haben, können Sie die Ansicht **Primärschlüssel-Profil** verwenden, um die Details der gefolgerten Primärschlüssel in der Datenquelle anzuzeigen.

In der folgenden Tabelle werden die gefolgerten Primärschlüsseigenschaften in der Ansicht **Primärschlüssel-Profil** beschrieben:

Eigenschaft	Beschreibung
Spalte	Name der Spalte im Profil.
% Konformität	Prozentsatz der einmaligen Werte in der Spalte.
Duplikate	Prozentsatz der doppelten Werte einer Spalte.
% Null	Prozentsatz der Nullwerte für die Spalte.
Überprüft	Ermittelt, ob die Spalte eine Primärschlüsselspalte ist.
Ableitungsstatus	Ableitungsstatus der Spalte.
Letzte Ausführung	Datum und Uhrzeit der letzten Ausführung des Primärschlüsselprofils.

Schlüsselverstoß-Eigenschaften

Nachdem Sie ein einzelnes Datenobjektprofil ausgeführt haben, können Sie die Ansicht **Primärschlüssel-Profiling** verwenden, um die Details der Primärschlüsselverstoße in der Datenquelle anzuzeigen.

In der folgenden Tabelle sind die Eigenschaften der Schlüsselverstoße in der Ansicht **Primärschlüssel-Profiling** beschrieben:

Eigenschaft	Beschreibung
Spalte(n)	Name der Spalte(n), aus der das Profil einen Kandidatenprimärschlüssel folgert.
Anzahl der Schlüsselverstoße	Anzahl der Schlüsselverstoße im Primärschlüsselkandidaten.

Funktionale Abhängigkeitserkennung

Die funktionale Abhängigkeitserkennung liefert Informationen über die Abhängigkeiten zwischen Spaltenpaaren in einer Datenquelle.

Ein Spaltenpaar ist funktionell abhängig, wenn anhand der Werte in einer Spalte die Werte in einer anderen Spalte zuverlässig vorausgesagt werden kann. Wenn ein Datensatz beispielsweise eine Spalte Arbeitgeber-ID und eine Spalte Geburtsdatum enthält, muss die Spalte Geburtsdatum in allen Zeilen mit einer bestimmten Arbeitgeber-ID gleich sein.

Anhand funktionaler Abhängigkeiten können potenzielle Qualitätsprobleme hervorgehoben werden, indem die Datensätze bestimmt werden, die in voneinander abhängigen Spalten nicht übereinstimmen. Sind zum Beispiel 99,8 % der Zeilen in einer Datenquelle funktionell voneinander abhängig, besteht eine hohe Wahrscheinlichkeit, dass die übrigen Zeilen falsche Daten enthalten.

Funktionaler Abhängigkeitsrückschluss - Eigenschaften

Die Ansicht **Funktionaler Abhängigkeitsrückschluss** liefert Informationen über die funktionalen Abhängigkeiten zwischen Spalten.

In der folgenden Tabelle werden die Eigenschaften des funktionalen Abhängigkeitsrückschlusses in der Ansicht **Funktionales Abhängigkeits-Profiling** beschrieben:

Eigenschaft	Beschreibung
Überschreiben der Standard-Rückschlussoptionen	Ermöglicht die Konfiguration benutzerspezifischer Einstellungen für den funktionale Abhängigkeitsrückschluss.
Max. Spalten in Determinante	Anzahl der Spalten, die das Profil kombinieren kann, um eine Determinante zu finden.
Max. Zeilen	Anzahl der zu profilierenden Zeilen.
Zurückgegebene Abhängigkeiten	Anzahl der vom Profil angezeigten Abhängigkeiten. Der Standard ist Min. Abdeckung , die den kleinsten Satz von Abhängigkeiten anzeigt, wobei jede Spalte mindestens einmal in einer Determinante erscheint.

Eigenschaft	Beschreibung
Max. zurückgegebene Abhängigkeiten	Maximale Anzahl der vom Profil angezeigten Abhängigkeiten.
Konformitätskriterien	Der Mindestprozentsatz oder die Maximalanzahl der Zeilen für Abhängigkeitsverstöße, die das Profil bei der Bestimmung der funktionalen Abhängigkeiten zulässt.

Gefolgerte funktionale Abhängigkeit - Eigenschaften

Nachdem Sie ein einzelnes Datenobjektprofil ausgeführt haben, können Sie die Ansicht **Funktionaler Abhängigkeitsrückschluss** verwenden, um die Details der gefolgerten funktionalen Abhängigkeiten zwischen Spalten in der Datenquelle anzuzeigen.

Die folgende Tabelle beschreibt die Eigenschaften der gefolgerten funktionalen Abhängigkeit in der Ansicht **Funktionaler Abhängigkeitsrückschluss**:

Eigenschaft	Beschreibung
Determinantenspalten	Name der auf funktionale Abhängigkeiten hin analysierten Spalte.
Abhängige Spalten	Name der Spalte, die von der Determinantenspalte abhängig ist.
% Null	Prozentsatz der Nullwerte für die Spalte.
% Konformität	Prozentsatz des Matches der funktionalen Abhängigkeit
Überprüft	Gibt an, ob die Spalten funktional abhängig sind.
Letzte Ausführung	Datum und Uhrzeit der letzten Ausführung des funktionalen Abhängigkeitsprofils.

Funktionale Abhängigkeitsverstöße - Eigenschaften

Die Ansicht zeigt Informationen über die funktionalen Abhängigkeiten zwischen Spalten. Nachdem Sie ein einzelnes Datenobjektprofil ausgeführt haben, können Sie die Ansicht **Funktionaler Abhängigkeitsrückschluss** verwenden, um die Details der funktionalen Abhängigkeitsverstöße in der Datenquelle anzuzeigen.

In der folgenden Tabelle werden die Eigenschaften der funktionalen Abhängigkeitsverstöße in der Ansicht **Funktionales Abhängigkeits-Profiling** beschrieben:

Eigenschaft	Beschreibung
Determinantenspalte	Name der auf funktionale Abhängigkeiten hin analysierten Spalte.
Eigenständige Abhängige	Anzahl der einmaligen funktionalen Abhängigkeiten.

Betriebssystemprofile in Informatica Developer

Sie können ein Betriebssystemprofil im Developer Tool auswählen. Nach Auswahl eines Betriebssystemprofils erstellt der Datenintegrationsdienst die Spalten- und Enterprise-Erkennungsprofile und führt sie aus. Darüber hinaus werden Scorecards basierend auf der Berechtigung des Betriebssystemprofilbenutzers erstellt.

Auswählen eines Betriebssystemprofils

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

1. Klicken Sie in Informatica Developer auf **Windows > Einstellungen**.
Das Dialogfeld **Einstellungen** wird geöffnet.
2. Klicken Sie auf **Informatica > Konfigurationen ausführen > Zuordnung**.
Das Dialogfeld **Zuordnung** wird geöffnet.
3. Deaktivieren Sie im Feld **Zuordnung** die Option **Standardmäßigen Datenintegrationsdienst verwenden**.
4. Klicken Sie auf **Durchsuchen**, um ein Betriebssystemprofil in der Liste auszuwählen.
5. Klicken Sie auf **OK**.

Erstellen eines einzelnen Datenobjektprofils in Informatica Developer

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

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

Hinweis: Wählen Sie zum Aktivieren einer Profioperation **Als Teil der Aktion "Profil ausführen" aktiviert** für die betreffende Operation. Standardmäßig ist Spalten-Profiling aktiviert.

9. Prüfen Sie die Optionen Ihres Profils.

Sie können die Spaltenauswahl für alle Profiltypen ändern. Überprüfen Sie die Filter- und Stichprobenoptionen für die Spaltenprofile. Sie können die Interferenzoptionen für den Primärschlüssel, die funktionale Abhängigkeit und die Datendomänenerkennung überprüfen. Sie können ebenfalls die Datendomänenauswahl für die Datendomänenerkennung überprüfen.

10. Überprüfen Sie die Drilldown-Optionen und bearbeiten Sie sie gegebenenfalls. Standardmäßig ist die Option **Zeilen-Drilldown aktivieren** ausgewählt. Sie können Drilldown-Optionen für Spaltenprofile bearbeiten. Die Optionen bestimmen auch, ob Drilldown-Vorgänge aus der Datenquelle oder aus Staging-Daten gelesen werden sollen und ob das Profil Ergebnisdaten aus vorherigen Profildurchläufen speichert.
11. Wählen Sie im Abschnitt **Einstellungen ausführen** eine Validierungsumgebung aus. Wählen Sie **Nativ** oder **Hadoop** als Validierungsumgebung aus. Sie können **Nativ**, **Blaze** oder **Spark** als Laufzeitumgebung auswählen. Nach Auswahl der Option **Blaze** oder **Spark** können Sie eine Hadoop-Verbindung auswählen.
12. Klicken Sie auf **Fertigstellen**.

Erstellen mehrerer Datenobjektprofile in Informatica Developer

Wenn Sie das Profil für mehrere Datenobjekte für mehrere Datenobjekte ausführen, verwendet das Developer Tool die Standardoptionen für Spaltenprofile, um Spaltenprofile für mindestens ein Datenobjekt zu erzeugen. Optional können Sie ein Enterprise-Erkennungsprofil erstellen, um ein Profil für mehrere Datenobjekte auszuführen.

1. Wählen Sie in der Ansicht **Objekt-Explorer** das Datenobjekt aus, das Sie profilieren möchten.
2. Klicken Sie auf **Datei > Neu > Profil** um den Assistenten **Neues Profil** zu öffnen.
3. Wählen Sie im Assistenten **Neu** die Option **Mehrere Profile** aus und klicken Sie auf **Weiter**.
4. Wählen Sie im Fenster **Mehrere Profile** den Speicherort aus, in dem die Profile erstellt werden sollen. Sie können jedes Profil an demselben Speicherort wie dessen profiliertes Objekt erstellen oder einen gemeinsamen Speicherort für die Profile angeben.
5. Überprüfen Sie, dass die Namen der von Ihnen ausgewählten Datenobjekte im Abschnitt **Datenobjekte** erscheinen.
Optional können Sie auch auf **Hinzufügen** klicken, um ein anderes Datenobjekt hinzuzufügen.
6. Geben Sie optional die Anzahl der zu profilierenden Zeilen an, und wählen Sie, ob das Profil ausgeführt werden soll, wenn der Assistent beendet ist.
7. Klicken Sie auf **Weiter**.
8. Wählen Sie im Abschnitt **Validierungsumgebung** die Option **Nativ** aus.
Hinweis: Wählen Sie nur die Option „Nativ“ aus, um das Profil für mehrere Datenobjekte auszuführen. Zum Ausführen mehrerer Datenobjekte mit der Blaze- oder Spark-Engine in der Hadoop-Laufzeitumgebung können Sie das Enterprise-Erkennungsprofil auswählen.
9. Klicken Sie auf **Fertigstellen**.

10. Geben Sie optional die Präfix- und Suffix-Strings ein, die den Profilnamen hinzugefügt werden sollen.
11. Klicken Sie auf **OK**.

Bearbeiten eines Profils

Sie können ein einzelnes Datenobjektprofil oder mehrere Datenobjektprofile bearbeiten. Falls das Versionsverwaltungssystem aktiviert ist, ist das Profil standardmäßig ausgecheckt.

1. Klicken Sie in der Ansicht **Object Explorer** mit der rechten Maustaste auf das Profil und klicken Sie auf **Öffnen**.

Die Ansicht **Ergebnisse** wird angezeigt.

2. Aktualisieren Sie in der Ansicht **Definition** nach Bedarf die Eigenschaften.
3. Klicken Sie auf **Team > Einchecken**, um das Profil einzuchecken.
4. Klicken Sie mit der rechten Maustaste auf das Profil und klicken Sie auf **Profil ausführen**, um das Profil auszuführen.

Die Profilergebnisse werden in der Ansicht **Ergebnisse** angezeigt.

Synchronisierungsoption

Wenn Sie die Metadaten einer externen Datenquelle ändern, werden die Datenobjekt-Metadaten im Modellrepository standardmäßig nicht aktualisiert. Verwenden Sie die Synchronisierungsoption, um die Datenobjekt-Metadaten mit den Metadaten der Datenquelle zu synchronisieren.

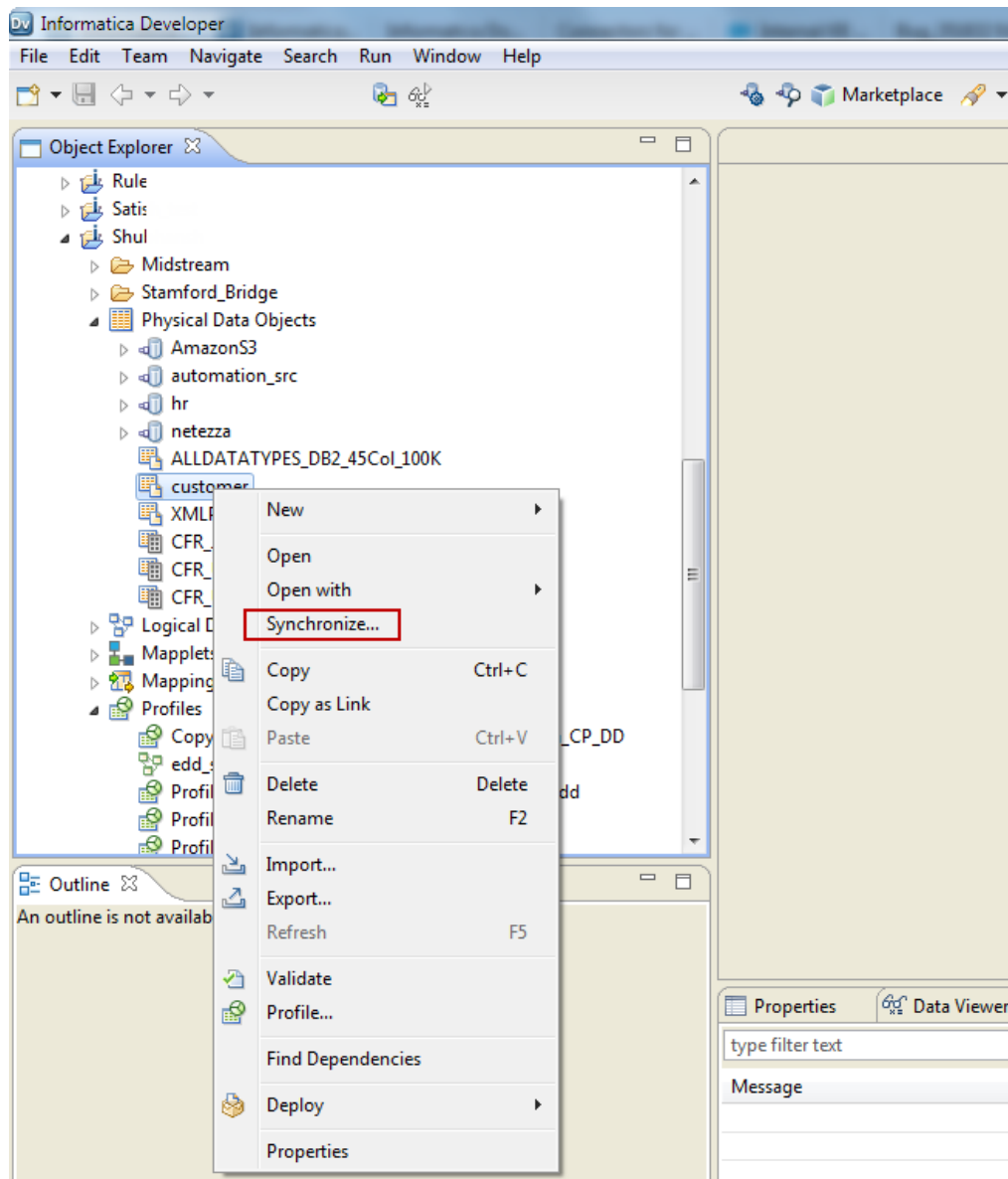
Wenn Sie im Developer Tool die Synchronisierungsoption verwendet haben, wird beim Öffnen eines Profils oder einer Scorecard, das bzw. die das Datenobjekt verwendet, neben dem Namen des Profils oder der Scorecard im Editor ein Sternchen angezeigt. Das Sternchen gibt an, dass die Metadaten des Datenobjekts für das Profil oder die Scorecard geändert wurden. Öffnen und speichern Sie das Profil oder die Scorecard, um die Profildefinition im Modellrepository zu aktualisieren. Beachten Sie, dass im Analyst Tool keine sichtbaren Änderungen angezeigt werden, wenn Sie das Profil oder die Scorecard nach der Synchronisierung des Datenobjekts für das Profil oder die Scorecard öffnen. Sie können die Synchronisierungsoption für Spaltenprofile, Enterprise-Erkennungsprofile und Scorecards verwenden. Die externe Datenquelle kann eine relationale Datenquelle oder eine Einfachdatei-Datenquelle sein.

Synchronisieren eines Einfachdatei-Datenobjekts in Informatica Developer

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

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

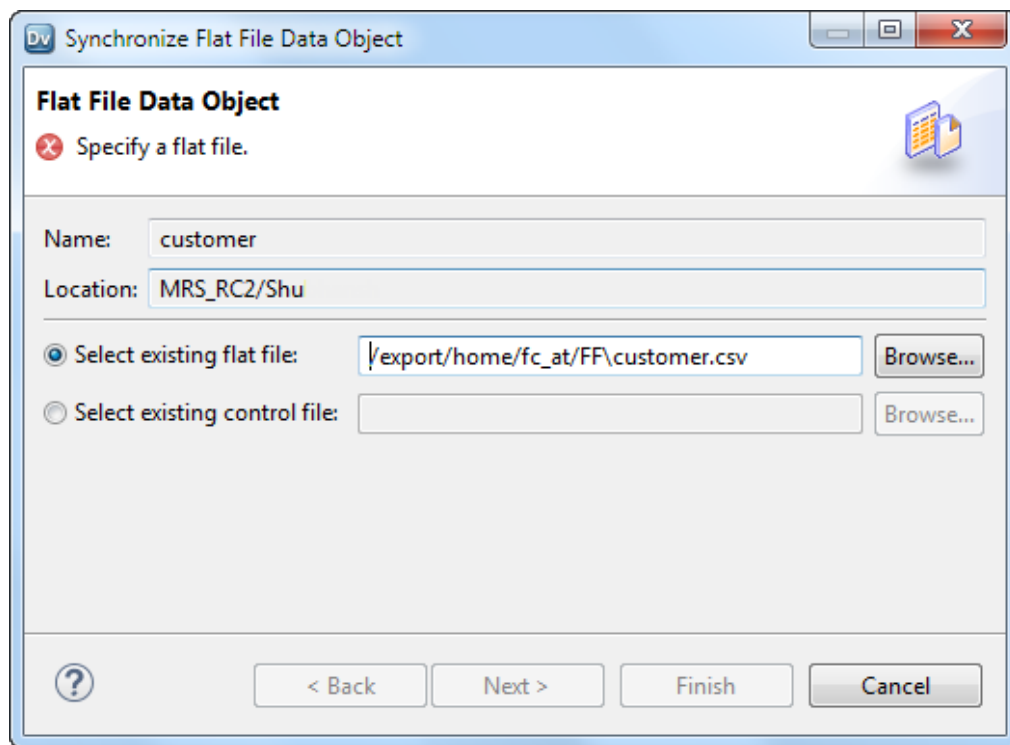
Die folgende Abbildung zeigt die Synchronisierungsoption für ein Datenobjekt:



Der Assistent zum **Synchronisieren eines Einfachdatei-Datenobjekts** wird angezeigt.

3. Überprüfen Sie im Assistenten zum **Synchronisieren eines Einfachdatei-Datenobjekts** den Pfad der Einfachdatei im Feld **Vorhandene Einfachdatei auswählen**.

Die folgende Abbildung zeigt den Assistenten zum Synchronisieren eines Einfachdatei-Datenobjekts:



4. Klicken Sie auf **Weiter**.
5. Wählen Sie optional die Codeseite, das Format, die Eigenschaften für das Trennzeichenformat und die Spalteneigenschaften.
6. Klicken Sie auf **Fertig stellen** und anschließend auf **OK**.

Synchronisieren eines relationalen Datenobjekts in Informatica Developer

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

1. Wählen Sie in der Ansicht **Objekt-Explorer** ein relationales Datenobjekt aus.
2. Klicken Sie mit der rechten Maustaste, und wählen Sie **Synchronisieren**.
Sie werden aufgefordert, die Aktion zu bestätigen.
3. Um den Synchronisierungsprozess abzuschließen, klicken Sie auf **OK**.
Es wird eine Statusmeldung für den Synchronisierungsprozess angezeigt.
4. Wenn Sie die Meldung **Synchronisierung beendet** sehen, klicken Sie auf **OK**.
In Verbindung mit der Meldung wird eine Zusammenfassung der an dem Datenobjekt vorgenommenen Metadatenänderungen angezeigt.

Kommentare

Sie können eine Beschreibung als Kommentar zu einem Profil hinzufügen. Außerdem können Sie den Spalten in den Spaltenprofilergebnissen Kommentare hinzufügen.

Sie können Spalten mehrere Kommentare hinzufügen. Sie können die Ansicht **Kommentare** im Developer Tool verwenden, um die Kommentare hinzuzufügen oder anzuzeigen.

Hinzufügen von Kommentaren in Informatica Developer

Fügen Sie Spalten in Spaltenprofilergebnissen Kommentare hinzu. Beim Exportieren der Profilergebnisse enthält das Developer Tool die Kommentare.

1. Öffnen Sie in der Ansicht **Objekt-Explorer** ein Profil.
2. Optional führen Sie das Profil aus, um die Profilergebnisse zu aktualisieren.
3. Wählen Sie die Ansicht **Kommentare** aus.
4. Klicken Sie auf **Hinzufügen**, um das Dialogfeld **Kommentar hinzufügen** zu öffnen.
5. Wählen Sie in der Liste den Profilnamen oder eine der Spalten aus.
Wenn Sie zuvor Kommentare hinzugefügt haben, können Sie die Kommentare im Dialogfeld anzeigen.
6. Geben Sie im Feld **Kommentar** eine Beschreibung ein.
7. Klicken Sie auf **OK**.
Das Developer Tool zeigt den Kommentar in der Ansicht **Kommentare** an.
8. Wählen Sie einen zu löschenden Kommentar in der Ansicht **Kommentare** aus und klicken Sie auf **Entfernen**.

KAPITEL 19

Spaltenprofile in halbstrukturierten Datenquellen

Dieses Kapitel umfasst die folgenden Themen:

- [Spaltenprofile in halbstrukturierten Datenquellen - Übersicht, 164](#)
- [JSON- und XML-Datenobjekte, 165](#)
- [Komplexe Dateidatenobjekte für halbstrukturierte Datenquellen in HDFS, 166](#)
- [Erstellen einer HDFS-Verbindung, 167](#)
- [Erstellen eines komplexen Dateidatenobjekts anhand einer JSON- oder XML-Datei in HDFS, 167](#)
- [Erstellen eines komplexen Dateidatenobjekts aus einer Avro- oder Parquet-Datenquelle, 168](#)
- [Erstellen eines Spaltenprofils in einer halbstrukturierten Datenquelle, 169](#)

Spaltenprofile in halbstrukturierten Datenquellen - Übersicht

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

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

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

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

JSON- und XML-Datenobjekte

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

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

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

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

Erstellen eines Datenobjekts aus einer JSON- oder XML-Datenquelle

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

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

Das Dialogfeld **Neu** wird eingeblendet.

3. Sie können ein Einfachdatei-Datenobjekt und ein komplexes Dateidatenobjekt erstellen.
 - Führen Sie zum Erstellen eines Einfachdatei-Datenobjekts die folgenden Aufgaben durch:
 1. Wählen Sie **Physische Datenobjekte > Einfachdatei-Datenobjekt** aus und klicken Sie auf **Weiter**. Das Dialogfeld **Neues Einfachdatei-Datenobjekt** wird eingeblendet.
 2. Wählen Sie **Aus vorhandener Einfachdatei erstellen** aus und klicken Sie auf **Durchsuchen**, um die Textdatei auszuwählen. Klicken Sie auf **Weiter**.
 3. Stellen Sie sicher, dass die Codepage **MS Windows Latin 1 (ANSI)**, eine Obermenge von Latin 1 und das delimitergetrennte Format verwendet werden. Klicken Sie auf **Weiter**.
 4. Stellen Sie sicher, dass der Delimiter auf **Komma** gesetzt ist. Klicken Sie auf **Fertig stellen**.
 - Führen Sie zum Erstellen eines komplexen Dateidatenobjekts die folgenden Aufgaben durch:
 1. Wählen Sie **Physische Datenobjekte > Komplexes Dateidatenobjekt** aus und klicken Sie auf **Weiter**. Das Dialogfeld **Neues komplexes Dateidatenobjekt** wird eingeblendet.
 2. Geben Sie einen Namen für das Datenobjekt ein. Wählen Sie **Datei** als Zugriffstyp aus.
 3. Klicken Sie auf **Durchsuchen**, um eine JSON- oder XML-Datei auszuwählen. Klicken Sie auf **Fertigstellen**.

Wenn der Developer-Server unter Linux ausgeführt wird, müssen Sie den Dateipfad der Datenquelle zum Speicherort auf dem Server aktualisieren. Wählen Sie zum Aktualisieren des Dateipfads das komplexe Datei-Datenobjekt aus, klicken Sie auf der Registerkarte **Datenobjekt-Vorgänge** auf **Lesen** und fügen Sie den Dateipfad auf der Registerkarte **Erweitert** im Bereich **Details zu Datenobjekt-Vorgang** hinzu.

Das Datenobjekt wird im Projektordner angezeigt.

Komplexe Dateidatenobjekte für halbstrukturierte Datenquellen in HDFS

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

Komplexes Dateidatenobjekt aus einer JSON- oder XML-Datenquelle in HDFS

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

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

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

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

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

Komplexes Dateidatenobjekt aus einer Avro- oder Parquet-Datenquelle in HDFS

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

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

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

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

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

Erstellen einer HDFS-Verbindung

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

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

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

Sie können ein komplexes Dateidatenobjekt anhand einer JSON- oder XML-Quelldatei erstellen, die HDFS verwendet. Weiterhin können Sie ein Spaltenprofil anhand des Datenobjekts erstellen.

1. Wählen Sie im Developer Tool in der Ansicht **Objekt-Explorer** das Projekt aus, in dem Sie das physische Datenobjekt und das Spaltenprofil erstellen möchten.
2. Klicken Sie auf **Datei > Neu > Datenobjekt**.
Das Dialogfeld **Neu** wird eingeblendet.
3. Wählen Sie **Physische Datenobjekte > Komplexes Dateidatenobjekt** aus und klicken Sie auf **Weiter**.
Das Dialogfeld **Neues komplexes Dateidatenobjekt** wird eingeblendet.
4. Geben Sie einen Namen für das Datenobjekt ein. Wählen Sie **Verbindung** als Zugriffstyp aus.
5. Sie können ein Datenobjekt anhand einer JSON- oder XML-Datei oder in einem Ordner erstellen, der mehrere JSON- und XML-Dateien enthält.
 - Führen Sie die folgenden Schritte durch, um ein komplexes Dateidatenobjekt anhand einer JSON- oder XML-Datei zu erstellen:
 1. Klicken Sie auf **Durchsuchen**, um eine Verbindung auszuwählen.
 2. Klicken Sie im Dialogfeld **Ressource hinzufügen** auf **Hinzufügen**, um eine JSON- oder XML-Datei auszuwählen.
 3. Klicken Sie auf **Fertig stellen**.
Das Datenobjekt wird im Projektordner angezeigt.

- Führen Sie die folgenden Schritte durch, um ein komplexes Dateidatenobjekt in einem Ordner zu erstellen, der mehrere JSON- oder XML-Dateien enthält:
 1. Klicken Sie auf **Durchsuchen**, um eine Verbindung auszuwählen.
 2. Klicken Sie im Dialogfeld **Ressource hinzufügen** auf **Hinzufügen**, um eine JSON- oder XML-Datei im Ordner auszuwählen.
 3. Klicken Sie auf **Fertig stellen**.
Das Datenobjekt wird im Projektordner angezeigt.
 4. Wählen Sie das Datenobjekt im Projektordner aus und klicken Sie auf **Erweitert > Laufzeit: Lesen > Quelldateiverzeichnis**.
 5. Entfernen Sie den Dateinamen und behalten Sie den Ordernamen im Dateipfad bei.

Erstellen eines komplexen Dateidatenobjekts aus einer Avro-oder Parquet-Datenquelle

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

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

7. Geben Sie **com.informatica.avro.AvroToXML** im Feld **Eingabeformat** für Avro-Datenquellen und **com.informatica.parquet.ParquetToXML** im Feld **Eingabeformat** für Parquet-Datenquellen ein. Wenn Sie das Eingabeformat hinzufügen, verarbeitet die Datenprozessorumwandlung die Datenquellen im Avro- oder Parquet-Format und wandelt sie zur Laufzeit in eine Datenquelle im XML-Format um.
- Wenn Sie **Datei** als Zugriffstyp und **Avro** oder **Parquet** als Ressourcenformat auswählen, führen Sie die folgenden Schritte durch:
 1. Klicken Sie auf **Durchsuchen**, um eine Avro- oder Parquet-Datei auf dem lokalen Computer auszuwählen.
 2. Klicken Sie im Dialogfeld **Neues komplexes Dateidatenobjekt** auf **Fertigstellen**. Das Datenobjekt wird im Projektordner angezeigt.
 3. Navigieren Sie nach dem Erstellen des Datenobjekts zu **Datenobjektvorgänge > Lesen > Erweitert** (Registerkarte) und stellen Sie sicher, dass der Dateipfad im Feld **Dateipfad** der Datenquelle auf dem Linux- oder Windows-Computer entspricht.

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

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

Erstellen eines Spaltenprofils in einer halbstrukturierten Datenquelle

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

1. Wählen Sie in der Ansicht **Objekt-Explorer** das Datenobjekt für die Avro-, JSON-, Parquet- oder XML-Datei aus.
2. Klicken Sie auf **Datei > Neu > Profil**.
Das Dialogfeld **Neu** wird eingeblendet.
3. Wählen Sie **Profil** aus. Klicken Sie auf **Weiter**.
Das Dialogfeld **Neues Profil** wird geöffnet.
4. Fügen Sie im Dialogfeld **Neues Profil** einen Namen für das Profil und eine optionale Beschreibung hinzu.
5. Wählen Sie die Option **Erweiterte Dateiformate verarbeiten** aus. Klicken Sie auf **Weiter**.
Die folgende Abbildung zeigt den Assistenten **Neues Profil** mit der Option **Erweiterte Dateiformate verarbeiten**:

1. Erweiterte Dateiformate verarbeiten. Wählen Sie diese Option aus, um teilstrukturierte Datenquellen zu verarbeiten.

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

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

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

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

KAPITEL 20

Regeln in Informatica Developer

Dieses Kapitel umfasst die folgenden Themen:

- [Regeln in Informatica Developer - Übersicht, 171](#)
- [Erstellen einer Regel in Informatica Developer, 172](#)
- [Anwenden einer Regel in Informatica Developer, 172](#)

Regeln in Informatica Developer - Übersicht

Eine Regel ist eine Geschäftslogik, die Bedingungen definiert, welche auf Quelldaten angewendet werden, wenn ein Spaltenprofil ausgeführt wird. Sie können eine Regel zum Profil hinzufügen, um Daten zu validieren. Sie können Mapplets verwenden, die als Regeln, vordefinierte Regeln oder wiederverwendbare Regeln in den Spaltenprofilen validiert werden.

Sie können die folgenden Methoden verwenden, um Regeln in den Spaltenprofilen zu verwenden:

- Erstellen Sie im Developer Tool ein Mapplet und validieren Sie es als Regel. Die Regel erscheint im Analyst Tool als wiederverwendbare Regel. Sie können die Regel auf Spaltenprofile im Analyst Tool und Developer Tool anwenden.
- Sie können vordefinierte Regeln in den Spaltenprofilen verwenden. Informatica stellt die vordefinierten Regeln mit dem Developer Tool und dem Analyst Tool bereit.
- Erstellen Sie im Analyst Tool eine Regelspezifikation und generieren Sie ein Mapplet. Sie können die Regelspezifikation auf Spaltenprofile im Analyst Tool anwenden. Validieren Sie im Developer Tool das Mapplet als Regel. Die Regel wird als wiederverwendbare Regel angezeigt, die Sie in den Spaltenprofilen verwenden können.

Hinweis: Im Developer Tool können Sie keine Regelspezifikationen in einem Spaltenprofil hinzufügen, bearbeiten oder löschen.

Eine Regel muss folgende Anforderungen erfüllen:

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

Erstellen einer Regel in Informatica Developer

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

Erstellen eines Mapplets im Developer-Tool.

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

Anwenden einer Regel in Informatica Developer

Sie haben die Möglichkeit, einer Regel ein gespeichertes Spaltenprofil hinzuzufügen. Sie können einem Profil, das für die Join-Analyse konfiguriert wurde, keine Regel hinzufügen.

1. Durchsuchen Sie die Ansicht **Objekt-Explorer** nach dem benötigten Profil.
2. Klicken Sie mit der rechten Maustaste auf das Profil, und wählen Sie **Öffnen** aus.
Das Profil wird im Editor geöffnet.
3. Klicken Sie auf die Registerkarte **Definition** und wählen Sie Regeln aus.
4. Klicken Sie auf **Hinzufügen**.
Das Dialogfenster **Regel anwenden** wird geöffnet.
5. Klicken Sie auf **Durchsuchen**, um die Regel zu finden, die Sie anwenden möchten.
Wählen Sie eine Regel aus einem Repository-Projekt, und klicken Sie auf **OK**.
6. Klicken Sie auf die Spalte **Wert** unter **Eingabewerte**, um einen Eingabeport für die Regel auszuwählen.
7. Optional klicken Sie auf die Spalte **Wert** unter **Ausgabewerte** zum Bearbeiten des Namens des Regelausgabeports.
Die Regel wird nun auf der Registerkarte **Definition** angezeigt.

KAPITEL 21

Mapplet- und Mapping-Profiling

Dieses Kapitel umfasst die folgenden Themen:

- [Mapplet- und Mapping-Profiling - Übersicht, 173](#)
- [Ausführen eines Profils für ein Mapplet- oder Mapping-Objekt, 173](#)
- [Vergleichen von Profilen für Mapping- oder Mapplet-Objekte, 174](#)
- [Generieren eines Mapping aus einem Profil, 174](#)

Mapplet- und Mapping-Profiling - Übersicht

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

Ausführen eines Profils für ein Mapplet- oder Mapping-Objekt

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

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

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

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

4. Wenn die Umwandlung mehrere Ausgabegruppen enthält, wählen Sie nach Bedarf die Ausgabegruppen.

5. Klicken Sie auf **OK**.

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

Vergleichen von Profilen für Mapping- oder Mapplet-Objekte

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

Wie Spaltenprofile einzelner Mapping- oder Mapplet-Objekte werden Vergleiche für alle Datenspalten ausgeführt und ermöglichen Drilldown-Operationen mit den Staging-Daten für die Datenobjekte. Nachdem Sie Daten aus einer Quelltable in eine Zieltabelle verschoben haben, können Sie Profile vergleichen, um die Migration der Daten zu überprüfen. Sie können auch Profile für eine Datenquelle vergleichen, die sich im Lauf der Zeit ändert.

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

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

Generieren eines Mapping aus einem Profil

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

1. Suchen Sie in der Ansicht **Objekt-Explorer** das Profil, für das Sie das Mapping erstellt haben.
2. Klicken Sie mit der rechten Maustaste auf den Profilnamen, und wählen Sie **Mapping generieren**.
Das Dialogfenster **Mapping generieren** wird angezeigt.
3. Geben Sie einen Mapping-Namen ein. Optional können Sie eine Beschreibung für das Mapping eingeben.
4. Bestätigen Sie den Ordnerspeicherplatz für das Mapping.

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

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

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

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

KAPITEL 22

Spaltenprofilerggebnisse in Informatica Developer

Dieses Kapitel umfasst die folgenden Themen:

- [Spaltenprofilerggebnisse in Informatica Developer, 176](#)
- [Spaltenwert-Eigenschaften, 177](#)
- [Spaltenmuster-Eigenschaften, 177](#)
- [Spaltenstatistik-Eigenschaften, 177](#)
- [Eigenschaften des Spaltendatentyps, 178](#)
- [Wiederherstellung in Informatica Developer, 179](#)
- [Exportieren von Profilergebnissen aus Informatica Developer, 180](#)

Spaltenprofilerggebnisse in Informatica Developer

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

Die folgende Tabelle beschreibt die Profilergebnisse für jeden Analysetyp:

Profiltyp	Profilergebnisse
Spaltenprofil	<ul style="list-style-type: none">- Prozentuale und Zählstatistiken für eindeutige und Nullwerte- Abgeleitete Datentypen- Der Datentyp, den die Datenquelle für die Daten deklariert- Die Höchst- und Mindestwerte- Datum und Uhrzeit des zuletzt ausgeführten Profildurchlaufs- Prozentuale und Zählstatistik für jedes eindeutige Datenelement in einer Spalte- Prozentuale und Zählstatistik für jedes eindeutige Zeichenmuster in einer Spalte
Primärschlüsselprofil	<ul style="list-style-type: none">- Gefolgte Primärschlüssel- Schlüssel-Verstöße
Funktionales Abhängigkeitsprofil	<ul style="list-style-type: none">- Gefolgte Funktionsabhängigkeiten- Funktionale Abhängigkeitsverstöße

Spaltenwert-Eigenschaften

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

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

Die folgende Tabelle beschreibt die Eigenschaften des Spaltenwertes:

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

Spaltenmuster-Eigenschaften

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

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

Die folgende Tabelle beschreibt die Eigenschaften für Spaltenwertmuster:

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

Spaltenstatistik-Eigenschaften

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

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

Die folgende Tabelle beschreibt die Spaltenstatistikeigenschaften:

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

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

Eigenschaften des Spaltendatentyps

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

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

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

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

Wiederherstellung in Informatica Developer

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

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

Genehmigen von Datentypen

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

1. Wählen und öffnen Sie ein Profil in der Ansicht **Objekt-Explorer**.
2. Stellen Sie sicher, dass Sie auf der Registerkarte **Ergebnisse** sind.
3. Wählen Sie in der Ansicht **Spalten-Profiling** im rechten Bereich eine Spalte zum Anzeigen der Werthäufigkeiten, Muster, Datentypen und Statistiken aus.
4. Wählen Sie im Bereich **Details** in der Liste **Anzeigen** die Option **Datentypen** aus.
Die abgeleiteten Datentypen für die Spalte werden angezeigt.
5. Klicken Sie mit der rechten Maustaste auf die Spalte, die Sie genehmigen wollen, und klicken Sie auf **Genehmigen**.
Der Status des Datentyps ändert sich in **Genehmigt**.
6. Klicken Sie zum Wiederherstellen des abgeleiteten Status des Datentyps mit der rechten Maustaste auf den Datentyp und klicken Sie auf **Zurücksetzen**.

Ablehnen von Datentypen

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

1. Wählen Sie in der Ansicht **Objekt-Explorer** ein Profil aus.
2. Doppelklicken Sie auf das Profil, um es zu öffnen.
Das Profil wird auf einer Registerkarte geöffnet.
3. Wählen Sie in der Ansicht **Spalten-Profiling** eine Zeile aus.
4. Wählen Sie zum Ablehnen abgeleiteter Spaltendatentypen die Ansicht **Datentypen** im rechten Bereich aus. Wählen Sie den abgeleiteten Datentyp aus, den Sie ablehnen möchten, klicken Sie mit der rechten Maustaste auf die Zeile und wählen Sie **Ablehnen** aus.
Informatica Developer blendet den abgelehnten Datentyp in der Liste der Datentypen aus.
5. Klicken Sie zum Ausblenden der abgelehnten Datentypen mit der rechten Maustaste auf die Zeile und wählen Sie **Abgelehnte ausblenden** aus.
6. Klicken Sie zur Anzeige der abgelehnten Datentypen mit der rechten Maustaste auf eine der Zeilen und wählen Sie **Abgelehnte einblenden** aus.

Exportieren von Profilergebnissen aus Informatica Developer

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

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

Das Dialogfenster **Daten in Datei exportieren** wird geöffnet.

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

KAPITEL 23

Scorecards in Informatica Developer

Dieses Kapitel umfasst die folgenden Themen:

- [Scorecards in Informatica Developer - Übersicht, 181](#)
- [Erstellen einer Scorecard, 181](#)
- [Exportieren einer Ressourcendatei zur Scorecard-Herkunftskontrolle, 182](#)
- [Anzeigen der Scorecard-Herkunft aus Informatica Developer, 182](#)

Scorecards in Informatica Developer - Übersicht

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

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

Erstellen einer Scorecard

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

1. Wählen Sie im **Objekt-Explorer** das Projekt oder den Ordner, in dem Sie die Scorecard erstellen möchten.
2. Klicken Sie auf **Datei > Neu > Scorecard**.
Das Dialogfenster **Neue Scorecard** wird aufgerufen.
3. Klicken Sie auf **Hinzufügen**.
Das Dialogfenster **Profil auswählen** wird angezeigt. Wählen Sie das Profil aus, das die Spalten enthält, die Sie hinzufügen möchten.
4. Klicken Sie auf **OK** und anschließend auf **Weiter**.
5. Wählen Sie die Spalten aus, die Sie zu der Scorecard hinzufügen möchten.

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

6. Klicken Sie auf **Fertigstellen**.

Das Developer Tool erstellt die Scorecard.

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

Exportieren einer Ressourcendatei zur Scorecard-Herkunftskontrolle

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

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

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

9. Klicken Sie auf **Fertigstellen**.

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

Anzeigen der Scorecard-Herkunft aus Informatica Developer

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

1. Wählen Sie in der Ansicht **Objekt-Explorer** das Projekt oder den Ordner, in dem sich die Scorecard befindet.
2. Doppelklicken Sie auf die Scorecard, um sie zu öffnen.
Die Scorecard wird auf einer Registerkarte angezeigt.
3. Klicken Sie auf **Mit Informatica Analyst öffnen**.

Das Analyst-Tool wird im Browserfenster geöffnet.

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

Das Scorecard-Herkunftsdiagramm erscheint in einem Dialogfenster.

KAPITEL 24

Datendomänenenerkennung in Informatica Developer

Dieses Kapitel umfasst die folgenden Themen:

- [Datendomänenenerkennung in Informatica Developer - Übersicht, 184](#)
- [Datendomänenglossar in Informatica Developer, 185](#)
- [Optionen der Datendomänenenerkennung in Informatica Developer, 189](#)
- [Erstellen eines Profils zur Ausführung der Datendomänenenerkennung in Informatica Developer, 191](#)
- [Bearbeiten eines Profils in Informatica Developer, 192](#)
- [Ausführen eines Profils zur Ausführung der Datendomänenenerkennung in Informatica Developer, 193](#)
- [Datendomänenenerkennungsergebnisse in Informatica Developer, 193](#)

Datendomänenenerkennung in Informatica Developer - Übersicht

Verwenden Sie das Datendomänenglossar zum Verwalten der Datendomänen. Zum Erstellen einer Datendomäne können Sie vordefinierte Regeln für Daten und Spaltennamen verwenden. Sie können eine Datendomäne basierend auf bestimmte Werten oder Mustern in den Spaltenprofilergebnissen erstellen.

Sie können die Quellspalten und Datendomänen auswählen, mit denen Sie die Spaltendaten und Spaltennamen, Stichprobenoptionen, Drilldown-Optionen und die Laufzeitumgebung abgleichen möchten. Sie können die maximale Anzahl an Zeilen auswählen, für die Datendomänenenerkennung durchgeführt werden soll. Sie können ein Konformitätskriterium für die Datendomänenenerkennung auswählen. Sie können Nullwerte aus der Datendomänenenerkennung ausschließen. Nachdem Sie ein Profil ausgeführt haben, können Sie eine Überprüfung, ein Drilldown und eine Wiederherstellung der Ergebnisse durchführen. Sie können auch die Ergebnisse einem Datenmodell über den Editor im Developer-Tool hinzufügen.

Sie können ein Profil mit einer Stichprobenoption und Filtern erstellen, um Datendomänenenerkennung durchzuführen. Beim Ausführen des Profils wenden Sie die Stichprobenoption und die Filter auf die Datenquelle an und erzeugen einen Datensatz. Der Datendomänenenerkennungs-Prozess verwendet den Datensatz zur Erkennung von Datendomänen.

Datendomänenglossar in Informatica Developer

Datendomänen und Datendomänengruppen werden im Datendomänenglossar verwaltet. Sie können Datendomänen und Datendomänengruppen hinzufügen, bearbeiten und entfernen. Außerdem können Sie nach bestimmten Datendomänen und Datendomänengruppen suchen.

Sie können Datendomänen aus dem Datendomänenglossar in eine XML-Datei exportieren. Darüber hinaus können Sie Datendomänen aus einer XML-Datei in das Datendomänenglossar importieren. Mithilfe einer Datendomänengruppe ordnen Sie Datendomänen in bestimmten Gruppen an, wie beispielsweise Personal Health Information (PHI), Personally Identifiable Information (PII) oder jede andere konzeptionelle Gruppe, die für das Projekt relevant sind. Eine Datendomäne kann in mehreren Datendomänengruppen stehen. So kann zum Beispiel eine Telefonnummer zu den beiden Datendomänengruppen PII und PHI gehören.

Erstellen einer Datendomänengruppe in Informatica Developer

Sie können Datendomänen zu Datendomänengruppen für eine effektive Datenanalyse hinzufügen.

1. Klicken Sie auf **Fenster > Einstellungen**.
Das Dialogfeld **Einstellungen** wird eingeblendet.
2. Wählen Sie im Dialogfeld **Einstellungen** die Option **Informatica > Datendomänenglossar**.
Das Developer-Tool zeigt eine Liste aller Datendomänen im Bereich **Datendomänenglossar**.
3. Wählen Sie im Feld **Anzeigen** die Option **Datendomänengruppen**.
Im Bereich **Datendomänenglossar** wird die Liste der Datendomänen entsprechend den Datendomänengruppen angeordnet.
4. Wählen Sie im Bereich **Datendomänenglossar** die Option **Datendomänengruppen**.
5. Klicken Sie auf **Hinzufügen**.
Das Dialogfeld **Datendomänengruppe** wird angezeigt.
6. Geben Sie einen Namen und eine Beschreibung ein.
7. Klicken Sie auf **Weiter**.
8. Klicken Sie auf **Auswählen**, um das Dialogfeld **Datendomänen auswählen** zu öffnen.
9. Wählen Sie die Datendomänen aus, die Sie zur Datendomänengruppe hinzufügen möchten, und klicken Sie auf **OK**.
Das Developer Tool führt die ausgewählten Datendomänen im Bereich **Ausgewählte Datendomänen** auf.
10. Klicken Sie auf **Fertig stellen**.
Das Developer Tool fügt die Datendomänengruppe dem Datendomänenglossar hinzu.

Erstellen einer Datendomäne in Informatica Developer

Sie können Datendomänen erstellen und sie dann in das Datendomänenglossar aufnehmen. Sie können Datendomänen auch einer oder mehreren Datendomänengruppen hinzufügen.

1. Klicken Sie auf **Fenster > Einstellungen**.
Das Dialogfeld **Einstellungen** wird eingeblendet.
2. Wählen Sie im Dialogfeld **Einstellungen** die Option **Informatica > Datendomänenglossar**.
Das Developer Tool zeigt eine Liste aller Datendomänen im Bereich **Datendomänenglossar**.
3. Wählen Sie im Bereich **Datendomänenglossar** die Option **Datendomänen**.

4. Klicken Sie auf **Hinzufügen**.
Das Dialogfeld **Datendomäne** wird angezeigt.
5. Geben Sie einen Namen und eine Beschreibung ein.
6. Klicken Sie auf **Datenregel anwenden**, um Datendomänen anhand der Spaltendaten zu erkennen. Sie können auch **Spaltennamenregel anwenden** aktivieren, um Datendomänen anhand der Spaltennamen in der Datenquelle zu erkennen.
Die Schaltfläche **Durchsuchen** wird aktiviert.
7. Klicken Sie auf **Durchsuchen**, um das Dialogfeld **Speicherort auswählen** zu öffnen.
8. Wählen Sie die passende Regel aus und klicken Sie auf **OK**.
Wenn Sie ein Datendomäne erstellen, kopiert das Developer Tool Regeln und andere abhängige Objekte, die mit der Datendomäne verknüpft sind, in das Datendomänenglossar. Um eine mit der Datendomäne verknüpfte Regel zu bearbeiten, müssen Sie Änderungen an der ursprünglichen Regel vornehmen. Dann können Sie die geänderte Regel wieder mit der Datendomäne verknüpfen.
Die ausgewählten Regeln erscheinen in den Feldern **Datenregel** und **Spaltennamenregel**.
9. Klicken Sie auf **Weiter**.
10. Klicken Sie auf **Auswählen**, um das Dialogfeld **Datendomänengruppen auswählen** zu öffnen.
11. Wählen Sie die Datendomänengruppen aus, denen Sie die Datendomäne hinzufügen möchten, und klicken Sie auf **OK**.
Das Developer Tool zeigt die ausgewählten Datendomänengruppen im Bereich **Zu Datendomänengruppen zuweisen** an.
12. Klicken Sie auf **Fertig stellen**.
Das Developer Tool fügt die Datendomäne dem Datendomänenglossar hinzu.

Erstellen einer Datendomäne aus Profilergebnissen in Informatica Developer

Nachdem Sie ein Spaltenprofil ausgeführt haben, können Sie die Werte und Muster der Datenquelle anzeigen. Sie können dann eine Datendomäne daraus erstellen.

1. Führen Sie ein Spaltenprofil aus, um deren Ergebnisse anzuzeigen.
2. Wählen Sie die Werte oder Muster aus, auf deren Basis Sie eine Datendomäne erstellen möchten.
Die Werte, Muster und Statistiken erscheinen in der Ansicht **Ergebnisse**.
3. Klicken Sie mit der rechten Maustaste auf die Werte oder Muster, und wählen Sie dann **Senden an > Neue Datendomäne**.
Das Dialogfeld **Datendomäne** wird angezeigt.
4. Geben Sie den Namen der Datendomäne und eine optionale Beschreibung ein.
Der Speicherort wird standardmäßig auf Datendomänenglossar gesetzt.
5. Klicken Sie auf **Fertig stellen**.
Die Datendomäne wird zum Datendomänenglossar hinzugefügt.

Suchen von Datendomänen in Informatica Developer

Das Datendomänenglossar zeigt standardmäßig alle Datendomänen an. Sie können nach bestimmten Datendomänen und Datendomänengruppen suchen.

Im Bereich **Datendomänenglossar** im Dialogfeld **Einstellungen** werden alle Datendomänen und Datendomänengruppen angezeigt. Sie können wie folgt nach weiteren Informationen zu Datendomänen und Datendomänengruppen suchen und diese anzeigen:

Suchen nach Datendomänen und Datendomänengruppen.

Geben Sie in das Textfeld oben im Bereich **Datendomänenglossar** einen Teil des Datendomänennamens oder des Datendomänengruppennamens ein. Wenn Sie sich in der Ansicht **Datendomänengruppen** befinden, listet das Developer-Tool die Datendomänen, deren Namen die Suchzeichenfolge enthalten, und die mit ihnen verknüpften Datendomänengruppen auf. Wenn Sie sich in der Ansicht **Datendomäne** befinden, listet das Developer-Tool alle Datendomänen auf, deren Namen die Suchzeichenfolge enthalten.

Anzeigen von Datendomänengruppen und der darin enthaltenen Datendomänen.

Wählen Sie im Feld **Anzeigen** die Option **Datendomänengruppen**.

Anzeigen aller Datendomänen.

Wählen Sie im Feld **Anzeigen** die Option **Datendomänen**.

Anzeigen der Eigenschaften einer Datendomäne.

Klicken Sie auf einen Datendomänennamen, um dessen Eigenschaften unter dem Bereich **Datendomänenglossar** anzuzeigen. Sie können die Beschreibung und die verknüpften Regeln einsehen. Zur Anzeige der Domänengruppen, zu denen eine Datendomäne gehört, klicken Sie auf **In Datendomänengruppen anzeigen**.

Anzeigen der Eigenschaften einer Datendomänengruppe.

Klicken Sie auf den Namen einer Datendomänengruppe, um deren Beschreibung unter dem Bereich **Datendomänenglossar** anzuzeigen.

Importieren von Datendomänen

Sie können Datendomänen aus einer XML-Quelldatei in das Datendomänenglossar im Developer Tool importieren. Sie müssen sicherstellen, dass die Datei Informationen zu den Datendomänen enthält, die Sie importieren möchten.

1. Öffnen des Datendomänenglossars.
2. Stellen Sie sicher, dass **Datendomänen** oder **Datendomänengruppen** ausgewählt ist.
3. Klicken Sie auf **Importieren**.
Das Dialogfeld **Importieren** wird angezeigt.
4. Geben Sie im Feld **Dateiname** den Name der XML-Datei ein, aus der Sie Datendomänen importieren möchten.
Klicken Sie auf **Durchsuchen**, um die Datei auszuwählen.
5. Klicken Sie auf **Weiter**.
Der Bereich **Zu importierenden Objekte auswählen** wird angezeigt, in dem Sie Quelle und Target angeben können.
6. Wählen Sie im Bereich **Quelle** die zu importierenden Datendomänen aus.
Hinweis: Um mehrere Datendomänen auszuwählen, halten Sie die Umschalttaste gedrückt.

7. Klicken Sie auf **Auto-Match zum Target**, um die Datendomänen in den Bereich **Target** zu verschieben.
Das Developer Tool versucht, die Ableitungen der aktuellen Quellenauswahl einzeln nach Namen, Typ und übergeordneter Hierarchie abzugleichen und fügt die passenden Objekte hinzu.
8. Klicken Sie auf **Auflösung**, um anzugeben, was mit doppelten Objekten geschehen soll.
Sie können das importierte Objekt umbenennen, das vorhandene Objekt durch das importierte Objekt ersetzen, oder das vorhandene Objekt wiederverwenden. Standardmäßig benennt das Developer Tool alle doppelten Objekte um.
9. Klicken Sie auf **Weiter**.
Das Developer Tool fasst die Importeinstellungen für zusammen, sodass Sie sie überprüfen können. Im Bereich **Zusätzliche Importeinstellungen** können Sie zusätzliche Importeinstellungen angeben.
10. Klicken Sie auf **Fertig stellen**.

Exportieren von Datendomänen

Sie können Datendomänen und Datendomänenregeln aus dem Datendomänenglossar im Developer-Tool in eine XML-Datei exportieren.

1. Öffnen des Datendomänenglossars.
2. Stellen Sie sicher, dass **Datendomänen** oder **Datendomänengruppen** ausgewählt ist.
3. Klicken Sie auf **Exportieren**.
Das Dialogfeld **Export** wird eingeblendet
4. Um Datendomänen zu exportieren, wählen Sie **Datendomänen exportieren** aus. Wählen Sie **Datendomänenregeln exportieren** aus, um Datendomänenregeln zu exportieren.
5. Klicken Sie auf **Weiter**.
6. Wählen Sie im Bereich **In Datei exportieren** die Datendomänen oder Datendomänenregeln aus, die Sie exportieren möchten.
7. Um Datendomänen zu exportieren, klicken Sie auf **Durchsuchen**, um die Exportdatei und ihren Speicherort auszuwählen. Um Datendomänenregeln in ein anderes Projekt im Model Repository Service zu exportieren, wählen Sie **In Projekt kopieren** aus und wählen Sie dann das Projekt aus, in das Sie die Datendomänenregeln kopieren möchten.
8. Klicken Sie auf **Weiter**.
Im Bereich **Abhängigkeiten** wird eine Liste der abhängigen Objekte angezeigt.
9. Klicken Sie auf **Weiter**.
Der Bereich **Content-Export-Einstellungen** wird angezeigt. Sie können die zugehörigen Referenztabellen für den Export auswählen.
10. Klicken Sie auf **Fertig stellen**.
Wenn Sie eine Regel zuweisen, die Referenztabellen mit einer Datendomäne verwendet, können Sie die Referenztabellen möglicherweise nicht in derselben Developer Tool-Sitzung exportieren, in der Sie die Datendomäne erstellen. Klicken Sie im Datendomänenglossar auf **Exportieren** und trennen Sie dann die Verbindung zum Model Repository Service und stellen Sie die Verbindung wieder her, damit Sie die Regel, die Referenztabellen verwendet, importieren können.

Optionen der Datendomänenenerkennung in Informatica Developer

Sie können die Quellspalten-, Datendomänen- und Rückschlussoptionen auswählen, wenn Sie ein Profil zur Ausführung der Datendomänenenerkennung erstellen. Sie können außerdem Spalten festlegen, die bei der Datendomänenenerkennung aufgrund ihrer Datentypen und Datenlänge ausgelassen werden.

Datenqualitäts-Funktionen in Informatica Developer

In den **Datendomänenauswahl**-Optionen werden alle Datendomänen aus dem Datendomänenglossar aufgeführt. Sie können nach bestimmten Datendomänen suchen und sie auswählen, bevor sie im Rahmen der Datendomänenenerkennung ausgeführt werden.

In der folgenden Tabelle werden die **Datendomänenauswahl**-Optionen für die Datendomänenenerkennung beschrieben:

Option	Beschreibung
Aktiviert als Teil der Aktion "Profil ausführen"	Schließt die Datendomänenenerkennungsoptionen bei der Ausführung des Profils ein.
Name	Datendomänenname.
Beschreibung	Beschreibung für die Datendomäne.
Datendomänenengruppe	Name der Datendomänenengruppen, zu denen die Datendomäne gehört.
Datendomänenengruppe in der Hierarchie anzeigen	Führt alle alle Datendomänenengruppen auf, wobei die Datendomänen unter der jeweiligen Datendomänenengruppe zusammengefasst sind.

Datendomänen-Spaltenauswahl in Informatica Developer

Mit den **Spaltenauswahl**-Optionen wählen Sie die Spalten aus, die im Rahmen der Datendomänenenerkennung ausgeführt werden sollen.

In der folgenden Tabelle werden die **Spaltenauswahl**-Optionen für die Datendomänenenerkennung beschrieben:

Option	Beschreibung
Spalte	Spaltenname.
Datentyp	Datentyp der Spalte.
Genauigkeit	Maximale Genauigkeit für die Spalte.
Skalierung	Skalierung der Spalte.
Auf Null setzbar	Gibt eine Spalte an, die Nullwerte enthalten kann.
Beschreibung	Beschreibung für die Spalte.

Ableitungsoptionen für Datendomänen in Informatica Developer

Durch Rückschlussoptionen wird bestimmt, ob die Datendomänenerkennung für Spaltendaten, Spaltenname oder beide ausgeführt werden muss. Sie können angeben, ob das Profil alle Zeilen in der Datenquelle verarbeiten muss. Sie können ein Konformitätskriterium für den Datendomänenabgleich auswählen und festlegen, dass Nullen aus der Datendomänenerkennung ausgeschlossen werden.

In der folgenden Tabelle werden die **Rückschluss**-Optionen für die Datendomänenerkennung beschrieben:

Option	Beschreibung
Überschreiben der Standardableitungsoptionen	Ermöglicht die Änderung der vordefinierten Ableitungsoptionen.
Daten	Profil wird für Spaltendaten ausgeführt.
Spaltenname	Profil wird für Spaltentitel ausgeführt.
Daten und Spaltenname	Profil wird für Spaltendaten und Spaltentitel ausgeführt.
Maximale Zeilen zum Profilieren	Die maximale Anzahl an Zeilen, für die ein Profil ausgeführt werden kann. Das Developer Tool wählt die Zeilen beginnend bei der ersten Zeile in der Quelle aus.
Mindestprozentsatz der Zeilen	Der minimale Konformitätsprozentsatz der Zeilen im Datensatz, die für einen Datendomänenabgleich notwendig sind.
Mindestanzahl an Zeilen	Die Mindestanzahl der Zeilen im Datensatz, die für einen Datendomänenabgleich notwendig sind.
Nullwerte aus der Datendomänenerkennung ausschließen	Schließt die Nullwerte aus dem Datensatz für die Datendomänenerkennung aus.

Minimaler Konformitätsprozentsatz

Sie können einen Mindestprozentsatz an Zeilen im Datensatz als Konformitätskriterien für die Datendomänenerkennung auswählen.

Der Konformitätsprozentsatz ist der Quotient aus der Anzahl übereinstimmender Zeilen geteilt durch die Gesamtanzahl der Zeilen.

Hinweis: Im Developer Tool werden Nullwerte als nicht übereinstimmende Zeilen angesehen. Für Spalten mit einer großen Anzahl an Nullwerten wird unter Umständen keine Datendomänenableitung durchgeführt, es sei denn, Sie legen einen niedrigen Wert für den Mindestkonformitätsprozentsatz fest.

Beispiel

Sie haben eine Datenquelle mit 10.000 Zeilen, wobei 2.500 Zeilen in der Kommentarspalte Sozialversicherungsnummern aufweisen. Sie erstellen ein Spaltenprofil und eine Datendomänenerkennung und legen für eine Mindestanzahl an Zeilen als Konformitätskriterium 30 % fest. Wenn Sie das Profil ausführen, werden die Sozialversicherungsnummern in den Profilergebnissen nicht als abgeleitete Datendomäne angezeigt, da das minimale Konformitätskriterium bei 30 % der Zeilen oder 3.000 Zeilen in der Datenquelle liegt.

Mindestanzahl konformer Zeilen

Sie können eine Mindestanzahl an Zeilen im Datensatz als Konformitätskriterium für die Datendomänenerkennung auswählen.

Beispiel

Sie haben eine Datenquelle mit 10.000 Zeilen, wobei die Kommentarspalte E-Mail-Adressen in drei Zeilen aufweist. Sie erstellen ein Spaltenprofil und ein Datendomänenerkennungs-Profil und legen als Konformitätskriterium die Mindestanzahl der Zeilen auf 1 fest. Beim Ausführen des Profils zeigen die Profilergebnisse die E-Mail-Adresse als abgeleitete Datendomäne mit drei konformen Zeilen sowie die anderen abgeleiteten Datendomänen an.

Ausschließen von Nullwerten

Sie können Nullwerte ausschließen, wenn Sie Datendomänenerkennung für eine Datenquelle durchführen. Wenn Sie den Mindestprozentsatz der Zeilen mithilfe der Option zum Ausschließen von Nullwerten auswählen, stellt der Konformitätsprozentsatz den Quotienten aus der Anzahl übereinstimmender Zeilen geteilt durch die Gesamtanzahl der Zeilen abzüglich der Nullwerte in der Spalte dar.

Der Datendomänenerkennungsprozess verläuft anders, wenn Sie die Option **Nullwerte aus der Datendomänenerkennung ausschließen** und die zahlreichen Stichprobenoptionen oder Filter auswählen.

In den folgenden Szenarien werden die Ergebnisse der Datendomänenerkennung erläutert, wenn Sie die Option zum Ausschließen von Nullwerten mit einer Stichprobenoption und Filtern verwenden:

- Mit **Alle Zeilen** als Stichprobenoption, aber ohne Filter. Bei der Datendomänenerkennung werden alle Nullwerte in der Spalte ignoriert.
- Mit einer Stichprobenoption, aber ohne Filter. Bei der Datendomänenerkennung werden alle Nullwerte in den Stichprobendaten ignoriert und die verbleibenden Stichprobendaten berücksichtigt.
- Mit **Alle Zeilen** als Stichprobenoption und Filtern. Bei der Datendomänenerkennung werden alle Nullwerte in den gefilterten Daten ignoriert und die verbleibenden gefilterten Daten berücksichtigt.
- Mit einer Stichprobenoption und Filtern. Bei der Datendomänenerkennung werden alle Nullwerte in den gefilterten Daten der Stichprobe ignoriert und die verbleibenden gefilterten Daten berücksichtigt.

Beispiel

Sie haben eine Datenquelle mit 10.000 Zeilen, von denen 3.000 Zeilen Sozialversicherungsnummern in der Kommentarspalte aufweisen. Erstellen Sie ein Spaltenprofil und eine Datendomänenerkennung und wählen Sie folgende Optionen aus:

- Wählen Sie die Option **Nullwerte aus der Datendomänenerkennung ausschließen** aus.
- Wählen Sie **Alle Zeilen** als Stichprobenoption aus.
- Wählen Sie die Option **Mindestprozentsatz der Zeilen** aus und setzen Sie die Option auf 12 %.

Bei Ausführung des Profils wird das Profil für den Datensatz ausgeführt und die Nullwerte werden bei der Datendomänenerkennung ignoriert.

Erstellen eines Profils zur Ausführung der Datendomänenerkennung in Informatica Developer

Sie können Datendomänen in einer Datenquelle entweder als Teil eines einzelnen Datenobjektprofils oder eines Enterprise-Erkennungsprofils erkennen. Nach der Durchführung der Datendomänenerkennung können

Sie Ergebnisse überprüfen, einen Drilldown durchführen und die Ergebnisse einem Datenmodell über den Editor im Developer-Tool hinzufügen.

1. Wählen Sie in der Ansicht **Objekt-Explorer** das Projekt aus, in dem das Datenobjekt für das Profil enthalten ist.
2. Klicken Sie mit der rechten Maustaste auf das Datenobjekt und wählen Sie **Profil** aus.
Der Assistent **Neu** wird angezeigt.
3. Wählen Sie **Profil** aus.
4. Klicken Sie auf **Weiter**.
Das Developer Tool zeigt einen weiteren Bereich an, in dem Sie die allgemeinen Eigenschaften des Profils konfigurieren können.
5. Ändern Sie den Namen und die Beschreibung des Profils nach Bedarf. Sie können Datenobjekte auch hinzufügen oder entfernen.
6. Klicken Sie auf **Weiter**.
7. Wählen Sie die Spalten aus, für die die Datendomänenerkennung ausgeführt werden soll, und die Datendomänen, mit denen die Spalten abgeglichen werden sollen.
8. Ändern Sie die Standardinferenzoptionen nach Bedarf.
9. Klicken Sie auf **Fertig stellen**, um das Profil zu erstellen.

Bearbeiten eines Profils in Informatica Developer

Nachdem Sie ein Profil für die Ausführung der Datendomänenerkennung ausgeführt haben, können Sie Änderungen daran vornehmen. Sie können Spalten mit bestimmten Datentypen ausschließen und die Spaltenauswahl, die Datendomänenauswahl und die Inferenzoptionen ändern.

1. Wählen Sie in der Ansicht **Objekt-Explorer** das Projekt bzw. den Ordner aus, der das zu bearbeitende Profil enthält.
2. Doppelklicken Sie auf das Profil, um es zu öffnen.
Das Profil wird in einer Registerkarte eingeblendet.
3. Sie können die Spaltenauswahl, Datendomänenauswahl und Rückschlussoptionen nach Bedarf ändern.
4. Im Abschnitt **Spaltenauswahl** können Sie auf **Spalten ausschließen** klicken, um Ausschlussoptionen basierend auf Datentypen einzurichten.
Das Dialogfeld **Spalten ausschließen** wird angezeigt.
5. Speichern Sie die Änderungen.

Ausführen eines Profils zur Ausführung der Datendomänenenerkennung in Informatica Developer

Sie können wählen, das Profil direkt nach seiner Erstellung auszuführen. Sie können ein Profil auch manuell nach seiner Erstellung ausführen.

1. Wählen Sie in der Ansicht **Objekt-Explorer** das Projekt bzw. den Ordner aus, der das auszuführende Profil enthält.

Um ein Profil automatisch auszuführen, wählen Sie bei der Erstellung des Profils **Ausführen des Profils nach Beenden** im Assistenten **Neues Profil**.

2. Doppelklicken Sie auf das Profil, um es zu öffnen.

Das Profil wird in einer Registerkarte eingeblendet.

3. Klicken Sie mit der rechten Maustaste auf das Profil, und wählen Sie **Profil ausführen**.

Das Dialogfeld **Profil ausführen** wird mit dem Status der Profilausführung angezeigt.

Datendomänenenerkennungsergebnisse in Informatica Developer

In den Ergebnissen der Datendomänenenerkennung werden Statistiken über Spalten angezeigt, die Datendomänen entsprechen, einschließlich Konformitätskriterien für den Datendomänenabgleich. Darüber hinaus enthalten die Ergebnisse Spaltennamen, die mit Datendomänen übereinstimmen.

Zur weiteren Analyse der Ergebnisse können Sie ein Drilldown durchführen. Sie können auch die Ergebnisse in allen Zeilen der Datenquelle überprüfen und die Ergebnisse einem Datenmodell über den Editor im Developer Tool hinzufügen. Sie können die Ergebnisse auf der Basis von Datendomänen, Datendomänengruppen und Spalten sortieren. Sie können Datendomänenenerkennungsergebnisse in eine Microsoft Excel-Datei exportieren.

In der folgenden Tabelle werden die Ergebnisse der Datendomänenenerkennung beschrieben:

Spaltenname	Beschreibung
Name	Name der Datendomäne, Datendomänengruppe oder Spalte, je nachdem, ob Sie die Ansicht Datendomäne , Datendomänengruppe oder Spalten gewählt haben.
Verbindung	Name der Verbindung.
Status	Der Ableitungsstatus der Spalte
% Datenkonformität	Der minimale Konformitätsprozentsatz der Zeilen, die für einen Datendomänenabgleich notwendig sind.
Konformitätszeilenzähler	Die Mindestanzahl an Zeilen, die für einen Datendomänenabgleich notwendig sind.
% Null	Der Prozentsatz der Nullwerte für die Spalte.
Gesamtzeilen	Die Gesamtanzahl an Zeilen.

Spaltenname	Beschreibung
Spaltennamenvergleich	Gibt an, ob der Spaltenname einem Datendomänennamen entspricht.
Datendomänengruppen	Die Datendomänengruppe, zu der die Datendomäne gehört.
Dokumentierter Datentyp	Datentyp, der für die Spalte im Profilobjekt angegeben wurde.
Drilldown	Bei Auswahl dieser Option wird ein Drilldown für Zeilen ausgeführt.
Überprüft	Gibt die Überprüfung des Datendomänenvergleichs für alle Zeilen der Datenquelle an.
Letzte Ausführung	Datum und Uhrzeit der letzten Profilausführung.

Anzeigen nach Datendomänengruppen

Sie können die Datendomänenerkennungsergebnisse nach Datendomänengruppen sortiert anzeigen.

1. Führen Sie das Profil aus, um dessen Ergebnisse anzuzeigen.
2. Klicken Sie auf **Ergebnisse**.
3. Klicken Sie auf **Datendomänenerkennung**.

Sie können die Datendomänenerkennungsergebnisse im rechten Fensterausschnitt sehen.

4. Stellen Sie sicher, dass die Option **Datendomäne** im Feld **Zeigen** ausgewählt ist.
5. Wählen Sie **Datendomänengruppe in Hierarchie anzeigen**, um die Ergebnisse nach Datendomänengruppen sortiert anzuzeigen.

Anzeigen nach Spalten

Die Datendomänenerkennungsergebnisse können nach den Quellspalten sortiert angezeigt werden, die den Datendomänen entsprechen.

1. Führen Sie das Profil aus, um dessen Ergebnisse anzuzeigen.
2. Klicken Sie auf **Ergebnisse**.
3. Klicken Sie auf **Datendomänenerkennung**.

Sie können die Datendomänenerkennungsergebnisse im rechten Fensterausschnitt sehen.

4. Wählen Sie **Spalten**, um die Ergebnisse nach den Quellspalten sortiert anzuzeigen, die den Datendomänen entsprechen.

Überprüfen der Ergebnisse

Bei der Ausführung eines Profils wird eine Stichprobe der Datenquelle analysiert, um auf die Profilergebnisse zu schließen. Sie können das Profil über alle Zeilen der Quelldatenbank ausführen, um die Rückschlussergebnisse zu verifizieren.

1. Führen Sie das Profil aus, um dessen Ergebnisse anzuzeigen.
2. Klicken Sie auf **Ergebnisse**.
3. Klicken Sie auf **Datendomänenerkennung**.

Sie können die Datendomänenenerkennungsergebnisse im rechten Fensterausschnitt sehen.

4. Wählen Sie im rechten Bereich eine zu überprüfende Spalte aus.
5. Klicken Sie mit der rechten Maustaste in die Spalte und wählen Sie **Überprüfen**, um das Profil für alle Zeilen der Datenquelle auszuführen.

Unter Umständen wird nach der Überprüfung der Ergebnisse eine Änderung des Werts % **Datenkonformität** oder des Werts **Anzahl konformer Zeilen** angezeigt.

6. Zum Überprüfen der Inferenzergebnisse von mehreren Spalten wählen Sie mehrere Spalten aus. Sie können dann mit der rechten Maustaste darauf klicken und **Alle überprüfen** auswählen.

Genehmigen von Datendomänen

Wenn Sie die Datendomänenenerkennung in einem einzigen Datenobjektprofil ausführen, können Sie die abgeleiteten Datendomänen für mehrere Spalten gleichzeitig genehmigen. Wenn Sie die Datendomänenenerkennung als Teil der Enterprise-Erkennung ausführen, können Sie die Datendomäne einer Quellspalte gleichzeitig genehmigen. Um Datendomänen aus mehreren Spalten nach Enterprise-Erkennung zu genehmigen, können Sie die einzelnen Datenobjektprofilaufgaben öffnen und die Datendomänen genehmigen.

1. Wählen Sie in der Ansicht **Objekt-Explorer** ein Profil aus.
2. Doppelklicken Sie auf das Profil, um es zu öffnen.
Das Profil wird auf einer Registerkarte geöffnet.
3. Wenn Sie ein einzelnes Datenobjektprofil ausgeführt haben, wählen Sie die Ansicht **Datendomänenenerkennung** und markieren Sie anschließend eine Zeile. Die Zeile enthält die Datendomänenenerkennungsergebnisse für jede Spalte.
4. Klicken Sie mit der rechten Maustaste in die Zeile und wählen Sie **Akzeptieren**.
Die Inferenzstatus der Datendomäne ändert sich in **Akzeptiert**.
5. Wenn Sie die Enterprise-Erkennung ausgeführt haben, wählen Sie die Ansicht **Datendomänen** und markieren Sie anschließend eine Datendomäne.
Die Spalten, die mit der Datendomäne übereinstimmen, erscheinen im rechten Bereich.
6. Klicken Sie mit der rechten Maustaste auf die zu genehmigende Spalte und wählen Sie **Akzeptieren** aus. Sie können auch mehrere abgelehnte Spalten auswählen und gegebenenfalls genehmigen.
Die Inferenzstatus der Datendomäne ändert sich in **Akzeptiert**.
7. Zum Wiederherstellen des abgeleiteten Status der Datendomäne klicken Sie mit der rechten Maustaste auf die Zeile und dann auf **Zurücksetzen**.

Ablehnen von Datendomänen

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

1. Wählen Sie in der Ansicht **Objekt-Explorer** ein Profil aus.
2. Doppelklicken Sie auf das Profil, um es zu öffnen.
Das Profil wird auf einer Registerkarte geöffnet.
3. Wählen Sie in der Ansicht **Datendomänenenerkennung** oder **Datendomänen** eine Zeile aus.

4. Wenn Sie abgeleitete Domains ablehnen wollen, klicken Sie mit der rechten Maustaste auf die Reihe und wählen Sie **Ablehnen** aus.
Informatica Developer blendet die abgelehnte Datendomäne in den Ergebnissen der Datendomänenerkennung aus.
5. Um abgelehnte Datendomänen auszublenden, klicken Sie mit der rechten Maustaste auf die Zeile und wählen **Abgelehnte ausblenden**.
6. Klicken Sie zur Anzeige der abgelehnten Datendomänen mit der rechten Maustaste auf eine der Zeilen und wählen Sie **Abgelehnte einblenden** aus.

Exportieren von Datendomänenerkennungsergebnissen aus Informatica Developer

Beim Exportieren der Ergebnisse der Datendomänenerkennung aus Informatica Developer in eine `.xlsx`-Datei können Sie die Datei entweder auf dem Server oder in einem bestimmten Speicherort auf dem Clientcomputer speichern.

1. Führen Sie ein Profil zur Durchführung der Datendomänenerkennung aus.
2. Klicken Sie auf die Ansicht **Ergebnisse**.
3. Klicken Sie auf das Symbol **Ergebnisse in Datei exportieren**.
Das Dialogfenster **Daten in Datei exportieren** wird geöffnet.
4. Geben Sie den Dateinamen ein. Optional können Sie den Standarddateinamen verwenden.
5. Wählen Sie unter **Speichern** die Option **Auf Client speichern** und klicken Sie auf **Durchsuchen**, um einen Speicherort zu wählen. Speichern Sie anschließend die Datei lokal auf Ihrem Computer. Standardmäßig schreibt Informatica Developer die Datei an einen in den Data Integration Service-Eigenschaften von Informatica Administrator festgelegten Serverspeicherort.
6. Klicken Sie auf **OK**.

KAPITEL 25

Enterprise-Erkennung in Informatica Developer

Dieses Kapitel umfasst die folgenden Themen:

- [Enterprise-Erkennung in Informatica Developer - Übersicht, 197](#)
- [Enterprise-Erkennungsprozess, 198](#)
- [Profilooptionen für Enterprise-Erkennung, 199](#)
- [Erstellen eines Enterprise-Erkennungsprofils in Informatica Developer, 204](#)
- [Bearbeiten eines Profils, 205](#)
- [Ausführen eines Enterprise-Erkennungsprofils, 206](#)
- [Fremdschlüsselerkennung, 207](#)
- [Join-Analyse, 208](#)
- [Überschneidungserkennung, 210](#)
- [DDL-Skriptdateien, 212](#)
- [Synchronisieren eines Enterprise-Erkennungsprofils, 213](#)

Enterprise-Erkennung in Informatica Developer - Übersicht

Enterprise-Erkennung ist der Prozess, der Spaltenprofilstatistiken, Datendomänen, Primärschlüssel und Fremdschlüssel in einer großen Anzahl von Datenquellen erkennt. Sie können die Enterprise-Erkennung über mehrere Verbindungen oder Schemata ausführen.

Als Datenanalytiker in einem Unternehmen möchten Sie möglicherweise wichtige Datenmerkmale über eine große Anzahl von Datenquellen erkennen. Zu den möglichen Anforderungen können Identifizierung von relationalen Datenbeständen, Ausführungen von Spaltenprofilen über die erkannten Datenbestände, Erkennung kritischer Datenmerkmale innerhalb des Unternehmens, Primärschlüssel und Schlüsselkandidaten gehören. Möglicherweise möchten Sie auch die Fremdschlüsselbeziehungen einsehen, die bei den Datenquellen vorliegen, damit Sie daraus ein Datenmodell auf der Basis der erkannten Beziehungen ableiten können.

Mit der Enterprise-Erkennung können Probleme, Muster, Trends und wichtigen Datenmerkmale in Datenbeständen in Ihrem Unternehmen erkannt werden. Sie können sowohl Datenquellen, die Sie in das Model Repository importieren möchten, als auch Datenquellen von externen relationalen Verbindungen

wählen. Der Datenerkennungsprozess beinhaltet die Erkennung der Spaltenprofilstatistik, Datendomänenanalyse, Datenobjektstrukturen einschließlich Schlüsselkandidaten und Beziehungen von Datenobjekten, die Fremdschlüssel enthalten. Sie führen die Enterprise-Erkennung im Developer Tool aus, wobei die folgenden Tasks bei den einzelnen Datenquellen ausgeführt werden:

- Ausführen eines Spaltenprofils.
- Erkennen von Datendomänen.
- Rückschluss der Primärschlüssel.

Nach der Ausführung von Spaltenprofilen, Datendomänenerkennung und Primärschlüsselprofilen führt das Developer Tool ein Fremdschlüsselprofil über alle Datenquellen aus. Wenn das Developer Tool die Profiling- und Erkennungs-Tasks abgeschlossen hat, wird eine konsolidierte Zusammenfassung der Ergebnisse sowohl in grafischer als auch tabellarischer Form erzeugt.

Sie können ein Betriebssystemprofil in Informatica Developer auswählen. Nach der Auswahl eines Betriebssystemprofils erstellt der Datenintegrationsdienst die Enterprise-Erkennungsprofile und führt sie basierend auf den Berechtigungen des Betriebssystembenutzers aus, den Sie im Betriebssystemprofil definiert haben.

Enterprise-Erkennungsprozess

Sie können das Enterprise-Erkennungsprofil ausführen, um eine Enterprise-Erkennung im Developer-Tool durchzuführen. Bevor Sie das Profil ausführen können, müssen Sie die Datenerkennungsoptionen für verschiedene Profiltypen konfigurieren.

Das Developer-Tool erstellt Datenobjekte für die ausgewählten Datenquellen und Profilaufgaben für jedes Datenobjekt. Das Tool führt dann Profilaufgaben zum Generieren von Profilergebnissen aus.

Führen Sie die folgenden Schritte aus, um eine Enterprise-Erkennung durchzuführen:

1. Erstellen Sie ein Enterprise-Erkennungsprofil, indem Sie mehrere in das Modellrepository importierte Datenobjekte sowie Datenquellen über mehrere externe relationale Verbindungen auswählen.
2. Definieren Sie die Konfigurationseinstellungen für Datendomänenerkennung, Spaltenprofil, Primärschlüssel- und Fremdschlüsselprofil.
3. Führen Sie das Enterprise-Erkennungsprofil aus.
4. Aktualisieren Sie den Modellrepository-Dienst.

Hinweis: Sie müssen diese Aktion durchführen, da der Import von Metadaten für externe Verbindungen im Modellrepository erfolgt. Sie müssen den Modellrepository-Dienst aktualisieren, sodass das Developer-Tool die Änderungen im Modellrepository anzeigt.

5. Überwachen Sie die Profilausführung und zeigen Sie, falls erforderlich, die Status der Profilaufgaben an, die das Developer-Tool ausführt.
6. Überprüfen Sie die Zusammenfassung der Enterprise-Erkennungsergebnisse. Die Zusammenfassung enthält eine interaktive grafische Benutzeroberfläche und eine Tabellenansicht.

Profilooptionen für Enterprise-Erkennung

Richten Sie die Profilooptionen ein, bevor Sie ein Profil zur Ausführung der Enterprise-Erkennung ausführen. Zu den Profilooptionen gehören Datendomänenenerkennungsoptionen, Spaltenprofil-Stichprobenoptionen und Rückschlussoptionen für Primär- und Fremdschlüssel.

Sie können wählen, das Enterprise-Erkennungsprofil nach der Einrichtung der Profilooptionen auszuführen. Außerdem können Sie wählen, nach der Einrichtung Profil-Tasks zu erstellen, ohne das Profil auszuführen.

Datendomänenauswahl für die Enterprise-Erkennung

Durch Rückschlussoptionen wird bestimmt, ob die Datendomänenenerkennung für Spaltendaten, Spaltenname oder beide ausgeführt werden muss. Sie können angeben, ob das Profil alle Zeilen in der Datenquelle verarbeiten muss. Darüber hinaus können Sie ein Konformitätskriterium für den Datendomänenabgleich auswählen.

In der folgenden Tabelle werden die Datendomänen-Rückschlussoptionen beschrieben, die Sie für die Enterprise-Erkennung konfigurieren können:

Option	Beschreibung
Überschreiben der Standardableitungsoptionen	Ändert die vordefinierten Ableitungsoptionen.
Daten	Das Profil wird für Spaltendaten ausgeführt.
Spaltenname	Das Profil wird für Spaltentitel ausgeführt.
Daten und Spaltenname	Das Profil wird für Spaltendaten und Spaltentitel ausgeführt.
Mindestprozentsatz der Zeilen	Der minimale Konformitätsprozentsatz der Zeilen im Datensatz, die für einen Datendomänenabgleich notwendig sind. Der Konformitätsprozentsatz ist der Quotient aus der Anzahl übereinstimmender Zeilen geteilt durch die Gesamtanzahl der Zeilen. Hinweis: Im Developer Tool werden Nullwerte als nicht übereinstimmende Zeilen angesehen.
Mindestanzahl an Zeilen	Die Mindestanzahl der Zeilen im Datensatz, die für einen Datendomänenabgleich notwendig sind.
Nullwerte aus der Datendomänenenerkennung ausschließen	Schließt die Nullwerte aus dem Datensatz für die Datendomänenenerkennung aus.
Alle Zeilen	Das Profil wird für alle Zeilen der Datenquelle ausgeführt.
Erste Stichprobe	Die maximale Anzahl an Zeilen, für die ein Profil ausgeführt werden kann. Das Developer Tool wählt die Zeilen beginnend bei der ersten Zeile in der Quelle aus.
Genehmigte Datentypen und Datendomänen in den nachfolgenden Profilausführungen aus der Datentyp- und Datendomänenableitung ausschließen	Schließt den genehmigten Datentyp oder die genehmigte Datendomäne bei der nächsten Profilausführung aus der Datentyp- und Datendomänenableitung aus.

Stichprobenoptionen des Spaltenprofils für die Enterprise-Erkennung

Durch die Stichprobenoptionen wird festgelegt, ob das Developer Tool ein Spaltenprofil über allen Zeilen der Datenquellen oder über eine begrenzte Anzahl von Zeilen ausführt.

In der folgenden Tabelle werden die die Spaltenprofil-Stichprobenoptionen beschrieben, die Sie für die Enterprise-Erkennung konfigurieren können:

Option	Beschreibung
Alle Zeilen	Führt ein Profil für alle Zeilen im Datenobjekt aus. Unterstützt in der nativen Laufzeitumgebung und in der Blaze- und Spark-Laufzeitumgebung.
Stichprobe erste <number> Zeilen	Führt ein Profil auf Stichprobenzeilen aus, die sich am Anfang der Zeilen des Datenobjekts befinden. Sie können maximal 2.147.483.647 Zeilen auswählen. Unterstützt in der nativen Laufzeitumgebung und in der Blaze-Laufzeitumgebung.
Limit n <number> Zeilen	Führt ein Profil basierend auf der Anzahl der Zeilen im Datenobjekt aus. Wenn Sie ein Profil in der Hadoop-Validierungsumgebung ausführen, sammelt die Spark-Engine Stichproben von mehreren Partitionen des Datenobjekts und überträgt die Stichproben in einen einzigen Knoten, um den Stichprobenumfang zu berechnen. Die Stichprobenoption „Limit n“ unterstützt nur Oracle-, SQL Server- und DB2-Datenbanken. Sie können den erweiterten Filter nicht zusammen mit der Stichprobenoption „Limit n“ anwenden. Unterstützt in der Spark-Laufzeitumgebung.
Zufallsprozentsatz	Führt ein Profil auf einem Prozentsatz von Zeilen im Datenobjekt aus. Unterstützt in der Spark-Laufzeitumgebung.
Datentypableitung für Spalten mit genehmigtem Datentyp ausschließen	Schließt Spalten mit einem genehmigten Datentyp aus der Datentypableitung der Spaltenprofilausführung aus.

Laufzeitumgebung Option

Wählen Sie die Option „Nativ“ oder „Hadoop“ als Laufzeitumgebung aus. Sie können in der Hadoop-Laufzeitumgebung die Option „Blaze“ oder „Spark“ wählen. Informatica Developer legt die Laufzeitumgebung in der Profildefinition fest, nachdem Sie die Laufzeitumgebung ausgewählt haben. Die Laufzeitumgebung hat keine Auswirkungen auf die Profilergebnisse.

In der folgenden Tabelle werden die Laufzeitumgebungsoptionen für ein Enterprise-Erkennungsprofil beschrieben:

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

Ableitungsoptionen der Primärschlüssel für die Enterprise-Erkennung

Sie können die Standard-Primärschlüssel-Rückschlussoptionen für die Enterprise-Erkennung überschreiben. Zu den Optionen zählen die maximale Anzahl der Zeilen, für die das Profil ausgeführt werden kann, sowie der minimale Konformitätsprozentsatz.

In der folgenden Tabelle werden die die Primärschlüssel-Rückschlussoptionen beschrieben, die Sie für die Enterprise-Erkennung konfigurieren können:

Optionen	Beschreibung
Überschreiben der Standardableitungsoptionen	Ermöglicht das Konfigurieren der benutzerdefinierten Primärschlüsselableitung.
Max. Schlüsselspalten	Maximale Anzahl an Spalten, die einen Primärschlüssel bilden können.
Max. Zeilen	Maximale Anzahl an Zeilen, für die das Profil ausgeführt werden kann.
Minimaler Prozentsatz	Der minimale Konformitätsprozentsatz der Spaltendaten, die für einen Primärschlüsselvergleich erforderlich sind.
Zeilen mit maximaler Verletzung	Die maximale Anzahl an Zeilen mit Schlüsselverstößen, die das Profil bei der Ermittlung der Primärschlüssel zulässt.

Ableitungsoptionen der Fremdschlüssel für die Enterprise-Erkennung

Richten Sie die Fremdschlüssel-Rückschlussoptionen ein, um die Spalteneinstellungen für die Erkennung von Fremdschlüsselbeziehungen zwischen Datenobjekten zu definieren. Fremdschlüssel-Rückschlussergebnisse sind abhängig von den Primärschlüssel-Rückschlussoptionen, die Sie für die Enterprise-Erkennung festlegen, von dokumentierten Primärschlüsseln und benutzerdefinierten Primärschlüsseln.

Sie können die Fremdschlüssel in Informatica Developer mithilfe einer der folgenden Methoden ableiten:

- Verwenden von Standardwerten.
- Konfigurieren der Ableitungsoptionen für Fremdschlüssel.
- Verwenden der Konfigurationsdatei des Fremdschlüssels zum Konfigurieren der automatischen Wiederherstellungsparameter.

In der folgenden Tabelle werden die Fremdschlüssel-Rückschlussoptionen beschrieben, die Sie für die Enterprise-Erkennung konfigurieren können:

Optionen	Beschreibung
Überschreiben der Standardableitungsoptionen	Ändert die vordefinierten Ableitungsoptionen.
In Vergleichen verwendete Datentypen	Der in Haupt- und Fremdschlüsselvergleichen verwendete Datentyp. Hinweis: Diese Option gilt, wenn Sie vor dem Fremdschlüsselrückschluss ein Spaltenprofil auf die Datenquelle ausführen.
Vergleich Groß-/Kleinschreibung	Unterscheidet zwischen Groß- und Kleinschreibung beim Vergleichen von Spaltendaten.

Optionen	Beschreibung
Werte vor dem Vergleich abgleichen	Gibt an, ob das Developer Tool bei der Verarbeitung führende oder nachfolgende Leerzeichen in Spaltendaten berücksichtigt.
In Vergleichen verwendete abgeleitete Primärschlüssel Höchstplatzierte Schlüssel verwenden	Die Anzahl der höchstplatzierten Primärschlüssel, die bei der Ableitung von Fremdschlüsseln verwendet werden, wenn das Developer Tool ein Fremdschlüsselprofil über alle Datenquellen hinweg ausführt. Das Developer Tool verwendet die Höchstplatzierten-Methode zusammen mit dokumentierten und benutzerdefinierten Primärschlüsseln zum Ableiten der Fremdschlüsselbeziehungen. Höchstplatzierungen von abgeleiteten Schlüsseln basieren auf dem absteigenden Konformitätsprozentsatz, gerundet auf eine einzige Dezimalstelle. Das Developer Tool sieht beispielsweise einen Konformitätsprozentsatz von 99,75 als 99,8 und 99,74 als 99,7 an. Der Standardwert ist 1. Setzen Sie den Wert auf -1, wenn das Developer Tool alle abgeleiteten Schlüssel bei der Fremdschlüsselableitung verwenden soll. Hinweis: Wenn Primärschlüssel von den Datenquellen des Primärschlüssels genehmigt wurden, verwendet das Developer Tool keine abgeleiteten Primärschlüssel für die Ableitung von Fremdschlüsseln.
Max. Fremdschlüssel zwischen Datenobjekten	Die maximale Anzahl abgeleiteter Spalten, die vom Developer Tool nach der Profilausführung zurückgegeben werden und die für die Fremdschlüsselerkennung notwendig sind.
Minimaler Konformitätsprozentsatz	Der Mindestprozentsatz, ab dem Spalten in die Fremdschlüsselergebnisse aufgenommen werden.
Signatur neu generieren	Lädt Spaltensignaturen neu, wenn sich die Quelldaten ändern.

Automatische Wiederherstellungsparmeter für die Fremdschlüsselableitung

Sie können die automatischen Wiederherstellungsparmeter konfigurieren, um Primär- und Fremdschlüsselbeziehungen ohne manuelles Eingreifen abzurufen. Die automatischen Wiederherstellungsparmeter sind benutzerdefinierte Attribute, die zur Angabe der Datenbeziehungen auf Basis bestimmter Bedingungen konfiguriert werden können.

Wenn die Erkennungsergebnisse eine große Anzahl an Primär- und Fremdschlüsselbeziehungen enthalten, gestaltet es sich unter Umständen schwierig, die entscheidenden Datenbeziehungen unter Hunderten von Datenbeziehungen zu erkennen. Darüber hinaus ist es unter Umständen schwierig, die Beziehungen basierend auf bestimmten Bedingungen, wie z. B. Datenübereinstimmung oder Datentyp, wiederherzustellen. Zur Behebung dieses Problems können Sie die automatischen Wiederherstellungsparmeter konfigurieren und das Enterprise-Erkennungsprofil ausführen.

Wenn die Datenquellen mehrere Fremdschlüsselkandidaten aufweisen und Regeln zum Auswählen eines Fremdschlüsselkandidaten bereitgestellt werden sollen, können Sie folgende Aktionen ausführen:

- Konfigurieren Sie die Optionen **Max. Anzahl an Fremdschlüsseln zwischen Datenobjekten** und **Minimaler Konformitätsprozentsatz** im Enterprise-Erkennungsprofil-Assistenten.
- Konfigurieren Sie die Gewichtungen und Scores für die automatischen Wiederherstellungsparmeter in der Datei `ForeignKeyConfig.xml`.

Ein Administrator kann die Konfigurationsdatei des Fremdschlüssels bearbeiten und speichern. Konfigurieren Sie die automatischen Wiederherstellungsparmeter in der Konfigurationsdatei des Fremdschlüssels. Der

Algorithmus ruft die Primär- und Fremdschlüsselbeziehungen zwischen mehreren Datenobjekten basierend auf den automatischen Wiederherstellungsparametern ab.

Die Konfigurationsdatei des Fremdschlüssels `ForeignKeyConfig.xml` steht in folgendem Verzeichnis zur Verfügung:

```
<Informatica installation directory>\services\DataIntegrationService\modules\ProfilingService
```

Bei den automatischen Wiederherstellungsparametern handelt es sich um Datenüberschneidungs-, Spaltennamen-, Beziehungstyp- und Datentypvergleiche.

Datenüberschneidungsvergleich

Bei einem Datenüberschneidungsvergleich handelt es sich um die geschätzte Überschneidung von Werten zwischen den Primär- und Fremdschlüsseln. Sie können den Überschneidungsvergleich im Enterprise-Erkennungsprofil-Assistenten mit der Option **Minimaler Konformitätsprozentsatz** festlegen. Standardmäßig ist die Option **Minimaler Konformitätsprozentsatz** auf 90 festgelegt.

Wenn der Datenüberschneidungsvergleich nicht dem minimalen Konformitätsprozentsatz entspricht, wird der Fremdschlüssel bei der automatischen Wiederherstellung nicht berücksichtigt. Wenn minimale Konformität für den Datenüberschneidungsvergleich erfüllt ist, werden die verbleibenden Parameter zum Berechnen des angepassten Scores verwendet.

Namensvergleich

Der Namensvergleichsparameter ist ein optionaler Parameter. Er verwendet den Algorithmus „Abstand bearbeiten“, um festzulegen, wie genau die Namen von Primär- und Fremdschlüsseln übereinstimmen, und legt den Score zwischen 0 und 1 fest. Setzen Sie die Gewichtung des Namensvergleichs auf 0, wenn dieser Parameter nicht zum Festlegen der Beziehung zwischen Primär- und Fremdschlüssel verwendet werden soll.

Beziehungstypvergleich

Beim Beziehungstypvergleich wird der Typ der Beziehung zwischen Primär- und Fremdschlüsselspalten festgelegt und ein fester Score zwischen 0 und 1 zugewiesen. Der Beziehungstypvergleich wird basierend auf dem Spaltentyp der Fremdschlüsselspalte berechnet.

Die folgenden Beziehungstypvergleiche können in der Datei `ForeignKeyConfig.xml` eingerichtet werden:

- Beziehung zwischen einem Fremd- und einem Primärschlüssel, wobei die Fremdschlüsselspalte eine Nichtschlüsselspalte darstellt. Der Standardwert für diesen Beziehungstypvergleich lautet 1. Sie finden diese Beziehung in vielen Datenquellen.
- Beziehung zwischen zwei Primärschlüsseln, wobei die Fremdschlüsselspalte eine Primärschlüsselspalte darstellt. Der Standardwert für diesen Beziehungstyp lautet 0,25. Sie finden diesen Beziehungstyp selten, da er eine Tabelle darstellt, die vertikal partitioniert wurde.
- Beziehung zwischen einem Primärschlüssel und einer Primärschlüsselabfolge, wobei die Fremdschlüsselspalte eine Primärschlüsselspalte darstellt und als Datentyp der Spalte ein Sequenzdatentyp verwendet wird. Die Spalte `OrderID` in einer Auftrags-tabelle weist beispielsweise einen Sequenzdatentyp auf. Der Standardwert für diesen Beziehungstyp lautet null, da die Sequenzschlüssel unter Umständen zahlreiche falsch positive Fremdschlüssel erzeugen, die der Primärschlüssel-Primärschlüssel-Algorithmus zu vermeiden versucht. Sie können den Beziehungstypvergleich auf einen höheren Score festlegen, wenn bekannt ist, dass die Datenquelle einige Sequenzdatentypen enthält.

Datentypvergleich

Beim Datentypvergleich werden die Datentypen der Primärschlüsselspalten mit den Datentypen der Fremdschlüsselspalten verglichen. Darüber hinaus wird ein fester Konformitäts-Score zugewiesen, der darauf basiert, wie genau die Datentypen der Spalten übereinstimmen.

In der folgenden Tabelle werden die festen Datentypvergleichs-Scores für verschiedene Kombinationen aus Primär- und Fremdschlüsseln aufgelistet:

	Numerischer Fremdschlüssel	Datumsfremdschlüssel	Zeichenfolgenfremdschlüssel
Numerischer Primärschlüssel	1.0	0.5	0.0
Datumsprimärschlüssel	0.5	1.0	0.5
Zeichenfolgenprimärschlüssel	0.0	0.0	1.0

Sie können die standardmäßigen Datentypvergleichs-Scores gegebenenfalls ändern.

Erstellen eines Enterprise-Erkennungsprofils in Informatica Developer

Sie können ein Profil über mehreren Datenquellen unter mehreren Verbindungen erstellen. Das Developer Tool erstellt für jede Quelle individuelle Profil-Tasks.

1. Wählen Sie in der Ansicht **Objekt-Explorer** mehrere Datenobjekte aus, für die Sie ein Profil ausführen möchten.
2. Klicken Sie zum Öffnen des Profilassistenten auf **Datei > Neu > Profil**.
3. Wählen Sie **Enterprise-Erkennungsprofil** aus und klicken Sie auf **Weiter**.
4. Geben Sie einen Namen für das Profil ein, und überprüfen Sie den Projektspeicherort. Suchen Sie bei Bedarf nach einem neuen Speicherort.
5. Stellen Sie sicher, dass der Name des von Ihnen ausgewählten Datenobjekts im Abschnitt **Datenobjekte** angezeigt wird. Klicken Sie auf **Auswählen**, um bei Bedarf mehrere Datenobjekte auszuwählen.
6. Klicken Sie auf **Weiter**.

Der Bereich **Ressourcen zu Profildefinition hinzufügen** wird angezeigt. In diesem Bereich können Sie mehrere externe relationale Datenquellen und Verbindungen auswählen.

7. Klicken Sie auf **Auswählen**, um das Dialogfeld **Ressourcen auswählen** zu öffnen.

Im Bereich **Ressourcen** werden alle internen und externen Verbindungen und Datenobjekte unter der Informatica-Domäne aufgeführt.

8. Klicken Sie auf **OK**, um das Dialogfeld zu schließen.
9. Klicken Sie auf **Weiter**.
10. Konfigurieren Sie die Profiltypen, die Sie ausführen möchten. Die folgenden Profiltypen können konfiguriert werden:
 - Datendomänenerkennung

- Spaltenprofil
- Primärschlüssel-Profiling
- Fremdschlüssel-Profil

Hinweis: Wählen Sie **Aktiviert als Teil der Aktion „Enterprise-Erkennungsprofil ausführen“** für die Profiltypen aus, die Sie als Teil des Enterprise-Erkennungsprofils ausführen möchten. Standardmäßig ist Spalten-Profiling aktiviert.

11. Überprüfen Sie die Optionen für das Profil.

Sie können die Stichprobenoptionen für Spaltenprofile bearbeiten. Sie können auch die Inferenzoptionen für Datendomänen-, Primärschlüssel- und Fremdschlüsselprofile bearbeiten.

12. Wählen Sie **Profile erstellen** aus.

Das Developer-Tool erstellt Profile für jede einzelne Datenquelle.

13. Wählen Sie **Enterprise-Erkennungsprofil beim Beenden ausführen** aus, um das Profil auszuführen, wenn Sie die Profilkonfiguration abschließen. Wenn Sie alle Profiling-Operationen aktiviert haben, führt das Developer-Tool Spalten-, Datendomänen- und Primärschlüsselprofile für alle ausgewählten Datenquellen aus. Anschließend führt das Developer-Tool ein Fremdschlüsselprofil über alle Datenquellen aus.

14. Klicken Sie auf **Fertig stellen**.

Nach dem Ausführen eines Enterprise-Erkennungsprofils müssen Sie den Modellrepository-Dienst aktualisieren, bevor Sie die Ergebnisse anzeigen können. Dieser Schritt ist erforderlich, da der Import von Metadaten für externe Verbindungen im Modellrepository erfolgt. Sie müssen den Modellrepository-Dienst aktualisieren, sodass das Developer-Tool die Änderungen im Modellrepository anzeigt.

Bearbeiten eines Profils

Nachdem Sie ein Enterprise-Erkennungsprofil eingerichtet haben, können Sie Änderungen daran vornehmen. Sie können Spalten mit bestimmten Datentypen ausschließen und die Spaltenauswahl, die Datendomänenauswahl und die Inferenzoptionen ändern.

1. Wählen Sie in der Ansicht **Objekt-Explorer** das Projekt bzw. den Ordner aus, in dem das zu bearbeitende Profil enthalten ist.

2. Klicken Sie auf **Team > Aussehen**, um das Profil auszuwählen.

3. Doppelklicken Sie auf das Profil, um es zu öffnen.

4. Klicken Sie auf die Ansicht **Eigenschaften**.

Die Eigenschaftsansicht befindet sich unter der Standardansicht.

5. Klicken Sie auf **Profile**, um die Profilaufgaben anzuzeigen.

6. Wählen Sie im rechten Bereich eine Profilaufgabe aus, die Sie bearbeiten möchten, und klicken Sie auf **Öffnen**.

Das Profil wird in einer Registerkarte eingeblendet.

7. Um Änderungen an den globalen Einstellungen des Enterprise-Erkennungsprofils vorzunehmen, wählen Sie das Profil oben in der **Liste der Profiling-Aufgaben** aus und klicken Sie auf **Konfigurieren**.

8. Nehmen Sie die erforderlichen Änderungen an den Profildefinitionsoptionen vor.

9. Speichern Sie die Änderungen.

10. Klicken Sie auf **Team > Einchecken**, um das Profil einzuchecken.

Ausführen eines Enterprise-Erkennungsprofils

Ein Enterprise-Erkennungsprofil kann auf verschiedene Weise ausgeführt werden. Sie können das Profil aus der Ansicht **Objekt-Explorer** heraus oder von der Registerkarte **Profile** im Fenster **Eigenschaften** ausführen. Sie können wahlweise einzelne und mehrere Profilaufgaben ausführen, die einen Teil des Enterprise-Erkennungsprofils bilden.

1. Wählen Sie in der Ansicht **Objekt-Explorer** das Projekt bzw. den Ordner mit dem enthaltenen Profil aus, das Sie ausführen möchten.

Um ein Profil automatisch auszuführen, wählen Sie beim Erstellen des Profils **Enterprise-Erkennungsprofil beim Beenden ausführen** im Assistenten **Neue Enterprise-Erkennung** aus.

2. Doppelklicken Sie auf das Profil, um es zu öffnen.

Das Profil wird in einer Registerkarte geöffnet.

3. Klicken Sie in der Ansicht **Objekt-Explorer** mit der rechten Maustaste auf das Profil und wählen Sie **Enterprise-Erkennungsprofil ausführen** aus.

Sie können auch **Profile** im Fenster **Eigenschaften**, dann den Namen des Profils in der **Liste der Profiling-Tasks** auswählen und auf **Ausführen** klicken.

Hinweis: Beim Ausführen eines Enterprise-Erkennungsprofils müssen Sie den Modellrepository-Dienst aktualisieren, bevor Sie die Ergebnisse anzeigen können. Dieser Schritt ist notwendig, da der Import von Metadaten für externe Verbindungen im Modellrepository erfolgt. Sie müssen den Modellrepository-Dienst aktualisieren, damit das Developer-Tool die Änderungen im Modellrepository entsprechend anzeigen kann.

4. Das Dialogfeld **Ausführen** wird angezeigt. In diesem Dialogfeld können Sie Änderungen an den globalen Einstellungen des Profils vornehmen.

Standardmäßig gelten die vorgenommenen Änderungen für die im Enterprise-Erkennungsprofil neu hinzugefügten Datenobjekte.

5. Um die Änderungen auf alle Profilaufgaben des Datenobjekts sowie auf die Profilaufgabe des vom System erzeugten Fremdschlüssels im Erkennungsprofil des Unternehmens anzuwenden, wählen Sie **Globale Einstellungen für aktuelle Profile verwenden** aus.

Das Developer-Tool aktualisiert alle Datenobjektprofil-Tasks und Fremdschlüsselprofil-Tasks entsprechend den veränderten Einstellungen.

6. Wählen Sie zum Ausführen bestimmter Profilaufgaben eine Aufgabe aus und klicken Sie auf **Ausführen**.

7. Klicken Sie zum Ausführen mehrerer Profilaufgaben auf **Mehrere ausführen**.

Das Dialogfeld **Mehrere ausführen** wird angezeigt.

Tipp: Wenn es zu lange dauert, bis die Ergebnisse der Enterprise-Erkennung geladen sind, sollten Sie die Statistiken der Datenbank des Profiling Warehouse aktualisieren. Eine mehrfache Ausführung des Enterprise-Erkennungsprofils führt möglicherweise zu deutlichen Veränderungen bei Datenmengen und Spaltenwerten. Nachdem Sie die Statistiken aktualisiert haben, kann die Datenbank basierend auf den aktuellsten Statistiken einen Ausführungsplan für die SQL-Anfragen abarbeiten und die Datenbankoperationen optimieren.

8. Gemäß Voreinstellung werden alle Aufgaben ausgewählt. Deaktivieren Sie die Aufgaben, die Sie nicht ausführen möchten, und klicken Sie auf **OK**.

Fremdschlüsselerkennung

Eine Spalte ist ein Fremdschlüssel, wenn ihre Datenwerte mit den Primärschlüssel-Spaltenwerten in einem anderen Datenobjekt übereinstimmen.

Mit dem Developer-Tool können Sie Fremdschlüsselanalysen für Mehrfachdatenobjekte durchführen. Erstellen Sie ein Enterprise-Erkennungsprofil, um Datenobjekte auszuwählen und das Profil zu definieren.

Bevor Sie die Fremdschlüssel-Erkennung durchführen, müssen Sie die übergeordneten und untergeordneten Datenobjekte im Enterprise-Erkennungsprofil bestimmen. Das Profil nutzt einen oder mehrere Schlüssel im übergeordneten Objekt einschließlich dessen Primärschlüssel, um Fremdschlüssel im Kind-Objekt zu erkennen. Nachdem Sie die übergeordneten und Kind-Objekte definiert und die Schlüssel in dem übergeordneten Objekt identifiziert haben, erstellen Sie das Profil und führen es aus.

Definieren übergeordneter und Kind-Objektbeziehungen

Um Fremdschlüsselbeziehungen zwischen zwei Datenobjekten zu suchen, müssen Sie ein übergeordnetes Datenobjekt auswählen und den Primärschlüssel in diesem Objekt angeben.

1. Öffnen Sie ein Enterprise-Erkennungsprofil, das die Datenobjekte enthält, die Sie analysieren möchten.
2. Wählen Sie das übergeordnete Objekt aus.
3. Wählen Sie den Primärschlüssel im übergeordneten Objekt aus:
 - Klicken Sie auf die Registerkarte **Eigenschaften** und auf **Schlüssel**.
 - Klicken Sie auf **Hinzufügen**, und wählen Sie die Primärschlüsselspalte im Dialogfenster Neuer Schlüssel.
 - Klicken Sie auf **OK** im Dialogfenster **Neuer Schlüssel**. Vergewissern Sie sich, dass der Primärschlüssel im Abschnitt **Ausgewählte Felder** steht und dass die Option **Primärschlüssel** aktiviert ist.

Erstellen Sie ein Fremdschlüsselprofil, um das Kindobjekt auf Fremdschlüssel zu untersuchen.

Erkennen von Fremdschlüsselbeziehungen zwischen Datenobjekten

Mit einem Enterprise-Erkennungsprofil im Developer-Tool können Sie Schlüsselbeziehungen zwischen zwei Datenobjekten suchen.

Das Datenobjekt, das den Primärschlüssel enthält, ist das übergeordnete Objekt, und das Datenobjekt, das den Fremdschlüssel enthält, ist das Kindobjekt.

1. Öffnen Sie ein Enterprise-Erkennungsprofil, das die Datenobjekte enthält, die Sie analysieren möchten.
2. Klicken Sie mit der rechten Maustaste auf den Namen eines Datenobjekts und wählen Sie **Fremdschlüsselprofil** aus.
3. Geben Sie einen Namen für das Profil ein und überprüfen Sie den Projektspeicherort. Suchen Sie ggf. nach einem neuen Speicherort. Geben Sie optional eine Textbeschreibung des Profils ein.
4. Wählen Sie die Schlüssel im übergeordneten Objekt aus, die vom Profil verwendet werden, um Fremdschlüssel im untergeordneten Objekt zu suchen.
5. Speichern Sie das Profil und führen Sie es aus.

Ergebnisse der Fremdschlüsselanalyse

Nachdem Sie ein Fremdschlüsselprofil ausgeführt haben, klicken Sie auf den Profilnamen unterhalb der Arbeitsfläche, um die Analysenergebnisse anzuzeigen.

In der Ergebnisansicht finden Sie die Spalten, die die von Ihnen definierten Primärschlüssel-Inferenzkriterien erfüllen. Klicken Sie auf die Schaltfläche **Optionen**, um die Inferenzeinstellungen zu bearbeiten. Um zu überprüfen, dass ein Inferenzschlüssel ein gültiger Schlüssel für die Datenobjekte ist, klicken Sie auf einen Spaltennamen und wählen **Validieren**.

Die Eigenschaften der Fremdschlüsselanalyse finden Sie in der folgenden Tabelle:

Eigenschaft	Beschreibung
Übergeordneter Primärschlüssel	Eine Primärschlüsselspalte im übergeordneten Datenobjekt, die vom Profil verwendet wird, um nach Fremdschlüsseln in einem untergeordneten Objekt zu suchen.
Untergeordneter Fremdschlüssel	Eine Spalte, die vom Profil als Fremdschlüssel für den übergeordneten Primärschlüssel in der aktuellen Zeile abgeleitet wird.
Inklusion %	Die Menge der Datenwerte, die bei Primär- und Fremdschlüssel übereinstimmen, ausgedrückt als Prozentsatz. Hinweis: Sie werden feststellen, dass der Wert Inklusion % für eine gefolgte Spalte in den Fremdschlüsselergebnissen sich ändert, nachdem Sie diese validiert haben. Bei einer gefolgten Spalte wird mit Inklusion % die Anzahl der einmaligen Fremdschlüsselspaltenwerte eines Kindobjekts angegeben, die mit den einmaligen Primärschlüsselwerten des übergeordneten Objekts übereinstimmen. Nachdem Sie eine Inferenzspalte validiert haben, stimmt die Anzahl der Fremdschlüssel-Spaltenwerte eines Kindobjekts mit den Primärschlüssel-Spaltenwerten des übergeordneten Objekts überein.
Beziehungstyp	Der vor dem Ausführen des Profils für die Primär- und Fremdschlüsselspalten definierte Beziehungstyp. Wenn Sie vor der Ausführung des Profils eine Beziehung definieren, gibt das Profil auch dann Daten für die Beziehung zurück, wenn der prozentuale Inklusionswert nicht innerhalb der für das Profil festgelegten Vertrauensschwelle liegt.
Überprüft	Gibt an, dass die Beziehung zwischen Primär- und Fremdschlüssel von einem Benutzer validiert wurde.
Letzte Ausführung	Datum und Uhrzeit der letzten Ausführung des Profils.
Beziehungstyp (im Modell)	Gibt an, dass das Profil die Beziehung zwischen den Spalten überprüft hat.

Join-Analyse

Die Join-Analyse beschreibt den Grad der potenziellen Joins zwischen zwei Datenspalten. Die Analyse der Spalten-Joins in einer Datenquelle bzw. über mehrere Datenquellen ist anhand eines Join-Profiles möglich.

In einem Join-Profil werden die Ergebnisse als Venn-Diagramm und als numerische und prozentuale Werte angezeigt. Ein Join-Profil wird aus einem Enterprise-Erkennungsprofil erstellt und ausgeführt.

Erstellen eines Join-Profiles

Sie können potenzielle Joins zwischen Datenobjekten in einem Enterprise-Erkennungsprofil analysieren. Das Join-Profil speichert die Analyse im Modell-Repository.

1. Erstellen oder öffnen Sie ein Enterprise-Erkennungsprofil.
2. Vergewissern Sie sich, dass das Enterprise-Erkennungsprofil die erforderlichen Datenobjekte enthält.
Um dem Join-Profil ein Datenobjekt hinzuzufügen, ziehen Sie es aus dem **Objekt-Explorer** auf die Arbeitsfläche.
3. Wählen Sie die zu profilierenden Datenobjekte aus.
4. Klicken Sie mit der rechten Maustaste auf die Objekte und wählen Sie **Join-Profil** aus.
Der Profilassistent wird geöffnet.
5. Geben Sie einen Namen für das Profil ein. Geben Sie optional eine Textbeschreibung des Profils ein.
6. Stellen Sie sicher, dass die Namen der Datenobjekte im Assistenten unter **Datenobjekte** angezeigt werden.
7. Wählen oder löschen Sie die Option zum **Ausführen des Profils nach Beenden**.
8. Klicken Sie auf **Weiter**.
9. Wählen Sie die in das Profil aufzunehmenden Datenspalten aus und klicken Sie auf **Weiter**.
Falls erforderlich, scrollen Sie durch die Datenobjekte, um alle verfügbaren Spalten anzuzeigen. Das Profil wird standardmäßig auf allen Spalten ausgeführt.
10. Klicken Sie auf **Hinzufügen**.
Das Dialogfeld **Join-Bedingung** wird angezeigt.
11. Klicken Sie auf **Neu**, um die Spaltenauswahlfelder zu aktivieren.
12. Wählen Sie die zu validierenden Datenobjekte und Spalten aus.
Zwischen zwei Spalten definieren Sie eine Join-Bedingung. Sie können mehrere Join-Bedingungen über ein oder mehrere Datenobjekte definieren.
13. Klicken Sie auf **OK**, um die Join-Bedingung zu erstellen.
Klicken Sie optional auf **Hinzufügen**, um weitere Bedingungen zu definieren.
14. Stellen Sie sicher, dass die linke und rechte Join-Spalte die richtigen Datenobjektnamen als Präfix erhalten.
15. Klicken Sie auf **Fertig stellen**.

Ergebnisse der Join-Analyse

Auf der Registerkarte Join-Analysen-**Ergebnisse** erhalten Sie Informationen über die Anzahl und den Prozentsatz der übergeordneten Waisenzeilen, der Kindwaisenzeilen und der Join-Zeilen. Die Join-Analysenergebnisse enthalten ebenfalls Venn-Diagramme, aus denen die Beziehungen zwischen Spalten hervorgehen.

In der folgenden Tabelle sind die auf der Registerkarte **Ergebnisse** angezeigten Eigenschaften beschrieben:

Eigenschaft	Beschreibung
Linke Tabelle	Name der linken Tabelle und in der Join-Analyse verwendete Spalten
Rechte Tabelle	Name der rechten Tabelle und in der Join-Analyse verwendete Spalten

Eigenschaft	Beschreibung
Nur linke Zeilen	Anzahl der Zeilen in der linken Tabelle, die nicht verbunden werden können.
Nur rechte Zeilen	Anzahl der Zeilen in der rechten Tabelle, die nicht verbunden werden können.
Join-Zeilen	Anzahl der in den Join einbezogenen Zeilen.

Wählen Sie eine Join-Bedingung für die Anzeige eines Venn-Diagramms zur Darstellung der Beziehungen zwischen den Spalten. Der Bereich unterhalb des Venn-Diagramms enthält außerdem die Anzahl und den Prozentsatz der verwaisten, Null- und verbundenen Werte in Spalten.

Mit einem Doppelklick auf einen Abschnitt im Venn-Diagramm können Sie Datensätze anzeigen, die der Abschnitt enthält. Diese Datensätze werden in der Daten-Viewer-Ansicht geöffnet.

Hinweis: Sie können die Liste der Datensätze aus der Daten-Viewer-Ansicht in eine Einfachdatei exportieren.

Exportieren von Join-Profilergebnissen in eine Datei

Sie können die für eine Join-Bedingung zurückgegebenen Datenzeilen in eine delimitierte Datei exportieren. Exportieren Sie die sich überschneidenden Zeilen der linken und rechten Quellen oder der Waisenzeilen in einer Quelle.

1. Öffnen Sie in der Ansicht **Objekt-Explorer** das Enterprise-Erkennungsprofil, in dem die Join-Analyse enthalten ist.
2. Führen Sie das Join-Profil aus.
3. Wählen Sie die Ansicht **Join-Ergebnisse** aus.
4. Klicken Sie auf der Registerkarte **Daten-Viewer** auf das Symbol **Drilldown-Ergebnisse in Datei exportieren**.

Das Dialogfenster **Daten exportieren** wird aufgerufen.

5. Geben Sie einen Dateinamen ein und klicken Sie auf **Speichern**.

Überschneidungserkennung

Durch die Überschneidungserkennung erhalten Sie Informationen zu Datenüberschneidungen in zwei Spalten einer Datenquelle oder mehreren Datenquellen. Sie können sich überschneidende Daten eines Enterprise-Erkennungsprofils suchen. Sie können die Profilergebnisse validieren und die Ergebnisse in Form eines Venn-Diagramms anzeigen.

Bei der Überschneidungserkennung werden sich überschneidende Daten basierend auf den Standardeinstellungen oder den von Ihnen vorgenommenen Einstellungen erkannt. Sie können die Standardeinstellungen überschreiben und Störungsoptionen angeben. Dazu zählt die maximale Anzahl an Top-Paaren, die die Überschneidungserkennung basierend auf dem Prozentsatz der Überschneidung zurückgibt. Sie können ebenfalls eine Truststufe angeben, mit der Sie die Eignung für die Überschneidungserkennung definieren.

Ergebnisse der Überschneidungserkennung

Die Registerkarte **Überschneidungserkennung** zeigt Ihnen Informationen über die betroffenen Spalten und den prozentualen Anteil der Überschneidungen an. In den Ergebnissen der Überschneidungserkennung finden Sie Venn-Diagramme zur Darstellung der sich überschneidenden Daten in Spaltenpaaren mit Angabe von Datum und Uhrzeit der zuletzt durchgeführten Überschneidungserkennung.

Sie können eine Spalte anklicken und **Überprüfen** wählen, um die Ergebnisse als Venn-Diagramm anzuzeigen.

In der folgenden Tabelle sind die Eigenschaften der Überschneidungserkennung beschrieben:

Eigenschaft	Beschreibung
Linke Spalte	Die primäre Spalten, mit der die übrigen Spalten bei der Überschneidungserkennung verglichen werden.
Rechte Spalte	Die Spalte, die mit der primären Spalte verglichen wird.
% Überschneidung	Der prozentuale Anteil der Überschneidung zwischen zwei Spalten.
Überprüft	Zeigt an, dass Sie die Überschneidungsergebnisse validiert haben.
Letzte Ausführung	Datum und Uhrzeit der letzten Überschneidungserkennung.

In Informatica Developer wird jedes sich überschneidende Paar zwei Mal in den Überschneidungserkennungs-Ergebnissen angezeigt. Beachten Sie die Datenquellen "Einträge" und "Aufträge". „Elemente“ enthält die Spalten „m“ und „n“. „Aufträge“ enthält die Spalten „p“ und „q“.

Die folgende Tabelle zeigt die Ergebnisse der Überschneidungserkennung für „Elemente“ und „Aufträge“:

Linke Spalte	Rechte Spalte
Einträge	-
m	Orders.p
m	Orders.q
n	Orders.p
n	Orders.q
Aufträge	-
p	Items.m
p	Items.n
q	Items.m
q	Items.m

Erkennen von sich überschneidenden Daten

Sie können sich überschneidende Daten zwischen Spaltenpaaren in einem Enterprise-Erkennungsprofil bestimmen. Die Überschneidungserkennung basiert auf eindeutigen Werten in den Spalten und berücksichtigt keine Null-Werte.

1. Erstellen oder öffnen Sie ein Enterprise-Erkennungsprofil, das die Datenobjekte enthält.
2. Wählen Sie die Datenobjekte aus, in denen Sie nach sich überschneidenden Daten suchen möchten.
Sie können ein einzelnes Datenobjekt auswählen, um nach sich überschneidenden Daten innerhalb von Spaltenpaaren zu suchen, oder mehrere Datenobjekte auswählen.
3. Klicken Sie mit der rechten Maustaste auf die Objekte und wählen Sie **Überschneidungserkennung** aus.
Das Dialogfeld **Neue Überschneidungserkennung** wird angezeigt.
4. Geben Sie einen Namen ein.
5. Geben Sie optional eine Textbeschreibung für die Überschneidungserkennung ein.
6. Stellen Sie sicher, dass die Namen der Datenobjekte im Assistenten unter **Datenobjekte** angezeigt werden.
7. Wählen Sie optional **Ausführen des Profils nach Beenden** aus, um das Profil auszuführen, wenn die Konfiguration der Einstellungen abgeschlossen ist.
8. Klicken Sie auf **Weiter**.
9. Wählen Sie die Spalten für die Überschneidungserkennung aus.
10. Klicken Sie auf **Weiter**.
Die Standardinferenzoptionen werden im Dialogfeld angezeigt.
11. Geben Sie optional die Inferenzoptionen für die Überschneidungserkennung an, um die Standardeinstellungen zu überschreiben.
12. Klicken Sie auf **Fertig stellen**.

DDL-Skriptdateien

Die Skriptdateien für die Datendefinitionssprache (DDL) enthalten die SQL-Anweisungen **CREATE**, **ALTER** und **DROP**.

Sie können einen Dateinamen, einen Speicherort und einen Zieldatenbanktyp beim Generieren von Skriptdateien angeben. Das Developer-Tool hängt an die Skriptdateinamen die Beschriftungen "_create" und "_drop". Virtuelle Spalten sind nicht Teil der DDL-Skriptdateien.

Erstellen von DDL-Skripts aus einem Enterprise-Erkennungsprofil

Beim Generieren von DDL-Skriptdateien aus einem Enterprise-Erkennungsprofil können Sie den Speicherort für die Skriptdateien auswählen. Sie können auch den Datenbanktyp wählen, für den Sie die Skripte ausführen möchten. Stellen Sie sicher, dass Sie alle erforderlichen Änderungen im Enterprise-Erkennungsprofil prüfen und einen Commit ausführen, bevor Sie DDL-Skripts generieren.

1. Wählen Sie in der Ansicht **Objekt-Explorer** ein Enterprise-Erkennungsprofil aus.
2. Klicken Sie mit der rechten Maustaste auf das Profil und wählen Sie **DDL generieren** aus.
Das Dialogfeld **DDL generieren** wird eingeblendet.

3. Klicken Sie auf **Durchsuchen**, um das Dialogfeld **Speichern unter** zu öffnen.
Die Standarddateierweiterung lautet `.sql`.
4. Wählen Sie einen Dateispeicherort aus und geben Sie einen Dateinamen ein.
5. Wählen Sie den Typ der Zieldatenbank aus.
6. Klicken Sie auf **OK**.
Das Developer-Tool generiert die DDL-Skriptdateien im Verzeichnis, das Sie angegeben haben.

Synchronisieren eines Enterprise-Erkennungsprofils

Sie können ein Enterprise-Erkennungsprofil im Developer Tool synchronisieren.

Nach der Aktualisierung von Version 9.5 oder früher auf Version 9.6 oder höher können Sie die Profile aus der vorherigen Version auf die aktualisierte Version migrieren. Wenn Sie bei Enterprise-Erkennungsprofilen in der vorherigen Version benutzerdefinierte Schlüssel, dokumentierte Schlüssel oder Beziehungen hinzugefügt haben, bleiben die Schlüssel- und Beziehungsinformationen nur im Modellrepository und nicht im Profil-Warehouse bestehen. Wenn Sie in der aktualisierten Version das Enterprise-Erkennungsprofil im Developer Tool öffnen, werden die dokumentierten oder benutzerdefinierten Schlüssel und Beziehungen nicht in den kuratierten Ergebnissen für das Profil angezeigt.

Um die benutzerdefinierten Schlüssel, die dokumentierten Schlüssel und die Beziehungen im Modellrepository mit dem Profiling Warehouse zu synchronisieren, verwenden Sie die Option „Enterprise-Erkennungsprofil synchronisieren“ im Developer Tool. Nachdem Sie das Enterprise-Erkennungsprofil synchronisiert haben, werden die benutzerdefinierten und dokumentierten Schlüssel und Beziehungen auf „Genehmigen“ festgelegt, und Sie können die kuratierten Ergebnisse im Developer Tool anzeigen.

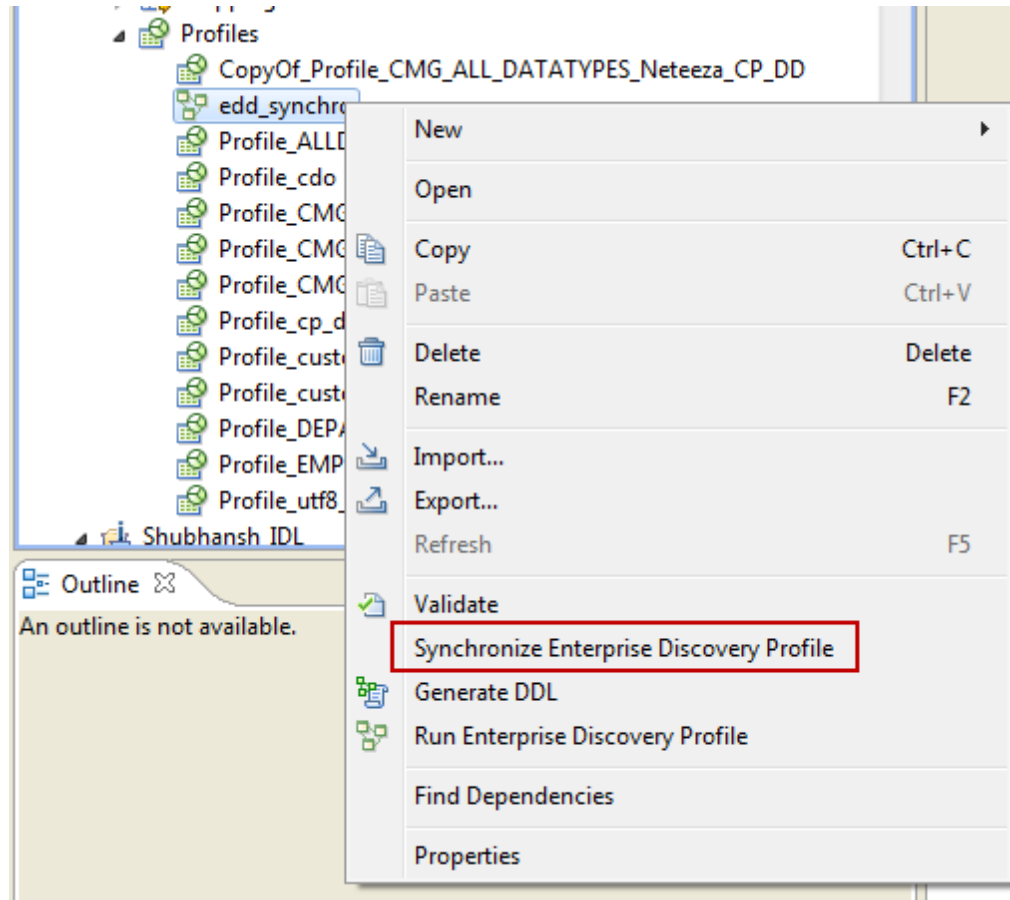
Synchronisieren eines Enterprise-Erkennungsprofils

In Informatica Developer können Sie die kuratierten Ergebnisse für ein Enterprise-Erkennungsprofil synchronisieren, nachdem Sie von Version 9.5 oder früher auf Version 9.6 oder höher aktualisiert haben.

1. Wählen Sie in der Ansicht **Objekt-Explorer** ein Enterprise-Erkennungsprofil aus.

2. Klicken Sie mit der rechten Maustaste auf das Profil, und wählen Sie **Enterprise-Profil synchronisieren** aus.

Die folgende Abbildung zeigt die Option „Enterprise-Erkennungsprofil synchronisieren“ im Developer Tool:



Die kuratierten Ergebnisse für das Profil sind synchronisiert.

KAPITEL 26

Ergebnisse der Enterprise-Erkennung

Dieses Kapitel umfasst die folgenden Themen:

- [Enterprise-Erkennungsergebnisse - Übersicht, 215](#)
- [Beziehungsansicht, 216](#)
- [Fremdschlüssel-Profiling-Ansicht, 217](#)
- [Tabellenansicht, 220](#)
- [Datendomänenansicht, 222](#)
- [Spaltenprofilansicht, 223](#)
- [Anzeigen von Spaltenprofilergebnissen während der Ausführung der Enterprise-Erkennung, 223](#)
- [Anzeigen von Datendomänenenerkennungsergebnissen während der Ausführung der Enterprise-Erkennung, 224](#)
- [Anzeigen des Laufzeitstatus der Enterprise-Erkennung, 224](#)
- [Enterprise-Erkennungs-Exportdateien, 225](#)

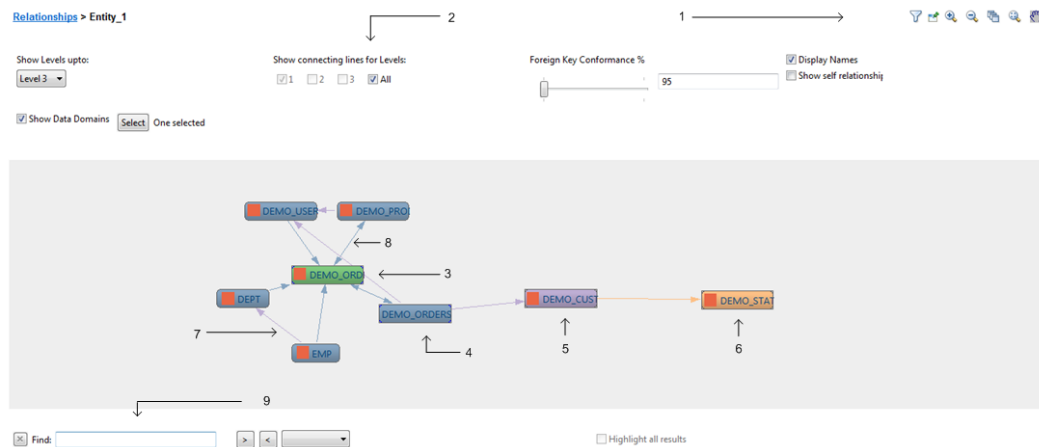
Enterprise-Erkennungsergebnisse - Übersicht

Sie können die Ergebnisse der Enterprise-Erkennung in mehreren Ansichten anzeigen. Die Ansichten sind **Beziehungen**, **Datendomänen**, **Spaltenprofil**, **Join-Profil** und **Überschneidungserkennung**.

In der Ansicht **Beziehungen** werden Gruppen von Datenobjekten als Kreise dargestellt. Aus dieser Ansicht heraus können Sie die Fremdschlüsselprofilergebnisse aufrufen. Die Fremdschlüsselprofilergebnisse werden sowohl in grafischer als auch tabellarischer Form angezeigt. In der Ansicht **Datendomänen** werden die Datendomänenenerkennungsergebnisse angezeigt. In der Ansicht **Spaltenprofil** werden Spaltenprofilergebnisse für die einzelnen Datenobjekte angezeigt. Die Ansicht **Join-Profil** zeigt die Anzahl der übergeordneten verwaisten Zeilen, untergeordneten verwaisten Zeilen und der in den Join einbezogenen Zeilen an. Auf der Registerkarte **Überschneidungserkennung** erhalten Sie Informationen über die betroffenen Spalten und den prozentualen Anteil der Überschneidungen.

Datenobjekte können über mehrere Beziehungen untereinander verfügen. In der grafischen Ansicht der Fremdschlüsselergebnisse werden Datenobjektbeziehungen angezeigt, die den höchsten Konformitätsprozentsatz aufweisen.

Die folgende Abbildung zeigt die grafische Ansicht einiger Beispieldatenobjekte aus den Enterprise-Erkennungsergebnissen:



1. Symbolleisten, die Symbole zum Filtern, Anheften von Datenobjekten, Heranzoomen, Herauszoomen, Anordnen aller Fenster, Anpassen des Inhalts an die Fenstergröße und Schwenken enthalten.
2. Filteroptionen zum Anzeigen von z. B. unterschiedlichen Ebenen von Datenobjektbeziehungen, Verbindungslinien und selbstbezogenen Datenobjekten.
3. Das ausgewählte Datenobjekt, auf Basis dessen der visuelle Editor den Rest der Beziehungen zwischen Datenobjekten anzeigt.
4. Erste Ebene von Datenobjektbeziehungen.
5. Zweite Ebene von Datenobjektbeziehungen.
6. Dritte Ebene von Datenobjektbeziehungen.
7. Connector zwischen Datenobjekten. Eine einzelne Pfeilspitze weist auf eine Primärschlüssel- zu Fremdschlüssel-Beziehung zwischen zwei Datenobjekten hin. Die Pfeilspitze zeigt auf das Daten-Objekt mit dem Primärschlüssel.
8. Connector zwischen Datenobjekten. Ein Connector mit einer doppelte Pfeilspitze weist auf eine Primärschlüssel- zu Primärschlüssel-Beziehung zwischen zwei Datenobjekten hin. Bewegen Sie die Maus auf den Connector, um die Join-Spalten mit maximaler Konformität für die abgeleitete Beziehung zu sehen.
9. Drücken Sie STRG+F, um das Suchfeld anzuzeigen, und nutzen Sie das Sternchen (*) als Platzhalter, um Datenobjekte in der grafischen Ansicht zu finden.

Beziehungsansicht

Sie können eine Zusammenfassung der Enterprise-Erkennungsergebnisse in der Ansicht **Beziehungen** einschließlich der Entitäten anzeigen. Entitäten sind Gruppen von Datenobjekten, die als Kreise dargestellt werden. Zu den Entitäten zählen zugehörige Objekte, auf sich selbst bezogene Objekte und Objekte ohne Bezug aus den Verbindungen und Schemata der Quelldatenbanken.

Ein auf sich selbst bezogenes Datenobjekt enthält Spalten, die Beziehungen haben. Ein Datenobjekt ohne Bezug hat weder Beziehungen zu anderen Datenobjekten in der Quelldatenbank, noch hat es Beziehungen zwischen den Spalten innerhalb des Datenobjekts. Das Entitäten-Beziehungsdiagramm der Datenobjekte in den Enterprise-Erkennungsergebnissen basiert auf abgeleiteten Beziehungen und nicht auf dokumentierten Beziehungen in den Datenquellen.

Suchen nach einem Datenobjekt

Sie können in der Ansicht **Beziehungen** oder **Fremdschlüssel-Profilng** nach einem Datenobjekt suchen. Dabei können Sie ein Sternchen (*) als Platzhalter verwenden, um Datenobjekte zu suchen.

1. Stellen Sie sicher, dass Sie in der Ansicht **Beziehungen** oder **Fremdschlüssel-Profilng** sind.

2. Geben Sie einen Teil des gesuchten Datenobjektnamens ein und fügen Sie je nach Bedarf vor oder hinter dem Suchstring das Platzhalterzeichen * ein. Um beispielsweise nach allen Datenobjekten zu suchen, die mit der Zeichenfolge "CA" beginnen, geben Sie "CA*" ein und drücken die **Eingabetaste**. Um nach allen Datenobjekten zu suchen, die in ihrem Namen die Zeichenfolge "ZIP" enthalten, geben Sie "*ZIP*" ein. Bei der Suche wird Groß-/Kleinschreibung unterschieden.

Navigieren zur Fremdschlüsselprofilansicht

In der Ansicht **Fremdschlüssel-Profiling** wird eine konsolidierte Sicht der Fremdschlüsselbeziehungen über mehrere Datenobjekte angezeigt, für die Sie das Profil ausgeführt haben. Die Kreise in der Ansicht kennzeichnen Entitäten, auf sich selbst bezogene Objekte und Objekte ohne Bezug.

1. Stellen Sie sicher, dass Sie in der Ansicht **Beziehungen** sind.
Den Link für das Fremdschlüsselprofil finden Sie im rechten Fensterausschnitt.
2. Klicken Sie zum Öffnen der Ansicht auf **Fremdschlüsselprofil**.
Die Ansicht wird auf einer neuen Registerkarte angezeigt. Die Ansicht zeigt Datenobjekte in verschiedenen Arten von Kreisen, je nach Beziehungstyp. Außerdem sehen Sie die Gesamtzahl der Datenobjekte, die Bestandteil der konsolidierten Fremdschlüsselansicht sind.
3. Optional können Sie auf die Verknüpfung **Beziehungen** klicken, um zur Ansicht **Beziehungen** zurückzukehren.

Fremdschlüssel-Profiling-Ansicht

Sie können eine Zusammenfassung der Enterprise-Erkennungsergebnisse in grafischer Form in der Ansicht **Fremdschlüssel-Profiling** anzeigen. Sie können die Profilergebnisse und Beziehungen auf Spaltenebene für ein Datenobjekt in tabellarischer Form über die Ansicht öffnen.

Die Profilergebnisse für ein Datenobjekt umfassen das Spaltenprofil, die Primärschlüsselinferenz, die funktionale Abhängigkeitsinferenz sowie die Ergebnisse der Datendomänenenerkennung. Nachdem Sie die Beziehungen auf Spaltenebene für ein Datenobjekt geöffnet haben, können Sie die Datenbeziehungen überprüfen und wiederherstellen. Wenn Sie eine Datenbeziehung überprüfen, führt das Developer-Tool das Profil über alle Zeilen der Datenquelle aus, um die Inferenzergebnisse zu verifizieren. Sie können Datenbeziehungen in der Ansicht **Fremdschlüssel-Profiling** genehmigen, ablehnen und zurücksetzen.

Hinweis: In der Ansicht **Fremdschlüssel-Profiling** werden keine Datenobjekte angezeigt, wenn Sie zum Erstellen eines Enterprise-Erkennungsprofils Hive-Datenquellen verwenden.

Anzeigen von Datenobjektbeziehungen

Sie können die Beziehungen zwischen Datenobjekten in grafischer Form anzeigen. Doppelklicken Sie auf einen Entitätskreis, um die Tabellen und deren Beziehungen anzuzeigen.

1. Stellen Sie sicher, dass Sie die Ansicht **Fremdschlüssel-Profiling** anzeigen.
2. Um Datendomänen in die konsolidierte grafische Ansicht aufzunehmen, wählen Sie **Datendomänen anzeigen** aus.
Die Schaltfläche **Auswählen** wird aktiviert.
3. Klicken Sie auf **Auswählen**, um die Datendomänen auszuwählen, die Sie in die grafische Ansicht aufnehmen möchten.

Das Dialogfeld **Datendomänen auswählen** wird angezeigt.

4. Wählen Sie die erforderlichen Datendomänen aus und klicken Sie auf **OK**.

Im Developer-Tool werden die Entitätskreise hervorgehoben, in denen die Datendomänen enthalten sind, die Sie ausgewählt haben.

5. Doppelklicken Sie auf einen Entitätskreis, um die visuelle Darstellung der Tabellenbeziehungen innerhalb der Entität anzuzeigen. Das Developer-Tool zeigt die Tabellen in grafischer Form an, wobei die Beziehungen jedes Datenobjekts mit anderen Datenobjekten innerhalb der Entität dargestellt werden.

Das Datenobjekt mit der höchsten Anzahl an Beziehungen mit anderen Datenobjekten oder die Datenobjekte, von denen aus Sie mit der Navigation beginnen, werden in Grün hervorgehoben. Wenn Sie Datendomänen aufgenommen haben, wird die Datendomänenauswahl im Developer-Tool links neben der visuellen Darstellung jedes Datenobjekts hervorgehoben.

6. Überprüfen Sie die Informationen zur direkten Beziehung und zur Datendomäne im rechten Bereich der Ansicht **Fremdschlüssel-Profiling**.
7. Optional können Sie auf die Verknüpfung **Beziehungen** klicken, um zur Ansicht **Beziehungen** zurückzukehren.

Vergrößern und Verkleinern der Ansicht

Die grafische Darstellung der Datenobjektbeziehungen kann zur deutlicheren Ansicht in der Ansicht **Fremdschlüssel-Profiling** vergrößert werden. Beim Vergrößern erhöht das Developer Tool den Zoomfaktor des Bildes. Mit Verkleinern reduzieren Sie den Zoomfaktor.

1. Stellen Sie sicher, dass Sie in der Ansicht **Fremdschlüssel-Profiling** sind.
2. Klicken Sie mit der rechten Maustaste auf die Ansicht und wählen Sie **Vergrößern**, um den Zoomfaktor des Bildes zu erhöhen.
3. Zum Reduzieren des Zoomfaktors der grafischen Darstellung klicken Sie mit der rechten Maustaste auf die Ansicht und wählen **Verkleinern**.

Suchen nach einem Datenobjekt

Sie können in der grafischen Ansicht der Fremdschlüsselergebnisse nach einem Datenobjekt suchen. Verwenden Sie ein Sternchen (*) als Platzhalterzeichen, um nach Datenobjekten zu suchen.

1. Stellen Sie sicher, dass Sie in der Ansicht **Fremdschlüssel-Profiling** sind.
Stellen Sie sicher, dass Sie die grafische Ansicht von Fremdschlüsselergebnissen in der Ansicht öffnen.
2. Drücken Sie Strg+F, um das Feld **Suchen** anzuzeigen.
3. Geben Sie im Feld **Suchen** einen Teil des Datenobjektnamens ein, nach dem Sie suchen, und fügen Sie basierend auf der Suchanforderung das Platzhalterzeichen * am Anfang oder Ende der Suchzeichenfolge hinzu. Um beispielsweise nach allen Datenobjekten zu suchen, die mit der Zeichenfolge "CA" beginnen, geben Sie "CA*" ein und drücken die **Eingabetaste**. Um nach allen Datenobjekten zu suchen, die in ihrem Namen die Zeichenfolge "ZIP" enthalten, geben Sie "*ZIP*" ein.
4. Klicken Sie auf die Schaltfläche **Nächster Treffer**, um zur nächsten Datenobjektübereinstimmung zu gelangen.
Verwenden Sie die Schaltfläche **Vorheriger Treffer**, um zur vorherigen Datenobjektübereinstimmung zu gelangen.
5. Wählen Sie **Alle Ergebnisse markieren**, um alle übereinstimmenden Datenobjekte zu markieren.
6. Zum Löschen der Suchzeichenfolge im Feld **Suchen** klicken Sie auf die Schaltfläche **Löschen** neben dem Feld.

Anzeigen von Spaltenbeziehungen

Sie können die Beziehung jeder Spalte in einem Datenobjekt zu Spalten in zugehörigen Datenobjekten anzeigen. Außerdem können Sie die Spaltenbeziehungen überprüfen und die Datenobjektbeziehungen in das Datenmodell übernehmen.

1. Stellen Sie sicher, dass Sie die Ansicht **Fremdschlüssel-Profiling** anzeigen.
2. Klicken Sie mit der rechten Maustaste auf ein Datenobjekt und wählen Sie **Spaltenbeziehungen anzeigen** aus.

Die Spaltenbeziehungen werden in einer Tabellenansicht angezeigt. In der Ansicht werden Beziehungsinformationen wie Quelldatenobjekt, verbundenes Datenobjekt und Spalten im verbundenen Datenobjekt angezeigt.

3. Überprüfen Sie den Inferenzstatus, Überprüfungsstatus und Wiederherstellungsstatus.
4. Wählen Sie **Alle Datenobjekte in der Gruppe** aus, um alle Datenobjekte in der übergeordneten Entität und ihre Spaltenbeziehungsinformationen anzuzeigen.

Gemäß Voreinstellung werden in der Ansicht die Beziehungsinformationen für das in der Ansicht ausgewählte Datenobjekt angezeigt.

5. Sie können auch auf die Verknüpfung **Entität** oben in der Ansicht klicken, um zur grafischen Darstellung der Datenobjekte zurückzukehren.

Speichern des Entitätsbeziehungsdiagramms als Symbol

Sie können das Entitätsbeziehungsdiagramm von Datenobjekten aus den Enterprise-Erkennungsergebnissen als PNG-Datei speichern.

1. Führen Sie ein Profil aus, um die Enterprise-Erkennung durchzuführen.
2. Stellen Sie sicher, dass Sie in der Ansicht **Fremdschlüssel-Profiling** sind.
3. Wechseln Sie aus der Ansicht ins Grafikformat der Datenobjektbeziehungen.
4. Klicken Sie mit der rechten Maustaste und wählen Sie **Als Bild speichern**.

Das Dialogfeld **Speichern unter** wird eingeblendet. Standardmäßig speichern Sie das Bild als PNG-Datei.

5. Wählen Sie den Dateispeicherort und geben Sie einen Dateinamen ein.
6. Klicken Sie auf **Speichern**.

Anzeigen von Datenobjektprofilergebnissen aus der Ansicht Fremdschlüssel-Profiling heraus

Sie können Spaltenprofil, Primärschlüssel und Datendomänenerkennungsergebnisse eines ausgewählten Datenobjekts aus der Ansicht **Fremdschlüssel-Profiling** heraus anzeigen. Vergewissern Sie sich, dass Sie das Datenobjekt durch Anheften als ausgewählte Tabelle auf der Arbeitsfläche ausgewählt haben.

1. Stellen Sie sicher, dass Sie in der Ansicht **Fremdschlüssel-Profiling** sind.
2. Klicken Sie mit der rechten Maustaste auf ein Datenobjekt und wählen Sie **Datenobjekt als Fokus anheften**, um die Tabelle auszuwählen.

Sie können auch das Symbol **Anheften** verwenden, um ein Datenobjekt auszuwählen.

3. Klicken Sie mit der rechten Maustaste auf eine beliebige Stelle auf der Arbeitsfläche und wählen Sie **Datenobjektprofil anzeigen**.

Die Datenobjektprofilergebnisse werden auf einer Registerkarte angezeigt.

Tabellenansicht

Wenn die Ansicht **Fremdschlüssel-Profiling** geöffnet wird, zeigt das Developer Tool standardmäßig die grafische Ansicht der Ergebnisse. Wechseln Sie zu einer Tabellenansicht, um die Tabellen und deren Beziehungsdetails in tabellarischer Form anzuzeigen.

Sie können die Anzahl der Datenobjekte in der Entität, die Namen der zugehörigen Tabellen, ihre Verbindungsinformationen und die Anzahl der Beziehungen zwischen den beiden Datenobjekten anzeigen. Außerdem können Sie die Spaltenbeziehungen überprüfen und sie einem Datenmodell hinzufügen.

Tabellendetailbereich

Sie können die Datenobjektdetails in der grafischen und in der tabellarischen Ansicht der Enterprise-Erkennungsergebnisse anzeigen. In der grafischen Ansicht sehen Sie im Tabellendetailbereich die Anzahl der Datenobjekte, die eine direkte Beziehung zu dem ausgewählten Datenobjekt haben, und die Datenobjektnamen.

In der folgenden Tabelle werden die Spalten des Tabellendetailbereichs in der Tabellenansicht beschrieben:

Spaltenname	Beschreibung
Tabellenname	Name des Datenobjekts, das eine direkte Beziehung zu dem ausgewählten Datenobjekt auf der linken Seite hat.
Verbindung	Name der Verbindung für das zugehörige Datenobjekt.
Beziehungen	Die Anzahl der Beziehungen zwischen dem ausgewählten Datenobjekt auf der linken Seite und dem zugehörigen Datenobjekt im Tabellendetailbereich.

Überprüfen der Enterprise-Erkennungsergebnisse

Wenn Sie die Ergebnisse der Enterprise-Erkennung überprüfen, führt das Developer Tool das Profil über alle Zeilen der Datenquelle aus. Der Wert des Konformitätsprozentsatzes kann nach der Überprüfung abhängig von den Spaltenwerte über alle Zeilen in der Datenquelle variieren.

1. Öffnen eines Profils nach dessen Ausführung.
2. Stellen Sie sicher, dass Sie in der Ansicht **Fremdschlüssel-Profiling** sind.
3. Klicken Sie oben in der Ansicht auf das Symbol **Tabellenansicht**.

In der Tabellenansicht werden die Entitäten im linken Bereich angezeigt.

4. Klicken Sie mit der rechten Maustaste im linken Bereich auf ein Datenobjekt und wählen Sie **Spaltenbeziehungen anzeigen**.

Sie können die Beziehungen der Spalten im ausgewählten Datenobjekt zu Spalten in anderen Datenobjekten anzeigen. Blättern Sie nach rechts, um Details wie Beziehungstyp, Konformitätsprozentsatz, Überprüfungsstatus und Übernahmestatus zu sehen.

5. Klicken Sie mit der rechten Maustaste auf eine Zeile, und wählen Sie **Überprüfen**.

Das Dialogfeld **Profil ausführen** wird angezeigt. Nach Abschluss der Überprüfung wählen Sie die Zeile aus, um die Überlappung von Primärschlüssel- und Fremdschlüssel-Beziehungen in einem Venn-Diagramm anzuzeigen.

Wiederherstellen von Spaltenbeziehungen

Sie können Datenbeziehungen in der Ansicht **Fremdschlüssel-Profil** genehmigen, ablehnen und zurücksetzen.

1. Öffnen Sie ein Profil nach dessen Ausführung.
2. Stellen Sie sicher, dass Sie die Ansicht **Fremdschlüssel-Profil** anzeigen.
3. Wenn Sie in der grafischen Ansicht eine abgeleitete Spaltenbeziehung ablehnen möchten, wählen Sie ein Datenobjekt aus, klicken mit der rechten Maustaste darauf und wählen eine der folgenden Optionen aus:
 - **Alle Beziehungen ablehnen mit > Abgeleitete Primärschlüssel.** Wählen Sie diese Option aus, um alle Beziehungen zwischen den Spalten mit abgeleiteten Primärschlüsseln im Datenobjekt und den Spalten mit abgeleiteten Fremdschlüsseln in den anderen verbundenen Datenobjekten abzulehnen.
 - **Alle Beziehungen ablehnen mit > Abgeleitete Fremdschlüssel.** Wählen Sie diese Option aus, um alle Beziehungen zwischen den Spalten mit abgeleiteten Fremdschlüsseln im Datenobjekt und den Spalten mit abgeleiteten Primärschlüsseln in den anderen verbundenen Datenobjekten abzulehnen.
 - **Alle Beziehungen ablehnen mit > Abgeleitete Primär- und Fremdschlüssel.** Wählen Sie diese Option aus, um alle Beziehungen zwischen den Spalten mit abgeleiteten Primärschlüsseln im Datenobjekt und den Spalten mit abgeleiteten Fremdschlüsseln in den anderen verbundenen Datenobjekten sowie alle Beziehungen zwischen den Spalten mit abgeleiteten Fremdschlüsseln im Datenobjekt und den Spalten mit abgeleiteten Primärschlüsseln in den anderen verbundenen Datenobjekten abzulehnen.
4. Klicken Sie in der grafischen Ansicht mit der rechten Maustaste auf ein Datenobjekt und wählen Sie **Spaltenbeziehungen anzeigen** aus.
5. Wählen Sie eine Datenobjektbeziehung aus, die Sie wiederherstellen möchten.
6. Um die Spaltenbeziehung zu genehmigen, klicken Sie mit der rechten Maustaste und klicken dann auf **Genehmigen**.
Der Status der Zeile ändert sich auf **Genehmigt**.
7. Um den abgeleiteten Status der Spaltenbeziehung wiederherzustellen, klicken Sie mit der rechten Maustaste und klicken dann auf **Zurücksetzen**.
8. Um abgelehnte Spaltenbeziehungen anzuzeigen, klicken Sie mit der rechten Maustaste auf eine der Zeilen und wählen dann **Abgelehnte einblenden** aus.
9. Klicken Sie zum Ausblenden der abgelehnten Datentypen mit der rechten Maustaste auf eine der Zeilen und wählen Sie **Abgelehnte ausblenden** aus.

Übergeben der Ergebnisse an das Modellrepository

Nachdem Sie ein Profil ausgeführt haben, können Sie Spaltenbeziehungen zwischen Datenobjekten im Modellrepository speichern. Sie können die Beziehungen an das Modellrepository über die Tabellenansicht der Ansicht **Fremdschlüssel-Profil** übergeben.

1. Öffnen Sie ein Profil nach dessen Ausführung.
2. Stellen Sie sicher, dass Sie die Ansicht **Fremdschlüssel-Profil** anzeigen.
3. Klicken Sie mit der rechten Maustaste im linken Bereich auf ein Datenobjekt und wählen Sie **Spaltenbeziehungen anzeigen** aus.
Sie können die Beziehungen der Spalten im ausgewählten Datenobjekt zu Spalten in anderen Datenobjekten anzeigen.
4. Klicken Sie mit der rechten Maustaste auf eine Zeile und wählen Sie **Genehmigen** aus.

Datendomänenansicht

In der Ansicht **Datendomänen** werden die Datendomänen und Statistiken der passenden Spalten aufgelistet, die das Developer Tool im Rahmen der Enterprise-Erkennung erkennt. Sie können in der Ansicht **Datendomänen** Spalten überprüfen, einen Drilldown auf Zeilen ausführen und Datenobjektprofilerggebnisse anzeigen.

Anzeigen von Datendomänenenerkennungsergebnissen

Sie können die Datendomänenenerkennungsergebnisse auf der Registerkarte **Datendomänen** sehen. Sie können nach Datendomänen suchen und die Datendomänen nach Datendomänengruppen sortiert anzeigen.

1. Ausführen eines Profils zur Ausführung der Enterprise-Erkennung.
2. Öffnen Sie das Profil.
3. Klicken Sie auf die Registerkarte **Datendomänen**, um die Datendomänenenerkennungsergebnisse anzuzeigen.

Die Datenobjektprofilerggebnisse werden im rechten Fensterbereich angezeigt.

4. Geben Sie einen Teil eines Datendomänennamens in das Suchfeld ein, um eine bestimmte Datendomäne zu suchen.

Wählen Sie **Datendomänengruppe in Hierarchie anzeigen**, um die Liste der Datendomänen nach Datendomänengruppen zu sortieren.

Überprüfen der Datendomänenenerkennungsergebnisse

Bei der Ausführung eines Profils wird eine Stichprobe der Datenquelle analysiert, um auf die Profilergebnisse zu schließen. Sie können das Profil über alle Zeilen der Quelldatenbank ausführen, um die Rückschlussergebnisse zu verifizieren.

1. Öffnen Sie ein Profil nach dessen Ausführung.
2. Klicken Sie auf die Registerkarte **Datendomänen**, um die Ergebnisse anzuzeigen.
Sie können die Datendomänenenerkennungsergebnisse im rechten Fensterausschnitt sehen.
3. Wählen Sie im rechten Bereich eine zu überprüfende Spalte aus.
4. Klicken Sie mit der rechten Maustaste in die Spalte und wählen Sie **Überprüfen**, um das Profil über allen Zeilen der Datenquelle auszuführen.

Unter Umständen wird nach der Überprüfung der Ergebnisse eine Änderung des Werts % **Datenkonformität** oder des Werts **Anzahl konformer Zeilen** angezeigt.

Zeilen-Drilldown

Zur weiteren Datenanalyse ist ein Drilldown der Datendomänenenerkennungsergebnisse möglich.

1. Öffnen eines Profils nach dessen Ausführung.
2. Klicken Sie auf die Registerkarte **Datendomänen**, um die Ergebnisse anzuzeigen.
Sie können die Datendomänenenerkennungsergebnisse im rechten Fensterausschnitt sehen.
3. Wählen Sie im rechten Bereich eine Spalte für den Drilldown aus.
4. Klicken Sie mit der rechten Maustaste auf die Spalte und wählen Sie **Drilldown** für den Drilldown der Quellzeilen.

Anzeigen von Datenobjektprofilergebnissen aus der Datendomänenansicht

Sie können die Datenobjektprofilergebnisse eines ausgewählten Datenobjekts aus der Ansicht **Datendomänen** heraus anzeigen.

1. Stellen Sie sicher, dass Sie sich in der Ansicht **Datendomänen** befinden.
2. Wählen Sie im Bereich **Profilierte Domänen** eine Datendomäne aus.
3. Wählen Sie im Bereich **Spalten** auf der rechten Seite eine Spalte aus.
4. Klicken Sie mit der rechten Maustaste auf die Spalte, und wählen Sie **Datenobjektprofil öffnen**.

Die Datenobjektergebnisse werden auf einer Registerkarte angezeigt.

Spaltenprofilansicht

Die Ansicht **Spaltenprofil** zeigt eine Zusammenfassung der Spaltenprofilergebnisse für einzelne Datenobjektprofile, die das Developer-Tool als Teil der Enterprise-Erkennung ausführt. Sie können die Spaltenstatistiken, wie z. B. eindeutige Werte, Nullwerte, Datentypen sowie Maximal- und Mindestwerte, für jede Spalte in den Datenobjekten anzeigen.

Anzeigen von Datenobjektprofilergebnissen

Die Enterprise-Erkennung umfasst die Ausführung eines Datenobjektprofils zur Erkennung der Spaltendatenstatistik, der Primärschlüssel und Schlüsselkandidaten und der Datendomänen. Sie können die Datenobjektprofilergebnisse eines ausgewählten Datenobjekts aus der Ansicht **Datenobjektprofil** heraus anzeigen.

1. Stellen Sie sicher, dass Sie sich in der Ansicht **Spaltenprofil** befinden.
2. Wählen Sie im Bereich **Profilierte Objekte** ein Datenobjekt aus.
3. Wählen Sie im Bereich **Spalten** auf der rechten Seite eine Spalte aus.
4. Klicken Sie mit der rechten Maustaste auf die Spalte, und wählen Sie **Datenobjektprofil anzeigen**.

Die Datenobjektprofilergebnisse werden auf einer Registerkarte angezeigt. Die Spaltenprofilergebnisse werden standardmäßig angezeigt.

5. Klicken Sie auf **Primärschlüsselerkennung**, um die Primärschlüsselergebnisse anzuzeigen.
6. Klicken Sie auf **Funktionale Abhängigkeitsinferenz**, um die funktionalen Abhängigkeitserkennungsergebnisse anzuzeigen.
7. Klicken Sie auf **Datendomänenerkennung**, um die Datendomänenerkennungsergebnisse anzuzeigen.

Anzeigen von Spaltenprofilergebnissen während der Ausführung der Enterprise-Erkennung

Die für die Ausführung der Enterprise-Erkennung benötigte Zeit hängt von der Anzahl der Profil-Tasks, der Größe der Datenquelle und dem Profiltyp ab. Während das Developer Tool die Datenerkennungs-Tasks weiter

ausführt, können Sie die Ergebnisse der Spaltenprofile anzeigen, die in den ersten Phasen der Datenerkennung abgeschlossen wurden.

1. Klicken Sie nach der Ausführung des Profils auf **Profile** im Fenster **Eigenschaften**.
2. Wählen Sie das Spaltenprofil aus, dessen Ergebnisse Sie anzeigen möchten. Stellen Sie sicher, dass der Status der Profilausführung **Erfolgreich** im Fenster **Eigenschaften** ist.
3. Klicken Sie auf **Öffnen**, um die Ergebnisse auf einer anderen Registerkarte anzuzeigen.
4. Wählen Sie im Bereich **Ergebnisse** die Option **Spalten-Profiling**, um die Ergebnisse im rechten Fensterbereich anzuzeigen.

Anzeigen von Datendomänenenerkennungsergebnissen während der Ausführung der Enterprise-Erkennung

Während das Developer Tool die Datenerkennungs-Tasks, die in der Enterprise-Erkennung enthalten sind, weiter ausführt, können Sie die Ergebnisse der Datendomänenenerkennung anzeigen, die vom Developer Tool in den ersten Phasen der Enterprise-Erkennung abgeschlossen wurden.

1. Klicken Sie auf **Profile** im Fenster **Eigenschaften**, nachdem Sie die Ausführung des Profil begonnen haben.
2. Wählen Sie das Profil aus, dessen Datendomänenenergebnisse Sie anzeigen möchten. Stellen Sie sicher, dass der Status der Profilausführung **Erfolgreich** im Bereich **Eigenschaften** ist.
3. Klicken Sie auf **Öffnen**, um die Ergebnisse auf einer anderen Registerkarte anzuzeigen.
4. Wählen Sie im Bereich **Ergebnisse** die Option **Datendomänenenerkennungs-Profiling**, um die Ergebnisse im rechten Fensterbereich anzuzeigen.

Anzeigen des Laufzeitstatus der Enterprise-Erkennung

Die Ansicht **Fortschritt** im Developer Tool zeigt den Fortschritt der Operationen, wie z. B. einer Profilausführung. Sie können den Laufzeitstatus der Enterprise-Erkennungs-Tasks aus der Ansicht **Fortschritt** heraus anzeigen.

1. Nachdem Sie ein Profil zur Ausführung der Enterprise-Erkennung für die Datenquellen ausgeführt haben, klicken Sie auf die Schaltfläche **Fortschritt anzeigen** unten rechts im Developer-Tool.
Der Bereich **Fortschritt** wird angezeigt, sofern er nicht bereits geöffnet ist.
2. Klicken Sie auf die Verknüpfung **Enterprise-Erkennung wird ausgeführt: Aufgabenstatus anzeigen**, um das Dialogfeld für Unteraufgaben zu öffnen.
In dem Dialogfeld werden mehrere Profil-Tasks aufgelistet, die Bestandteil der Enterprise-Erkennung sind. Sie können den Profilnamen, den Typ und die verknüpften Regeln einsehen.
3. Klicken Sie auf die Spaltenüberschrift, um die Profilaufgaben zu sortieren. Klicken Sie beispielsweise auf die Spaltenüberschrift **Status**, um die Profilaufgaben nach Status zu sortieren.

4. Wenn Sie eine bestimmte Profilaufgabe abbrechen müssen, wählen Sie die Aufgabe aus und klicken Sie auf **Abbrechen**.

Der Status des abgebrochenen Task ändert sich auf **Beendet**.

Enterprise-Erkennungs-Exportdateien

Nach dem Ausführen eines Enterprise-Erkennungsprofils können Sie Informationen, u. a. alle Datenobjektbeziehungen, Datendomänen und einzelne Fremdschlüssel-Aufgabenergebnisse exportieren. Sie können das grafische Bild der Datenobjektbeziehungen als **JPG**-Datei abspeichern.

Beim Exportieren der Profilergebnisse speichert das Developer-Tool alle Enterprise-Erkennungsergebnisse in mehreren Microsoft Excel-Dateien. Sie können Datenobjektbeziehungen, Spaltenprofilergebnisse, Datendomänenerkennungsergebnisse, Entitäten und einzelne Fremdschlüssel-Aufgabenergebnisse in separaten Dateien anzeigen.

Exportieren der Enterprise-Erkennungsergebnisse

Sie können die Liste der Entitäten und die Zusammensetzung jeder Entität, alle Datenobjekte und Datenobjektbeziehungen einer Entität auf Spaltenebene, Datendomänen und Spalten-Profiling-Ergebnisse exportieren.

1. Führen Sie ein Profil aus, um die Enterprise-Erkennung durchzuführen.
2. Klicken Sie in der Ansicht **Beziehungen**, **Datendomänen** oder **Spaltenprofil** im oberen rechten Bereich des Fensters auf das Symbol **Exportieren**.

Das Dialogfenster **Daten in Datei exportieren** wird geöffnet.

3. Geben Sie den Dateinamen ein. Optional können Sie den Standarddateinamen verwenden.
4. Wählen Sie unter **Speichern** die Option **Auf Client speichern** und klicken Sie auf **Durchsuchen**, um einen Speicherort zu wählen. Speichern Sie anschließend die Datei lokal auf Ihrem Computer. Standardmäßig schreibt Informatica Developer die Datei an einen in den Data Integration Service-Eigenschaften von Informatica Administrator festgelegten Speicherort.
5. Klicken Sie auf **OK**.

KAPITEL 27

Business Glossary Desktop in Informatica Developer

Dieses Kapitel umfasst die folgenden Themen:

- [Business Glossary-Suche, 226](#)
- [Suchen einer Geschäftsbedingung, 227](#)
- [Anpassen der Abkürzungstasten zum Nachschlagen einer Geschäftsbedingung, 227](#)

Business Glossary-Suche

Schlagen Sie die Bedeutung eines Developer-Tool-Objektnamens als Geschäftsbegriff im Business Glossary-Desktop nach, um dessen Geschäftsanforderung und aktuelle Implementierung zu verstehen.

Bei einem Business Glossary handelt es sich um eine Reihe von Bedingungen, die Geschäftssprache verwenden, um Konzepte für Geschäftsbenutzer zu definieren. Ein Geschäftsbegriff stellt die Geschäftsdefinition sowie die Verwendung eines Konzepts bereit. Der Business Glossary-Desktop ist ein Client, der mit dem Metadata Manager-Dienst verbunden ist, welcher das Business Glossary beinhaltet. Verwenden Sie den Business Glossary-Desktop, um Geschäftsbedingungen in einem Business Glossary nachzuschlagen.

Wenn der Business Glossary-Desktop auf Ihrem Rechner installiert ist, können Sie ein Objekt im Developer-Tool auswählen und Abkürzungstasten oder das Suchmenü verwenden, um den Namen des Objekts im Business Glossary zu suchen. Sie können nach Namen von Objekten in den Ansichten des Developer Tools, wie z. B. in der Ansicht **Objekt-Explorer**, oder nach Namen von Spalten, Profilen und Umwandlungsports im Editor suchen.

Beispielsweise möchte ein Entwickler einen Geschäftsbegriff in einem Business Glossary suchen, der dem Sales_Audit-Datenobjekt im Developer-Tool entspricht. Der Entwickler möchte die Details der Geschäftsbedingung zum Verständnis der geschäftlichen Anforderungen und der aktuellen Implementierung des Sales_Audit-Objekts im Developer tool ansehen. Dies kann dem Entwickler helfen zu verstehen, was das Datenobjekt bedeutet und welche Änderungen möglicherweise in das Objekt implementiert werden.

Suchen einer Geschäftsbedingung

Suchen Sie einen Developer-Tool-Objektnamen im Unternehmensglossar-Desktop als Geschäftsbedingung, um seine Geschäftsanforderungen und aktuelle Anwendung zu verstehen.

Der Unternehmensglossar Desktop muss auf Ihrem Rechner installiert sein.

1. Wählen Sie ein Objekt aus.
2. Verwenden Sie Abkürzungstasten oder das Menü "Suchen", um den Geschäftsglossar-Desktop zu öffnen.
 - Benutzen Sie die folgende Tastenkombination, um Abkürzungstasten zu verwenden.

CTRL+Shift+F

- Um das Menü "Suchen" zu verwenden, klicken Sie auf **Suche > Unternehmensglossar**.

Der **Unternehmensglossar-Desktop** wird angezeigt; es werden die Geschäftsbedingungen angezeigt, die mit dem Objektnamen übereinstimmen.

Anpassen der Abkürzungstasten zum Nachschlagen einer Geschäftsbedingung

Abkürzungstasten anpassen, um die Tastenkombination zu ändern, die den Unternehmensglossar-Desktop öffnen.

1. Klicken Sie im Developer-Tool-Menü auf **Fenster > Einstellungen > Allgemeines > Tasten**.
2. Um **Suche Geschäftsglossar** in der Liste der Befehle zu suchen oder zu finden, wählen Sie eine der folgenden Möglichkeiten:
 - Um nach den Tasten zu suchen, geben Sie "Suche Unternehmensglossar" im Suchfeld ein.
 - Um nach den Tasten zu scrollen, scrollen Sie, um den Befehl **Suche Unternehmensglossar** in der Spalte **Befehl** zu suchen.
3. Klicken Sie auf den **Befehl Suche Geschäftsglossar**.
4. Klicken Sie auf **Befehl trennen**.
5. Geben Sie im Feld **Bindung** eine Tastenkombination ein.
6. Klicken Sie auf **Anwenden** und anschließend auf **OK**.

Funktionsunterstützung auf Basis der Profiling-Warehouse-Verbindung

- [Unterstützung der Profiling-Funktion, 228](#)

Unterstützung der Profiling-Funktion

Sie können mittels JDBC- oder nativer Konnektivität eine Verbindung zum Profiling-Warehouse herstellen. Sie können basierend auf der Profiling-Warehouse-Verbindung bestimmte Funktionen beim Profiling durchführen.

In der folgenden Tabelle werden die Funktionen aufgeführt, die Sie auf Basis des in Data Engineering Quality und Data Engineering Integration ausgewählten Typs der Profiling-Warehouse-Verbindung durchführen können:

Funktion	JDBC-Verbindung	Native Verbindung
Spaltenprofil eines einzelnen Datenobjekts	Unterstützt	Unterstützt
Datendomänenerkennung eines einzelnen Datenobjekts	Unterstützt	Unterstützt
Einzelnes Datenobjektprofil mit Primärschlüsselerkennung	Nicht unterstützt	Nicht unterstützt
Einzelnes Datenobjektprofil mit funktionaler Abhängigkeitserkennung	Nicht unterstützt	Nicht unterstützt
Profil mit Stichprobe <i>Stichprobe erste <number> Zeilen</i>	Unterstützt	Unterstützt
Profil mit Stichprobe <i>Zufällige Stichprobe <number> Zeilen</i>	Unterstützt	Unterstützt
Scorecard-Metriken und Wertkosten	Nicht unterstützt	Unterstützt

Funktion	JDBC-Verbindung	Native Verbindung
Enterprise-Erkennungsprofil mit Primärschlüssel- und Fremdschlüsselerkennung	Nicht unterstützt	Nicht unterstützt
Enterprise-Erkennungsprofil mit Join-Analyse und Überschneidungserkennung	Nicht unterstützt	Nicht unterstützt
Drilldown für abgerufene Werte und Werthäufigkeiten	Unterstützt	Unterstützt
Profilerggebnisse exportieren	Unterstützt	Unterstützt
Zeilenzählerzuordnung*	Nicht unterstützt	Unterstützt
*Profiling führt die Zeilenzählerzuordnung durch, wenn das Statistikhilfsprogramm die Zeilenanzahl nicht melden kann.		

INDEX

D

Data Discovery

Prozess [19](#)

Datendomäne

Erstellen aus Profilergebnissen in Informatica Analyst [113](#)

Erstellen aus Profilergebnissen in Informatica Developer [186](#)

Erstellen in Informatica Analyst [112](#)

Erstellen in Informatica Developer [185](#)

Suchen in Informatica Developer [187](#)

Übersicht [29](#)

Datendomänen

Exportieren [188](#)

Importieren [187](#)

Suchen in Informatica Analyst [113](#)

Datendomänenerkennung

Informatica Analyst - Übersicht [111](#)

Informatica Developer - Übersicht [184](#)

Prozess [30](#)

Übersicht [28](#)

Datendomänenerkennungsergebnisse

Exportieren aus Informatica Analyst [121](#)

Exportieren aus Informatica Developer [196](#)

Exportieren in Informatica Analyst [121](#)

Informatica Analyst [119](#)

Informatica Developer [193](#)

Datendomänenglossar

Informatica Analyst [111](#)

Informatica Developer [185](#)

Übersicht [30](#)

Datendomänenengruppe

Erstellen in Informatica Analyst [112](#)

Erstellen in Informatica Developer [185](#)

Übersicht [29](#)

Datenerkennung

Übersicht [21](#)

Datenobjekte mit Scorecards

Informatica Analyst [102](#)

Datenobjektprofil

Kommentare [163](#)

Datenobjektprofile

Enterprise-Erkennung [204](#)

Erstellen eines einzelnen Profils [158](#)

Erstellen mehrerer Kundenprofile [159](#)

Übersicht [149](#)

E

Einfachdatei-Datenobjekt

Synchronisieren [45](#)

Enterprise-Erkennung

Anzeigen von Datenobjektbeziehungen [217](#)

bearbeiten [205](#)

Bearbeiten im Analyst Tool [127](#)

Beziehungsansicht [216](#)

Enterprise-Erkennung (Fortsetzung)

Datenvorschauansicht [222](#)

Fremdschlüssel-Profiling-Ansicht [217](#)

In Informatica Analyst ausführen [125](#)

Laufzeitstatus [224](#)

Prozess [198](#)

Spaltenprofilansicht [223](#)

Tabellenansicht [220](#)

Übersicht [197](#)

Enterprise-Erkennung im Analyst Tool

Übersicht [122](#)

Enterprise-Erkennung im Analyst-Tool

Datentypkonflikt [131](#)

Profilansicht [132](#)

Prozess [123](#)

Zusammenfassungsansicht [129](#)

Enterprise-Erkennungsergebnisse

Exportieren [225](#)

Speichern als Symbol [219](#)

Übersicht [215](#)

Enterprise-Erkennungsprofil

DDL-Skripts [212](#)

Erstellen von DDL-Skripts [212](#)

wird ausgeführt [206](#)

Ergebnisse der Enterprise-Erkennung in Informatica Analyst

Übersicht [128](#)

Ergebnisse der Erkennungssuche

Schnittstelle [137](#)

Ergebnisse der Erkennungssuche im Analyst-Tool

Übersicht [136](#)

Ergebnisse zum Datendomänenerkennungprofil

Microsoft Excel [121](#)

Erkennungssuche

Voraussetzungen [134](#)

Erkennungssuche im Analyst-Tool

Prozess [134](#)

Erstellen einer Ausdrucksregel

Regeln [50](#)

Erstellen eines Spaltenprofils

Profile [42](#)

Exportieren

Scorecard-Herkunft in XML [182](#)

F

Filter

Übersicht [53](#)

Fremdschlüssel-Profil

Erkennen [207](#)

Fremdschlüsselerkennung

Übersicht [207](#)

Funktionale Abhängigkeitserkennung

Übersicht [156](#)

G

- Geschäftsbedingungen
 - Anpassen der Abkürzungstasten [227](#)
 - Suchen [227](#)
- Geschäftsbegriff
 - Suchen eines Geschäftsbegriffs [144](#)

I

- Informatica Analyst
 - Regeln [48](#)
 - Spaltenprofile - Übersicht [36](#), [72](#)
 - Spaltenprofilerggebnisse [58](#), [70](#)
 - Verwaltung von Sperren und Versionen [41](#)
- Informatica Developer
 - Profilansichten [147](#)
 - Profilübersicht [146](#)
 - Regeln [171](#)

J

- Join-Analyse
 - Übersicht [208](#)

K

- Konfigurationsoptionen
 - Enterprise-Erkennung im Analyst-Tool [123](#)
- Kumulative Metriken, Bereich
 - Informatica Analyst [103](#)

L

- Laufzeitumgebung
 - Analyst Tool [39](#)
 - Hadoop [39](#), [153](#)

M

- Mapping-Objekt
 - Ausführen eines Profils [173](#)
- Maplet- und Mapping-Profilung
 - Übersicht [173](#)
- Metadata Manager-Geschäftsbegriff
 - Geschäftsbegriffe verwalten [144](#)
 - Projekte [143](#)

O

- Optionen der Datendomänenerkennung
 - Informatica Developer [189](#)

P

- Primärschlüssel-Erkennung
 - Übersicht [154](#)
- Profil
 - Avro- oder Parquet-Formate [166](#)
 - Komponenten [22](#)
 - XML- und JSON-Formate [165](#), [166](#)

Profil-Optionen

- Enterprise-Erkennung [199](#)

Profile

- Ausführen [44](#), [71](#), [72](#), [119](#)
- Bearbeiten eines Filters [56](#)
- Bearbeiten eines Spaltenprofils [43](#)
- Erstellen eines Filters [53](#)
- Erstellen eines Spaltenprofils [42](#)

Profilerggebnisse

- Ablehnen von Datentypen [79](#)
- Ablehnen von Datentypen im Developer Tool [179](#)
- Datendomänen ablehnen [120](#)
- Datendomänen ablehnen in Informatica Developer [195](#)
- Datendomänen genehmigen [120](#)
- Datendomänen in Informatica Developer genehmigen [195](#)
- Detailansicht [62](#)
- Drilldown [77](#)
- Excel [80](#)
- Exportieren [79](#)
- Exportieren aus Informatica Analyst [80](#)
- Exportieren in Informatica Developer [180](#)
- Genehmigen von Datentypen [78](#)
- Genehmigen von Datentypen in Informatica Developer [179](#)
- Geschäftsbegriffe [83](#)
- Hinzufügen von Kommentaren in Informatica Developer [163](#)
- Kommentare [83](#)
- Spaltendatentypen [64](#), [178](#)
- Spaltenmuster [66](#)
- Spaltenwerte [67](#)
- Tags [84](#)
- Wiederherstellen von Spaltenbeziehungen in Informatica Developer [221](#)
- Zusammenfassung [74](#), [77](#)
- Zusammenfassungsansicht [60](#)

Profiling

- Architektur [18](#)
- Übersicht [16](#)
- Verwaltung von Sperren und Versionen [26](#)

Projekte

- Metadata Manager-Geschäftsbegriff [143](#)

R

Regeln

- Anwenden einer vordefinierten Regel [49](#)
- Anwenden in Informatica Developer [172](#)
- Anwenden in PowerCenter Express [172](#)
- Ausdruck [50](#)
- Erstellen einer Ausdrucksregel [50](#)
- Erstellen einer Ausdrucksregel unter Verwendung der Regelspezifikation [52](#)
- Erstellen in Informatica Developer [172](#)
- vordefiniert [49](#)

S

Scorecard

- Konfigurieren von Benachrichtigungen [108](#)
- Konfigurieren von globalen Benachrichtigungseinstellungen [108](#)

Scorecard-Dashboard

- Informatica Analyst [99](#)

Scorecard-Ergebnisse

- Exportieren [105](#)
- Exportieren aus Informatica Analyst [105](#)
- Exportieren nach Excel [106](#)

- Scorecard-Herkunft
 - Anzeigen aus Informatica Developer [182](#)
 - Anzeigen in Informatica Analyst [110](#)
- Scorecards
 - anzeigen [90](#)
 - Ausführen [90](#)
 - bearbeiten [91](#)
 - Bearbeiten einer Metrikgruppe [94](#)
 - Benachrichtigungen [106](#)
 - Definieren von Schwellenwerten [93](#)
 - Drilldown [95](#)
 - Erstellen einer Metrikgruppe [93](#)
 - Feste Kosten [92](#)
 - Hinzufügen von Spalten zu einer Scorecard [89](#)
 - Informatica Analyst [85](#)
 - Informatica Analyst-Prozess [86](#)
 - Informatica Developer [181](#)
 - Kosten für ungültige Daten [92](#)
 - Löschen einer Metrikgruppe [95](#)
 - Metriken [92](#)
 - Metrikgewichtungen [92](#)
 - Metrikgruppen [93](#)
 - Trenddiagramm [95](#)
 - Übersicht [27](#)
 - Variable Kosten [92](#)
 - Verschieben von Scores [94](#)
- Scorecards nach Projekt, Bereich
 - Informatica Analyst [100](#)
- Sonderfall
 - Erkennen [66](#)
- Spaltenprofil
 - Betriebssystemprofil [41](#), [158](#)
 - Drilldown [77](#)
 - Informatica Developer [150](#)
 - Optionen [26](#)
 - Übersicht [25](#)
 - Verarbeiten [37](#)
- Spaltenprofilergebnisse
 - Informatica Developer [176](#)
- Spaltenprofilergebnisse im Analyst Tool
 - Schnittstelle [61](#), [73](#), [75](#)
 - Spaltendetails [61](#), [76](#)
 - Zusammenfassung [59](#)

- Sqoop-Konfiguration
 - Profiling [40](#), [154](#)
- Suche
 - Business Glossary [226](#)

T

- Tabellendatenobjekt
 - Synchronisieren [47](#)
- Trend bei der Scorecard-Ausführung, Bereich
 - Informatica Analyst [101](#)
- Trenddiagramme
 - anzeigen [97](#)
 - Exportieren aus Informatica Analyst [98](#)
 - Kostenbasiert [96](#)
 - Score [96](#)

U

- Überschneidungserkennung
 - durchführen [212](#)
 - Ergebnisse [211](#)
 - Übersicht [210](#)

V

- Vordefinierte Regeln
 - Prozess [49](#)

W

- Wiederherstellung
 - Aufgaben [33](#)
 - Informatica Analyst [78](#)
 - Informatica Developer [179](#)
 - Konzepte [32](#)
 - Prozess [33](#)