

IBM InfoSphere QualityStage Standardization
Rules Designer
Version 11 Release 3

*Lernprogramm: Adressregelsatz in
Standardization Rules Designer
erweitern*



IBM InfoSphere QualityStage Standardization
Rules Designer
Version 11 Release 3

*Lernprogramm: Adressregelsatz in
Standardization Rules Designer
erweitern*



Hinweis

Vor Verwendung dieser Informationen und des darin beschriebenen Produkts sollten die Informationen unter „Bemerkungen und Marken“ auf Seite 43 gelesen werden.

Diese Veröffentlichung ist eine Übersetzung des Handbuchs
IBM InfoSphere QualityStage Standardization Rules Designer, Version 11 Release 3, Tutorial: Enhancing an address rule set in the Standardization Rules Designer,
IBM Form SC19-4334-00,
herausgegeben von International Business Machines Corporation, USA

© Copyright International Business Machines Corporation 2012, 2014

Informationen, die nur für bestimmte Länder Gültigkeit haben und für Deutschland, Österreich und die Schweiz nicht zutreffen, wurden in dieser Veröffentlichung im Originaltext übernommen.

Möglicherweise sind nicht alle in dieser Übersetzung aufgeführten Produkte in Deutschland angekündigt und verfügbar; vor Entscheidungen empfiehlt sich der Kontakt mit der zuständigen IBM Geschäftsstelle.

Änderung des Textes bleibt vorbehalten.

Herausgegeben von:
TSC Germany
Kst. 2877
September 2014

© Copyright IBM Corporation 2012, 2014.

Inhaltsverzeichnis

Lernprogramm: Adressregelsatz in Standardization Rules Designer erweitern . . . 1

Umgebung des Lernprogramms konfigurieren	6
Lernprogrammdateien in ein Verzeichnis kopieren	7
Projekt für Lernprogramm erstellen.	7
Regelsatz für Lernprogramm importieren.	8
Modul 1: Überarbeitung für einen Regelsatz erstellen und Stichprobendaten importieren	9
Lerneinheit 1.1: Regelsätze anzeigen und Überarbeitung in Standardization Rules Designer erstellen.	9
Lerneinheit 1.2: Stichprobendaten importieren	11
Modul 2: Werte klassifizieren	12
Lerneinheit 2.1: Werte durch Hinzufügen von Klassifikationsdefinitionen zu einer vorhandenen Klasse hinzufügen	12
Lerneinheit 2.2: Angepasste Klasse hinzufügen und füllen.	16
Modul 3: Lookup-Tabelle hinzufügen.	20
Lerneinheit 3.1: Lookup-Tabelle hinzufügen	20
Lerneinheit 3.2: Lookup-Tabellendefinitionen importieren und hinzufügen	21
Modul 4: Regeln hinzufügen und modifizieren	24
Lerneinheit 4.1: Regel für die ordnungsgemäße Bearbeitung von Daten modifizieren	25

Lerneinheit 4.2: Unbearbeitete Datensätze ermitteln und Ausgabespalten anpassen	26
Lerneinheit 4.3: Regel für unbearbeitete Datensätze hinzufügen	29
Lerneinheit 4.4: Regel hinzufügen, um Werte auf verschiedene Ausgabespalten aufzuteilen	31
Lerneinheit 4.5: Überarbeitung publizieren	33

Anhang A. Eingabehilfen in den Produkten 35

Anhang B. Kontaktaufnahme mit IBM 37

Anhang C. Auf Produktdokumentation zugreifen 39

Anhang D. Feedback zur Produktdokumentation geben 41

Bemerkungen und Marken 43

Index 49

Lernprogramm: Adressregelsatz in Standardization Rules Designer erweitern

In diesem Lernprogramm werden Sie einen Regelsatz, der Adressdaten standardisiert, mithilfe von IBM® InfoSphere QualityStage Standardization Rules Designer erweitern. Wenn Sie Daten standardisieren, implementieren Sie Datenqualitätsstandards, die Datenwerte normalisieren und Daten für Verwendungen wie Abgleich und Berichterstellung vorbereiten.

In diesem Lernprogramm werden Sie die Daten des fiktiven Unternehmens Sample Outdoor Company verwenden, das Produkte an Einzelhandelsunternehmen und Konsumenten verkauft und verteilt. In den letzten Jahren hat das Unternehmen seinen Aktionsradius kontinuierlich ausgebaut und verkauft seine Produkte jetzt an Einzelhändler in fast allen Teilen der Welt.

Das fiktive Unternehmen Sample Outdoor Company hat vor Kurzem mehrere Firmen aufgekauft. Im Rahmen der Firmenübernahme hat das fiktive Unternehmen Sample Outdoor Company Kundendaten von allen Unternehmen erhalten. Das fiktive Unternehmen Sample Outdoor Company möchte die Adressdaten dieser neuen Kunden in seine aktuelle Datenbank integrieren, die neuen Daten enthalten jedoch neue Datentypen und sind inkonsistent formatiert. Sample Outdoor Company kann den Regelsatz, der diesen Datentyp standardisiert, mithilfe von IBM InfoSphere QualityStage Standardization Rules Designer erweitern.

Nachdem der Regelsatz erweitert wurde, wendet das Unternehmen den Regelsatz in einem Standardisierungsjob auf eine Stage **Standardize** (Standardisieren) an. Wenn der Standardisierungsjob ausgeführt wird, werden die Eingabedaten nach der Logik standardisiert, die im erweiterten Regelsatz angegeben ist.

Dieses Lernprogramm führt Sie durch einige der allgemeinen Tasks, die Sie möglicherweise ausführen, wenn Sie einen Regelsatz für Adressdaten in Standardization Rules Designer erweitern. Die folgenden Schritte veranschaulichen die Reihenfolge der Aktionen des Lernprogramms:

1. In Modul 1 öffnen Sie eine Überarbeitung für den Regelsatz, für den die Erweiterung in Standardization Rules Designer erforderlich ist. Außerdem importieren Sie Stichprobendaten für die Verwendung in Standardization Rules Designer.
2. In Modul 2 kategorisieren Sie die Teile der Daten. Sie fügen Klassifikationsdefinitionen hinzu, die vorhandenen Klassen neue Werte zuweisen, und für einen neuen Datentyp eine benutzerdefinierte Klasse hinzufügen. Abbildung 1 zeigt, wie jeder Wert in einem Stichprobendatensatz für eine Adresse des fiktiven Unternehmens Sample Outdoor Company einer Klasse zugeordnet werden kann.

Eingabedatensatz

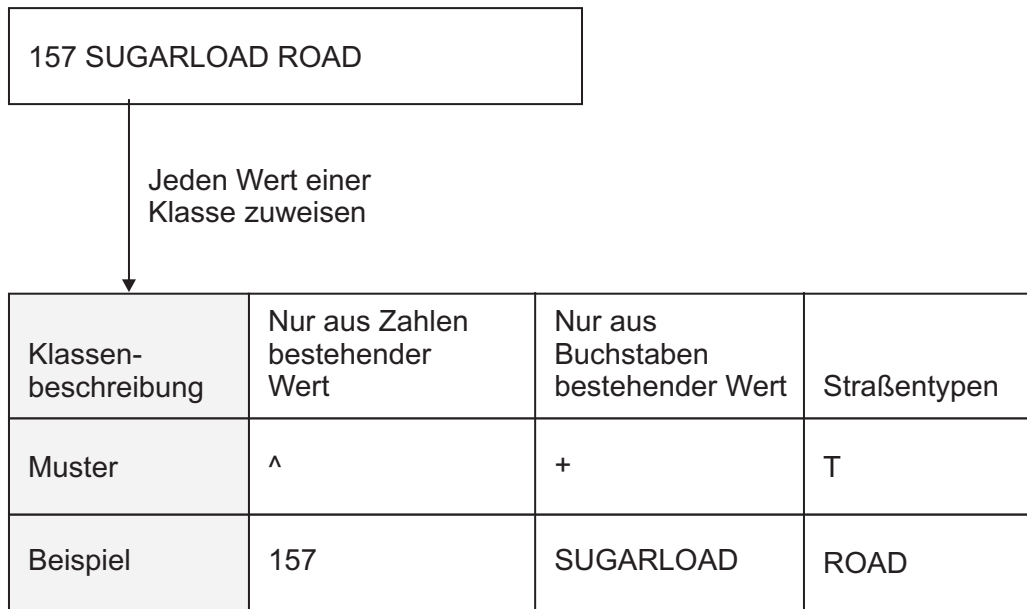


Abbildung 1. Jeder Wert eines Stichprobendatensatzes wird einer Klasse zugeordnet.

- In Modul 3 fügen Sie eine Lookup-Tabelle hinzu, die abgekürzte Straßennamen in vollständige Straßennamen konvertiert. Abbildung 2 zeigt einen Teil der Lookup-Tabelle, die in Standardization Rules Designer hinzugefügt wird.

Wert (abgekürzter Straßename)	Zurückgegebener Wert (vollständiger Straßename)
MLK	MARTIN LUTHER KING
BWAY	BROADWAY
SM	STUART MILL

Abbildung 2. Die Lookup-Tabelle konvertiert abgekürzte Straßennamen in vollständige Straßennamen.

- In der ersten Lerneinheit von Modul 4 modifizieren Sie eine Regel, die in Standardization Rules Designer hinzugefügt wurde. Die Regel bearbeitet die Daten für einige der neuen Adressen nicht ordnungsgemäß. Abbildung 3 zeigt die Ausgabe mit der aktuellen Regel und die Ausgabe nach der Modifizierung der Regel gemäß der Datenbereinigungsanforderungen des fiktiven Unternehmens Sample Outdoor Company.

Eingabedatensatz

^	+	T
588	SM	ST

Ausgabe mit aktueller Regel

House number	Street name	Street suffix type
588	SM	ST

Abgekürzten Straßennamen mithilfe von Lookup-Tabelle in vollständigen Straßennamen konvertieren

Wert (abgekürzter Straßename)	Zurückgegebener Wert (vollständiger Straßename)
MLK	MARTIN LUTHER KING
BWAY	BROADWAY
SM	STUART MILL

Ausgabe mit modifizierter Regel

House number	Street name	Street suffix type
588	STUART MILL	ST

Abbildung 3. Die aktuelle Regel entspricht nicht den Datenbereinigungsanforderungen des fiktiven Unternehmens Sample Outdoor Company. Die modifizierte Regel verwendet eine Lookup-Tabelle, um abgekürzte Straßennamen in vollständige Straßennamen zu konvertieren.

- In der zweiten und dritten Lerneinheit von Modul 4 stellen Sie ein unbearbeitetes Muster fest und fügen eine Regel hinzu, die Daten bearbeitet, die mit diesem Muster übereinstimmen. Abbildung 4 zeigt, wie eine Beispieladresse von der neuen Regel bearbeitet wird.

Eingabedatensatz für unbearbeitbares Muster

^	>	T
243	45E	AVE

Ausgabe

House number	Street name	Street suffix type
243	45E	AVE

Abbildung 4. Eine neue Regel für unbearbeitete Muster fügt den entsprechenden Ausgabespalten Adresswerte hinzu.

6. In der vierten Lerneinheit von Modul 4 fügen Sie eine Regel hinzu, die zwei unterschiedliche Werte bearbeitet, die in den Eingabedaten miteinander verknüpft sind. Wenn die Eingabedaten z. B. den Wert 673HEGENBERGER enthalten, können Sie eine Regel erstellen, die den Wert in die Werte 673 und HEGENBERGER aufteilt und in die entsprechenden Ausgabespalten stellt. Abbildung 5 zeigt, wie ein Beispieldatensatz für eine Adresse des fiktiven Unternehmens Sample Outdoor Company von dieser Regel bearbeitet wird.

Eingabedatensatz für verknüpfte Werte

>	T
673HEGENBERGER	ROAD

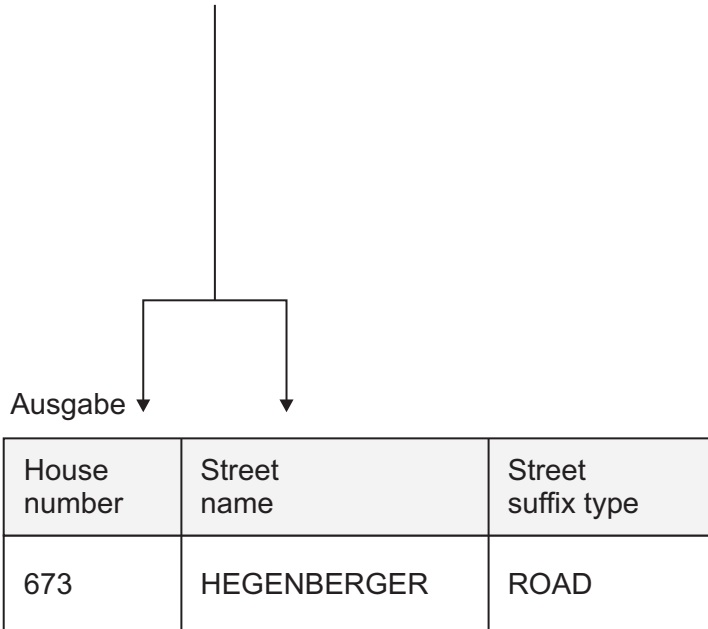


Abbildung 5. Die neue Regel teilt einen verknüpften Wert in zwei unterschiedliche Werte auf und ordnet diese in den entsprechenden Ausgabespalten an. Außerdem fügt die Regel den Straßentypenwert der Ausgabespalte für Straßentypen zu.

Lernziele

Während Sie die Module durcharbeiten, lernen Sie die Konzepte und Tasks zum Erweitern von Regelsätzen kennen:

- Importieren von Stichprobendaten, um zu sehen, wie sich Änderungen an Teilen des Regelsatzes auf Datensätze aus den Stichprobendaten auswirken
- Verwendung von Klassifikationen, um Teile der Daten zu kategorisieren
- Hinzufügen von Lookup-Tabellen, um Daten zu vergleichen oder in angegebene Werte zu konvertieren
- Hinzufügen von Regeln, die Aktionen auf eine Gruppe zusammengehöriger Datensätze anwenden

Erforderlicher Zeitaufwand

Bevor Sie mit dem Lernprogramm beginnen, müssen Sie Ihre Umgebung einrichten. Die für das Einrichten erforderliche Zeit hängt von Ihrer aktuellen Umgebung ab.

Für die Durchführung der einzelnen Module sind jeweils 20 - 60 Minuten erforderlich.

Systemvoraussetzungen

Die folgenden Komponenten und Anwendungen müssen auf Ihrem System installiert sein.

- IBM InfoSphere QualityStage Standardization Rules Designer
- IBM InfoSphere QualityStage mit den folgenden Clients:
 - IBM InfoSphere DataStage und QualityStage Designer
 - IBM InfoSphere DataStage und QualityStage Administrator

Voraussetzungen

Bevor Sie mit diesem Lernprogramm beginnen, sollten Sie mit Konzepten der Datenqualität wie Standardisierung, Klassifikation und Regeln für die Datenbereinigung vertraut sein. Kenntnisse der Konzepte von InfoSphere DataStage und QualityStage wie Jobs, Stages und Berichte können hilfreich sein, sind jedoch nicht erforderlich.

Bemerkungen: Sample Outdoor Company, GP Sales, alle Variationen des Unternehmensnamens sowie das Planungsbeispiel stellen fiktive Geschäftsvorgänge mit Beispieldaten dar, mit denen Beispieldaten für IBM und IBM Kunden erstellt wurden. Zu diesen fiktiven Datensätzen gehören Beispieldaten für Verkaufstransaktionen, Produktvertrieb, Finanzwesen und Personalwesen. Ähnlichkeiten mit tatsächlichen Namen, Adressen, Kontaktdaten oder Transaktionswerten sind rein zufällig. Andere Beispieldateien können Daten folgender Art enthalten: manuell oder vom System generierte fiktive Daten, aus wissenschaftlichen oder öffentlichen Quellen zusammengestellte Fakten sowie Daten, die mit Zustimmung der Copyrightinhaber als Beispieldaten zur Entwicklung von Beispieldaten genutzt werden dürfen. Referenzierte Produktnamen können Marken der jeweiligen Rechtsinhaber sein. Nicht autorisiertes Kopieren dieser Daten ist unzulässig.

Umgebung des Lernprogramms konfigurieren

Bevor Sie mit diesem Lernprogramm beginnen, müssen Sie Ihr System für die Ausführung des Lernprogramms vorbereiten.

Vorbereitende Schritte

- Stellen Sie sicher, dass die Serverkomponente von IBM InfoSphere QualityStage und die Clientkomponente von IBM InfoSphere DataStage und QualityStage Administrator ordnungsgemäß installiert und konfiguriert sind.
- Stellen Sie sicher, dass Benutzerberechtigungen den Berechtigungsnachweisen des Betriebssystembenutzers der Engineschicht zugeordnet sind, und dass Ihr Benutzeraccount über die folgenden Gruppen von Sicherheitsrollen verfügt:
 - Die Rollen Suiteadministrator sowie DataStage- und QualityStage-Administrator sind erforderlich, um ein Projekt zu erstellen. Wenn Sie nicht über diese Rollen verfügen, muss ein Benutzer mit diesen Rollen ein Projekt für das Lernprogramm erstellen und Ihnen für das Projekt die Rollen DataStage- und QualityStage-Entwickler oder DataStage- und QualityStage-Produktionsmanager zuweisen.
 - Die Rollen Suitebenutzer sowie DataStage- und QualityStage-Administrator oder DataStage- und QualityStage-Benutzer sind für die Anmeldung an Standardization Rules Designer erforderlich.

Wenn Sie sich nicht an Standardization Rules Designer anmelden können, ziehen Sie Ihren Systemadministrator zu Rate.

Informationen zu diesem Vorgang

IBM InfoSphere QualityStage Standardization Rules Designer ist eine webbasierte Anwendung, die mit den Client-, Engine- und Metadatenrepository-Schichten interagiert. Bevor Sie mit dem Lernprogramm beginnen, müssen Sie sicherstellen, dass die Stichprobendaten und Regelsätze für das Lernprogramm für Standardization Rules Designer verfügbar sind. Außerdem brauchen Sie ein Projekt, das Sie für das Lernprogramm verwenden können.

Lernprogrammdateien in ein Verzeichnis kopieren

Sie kopieren die Lernprogrammdateien in ein Verzeichnis, das Sie verwenden, während Sie mit dem Lernprogramm arbeiten. Damit Sie das Lernprogramm durcharbeiten können, müssen Sie auf die Lernprogrammdateien auf dem Computer zugreifen können, auf dem Sie Standardization Rules Designer verwenden.

Vorgehensweise

1. Suchen Sie die Datei `Standardization_Rules_Designer_address_tutorial.zip`, die sich auf den Installationsmedien befindet. Die Datei `Standardization_Rules_Designer_address_tutorial.zip` befindet sich im Verzeichnis `übergeordnetes_Verzeichnis\TutorialData\QualityStage` des Verzeichnisses mit den Installationsmedien. Die Datei `Standardization_Rules_Designer_address_tutorial.zip` kann sich beispielsweise im Verzeichnis `is-client\TutorialData\QualityStage` befinden.
2. Extrahieren Sie die Dateien aus der Datei `Standardization_Rules_Designer_address_tutorial.zip` in ein Verzeichnis, das sich auf dem Computer befindet, den Sie für die Verbindung mit Standardization Rules Designer verwenden werden.
3. Wenn sich das Verzeichnis, in dem Sie die Dateien extrahiert haben, nicht auf einem Computer befindet, auf dem die Client- oder Engine-Schicht installiert ist, kopieren Sie die Datei `Address_Rules_for_Tutorial.isx` auf einen Computer mit Client- oder Engineschicht.

Projekt für Lernprogramm erstellen

Sie erstellen ein neues Projekt für das Lernprogramm, damit die Übungen des Lernprogramms von den anderen Arbeiten in InfoSphere QualityStage getrennt sind.

Informationen zu diesem Vorgang

Wenn Sie ein Projekt für das Lernprogramm zum Erweitern eines Regelsatzes für die Produktdomäne erstellt haben, können Sie dasselbe Projekt auch für dieses Lernprogramm verwenden.

Vorgehensweise

1. Klicken Sie zum Starten des Clients von IBM InfoSphere DataStage und QualityStage Administrator auf **Start > Alle Programme > IBM InfoSphere Information Server > IBM InfoSphere DataStage und QualityStage Administrator**.
2. Geben Sie im Fenster **Verbindung zu DataStage herstellen** Ihren Benutzernamen und das Kennwort ein und klicken Sie dann auf **Anmelden**.
3. Klicken Sie auf der Seite **Projekte** auf **Hinzufügen**.
4. Geben Sie im Fenster **Projekt hinzufügen** `StandardizationRulesDesignerTutorial` in das Feld **Name** ein und klicken Sie dann auf **OK**.

5. Wenn andere Benutzer das Lernprogramm durcharbeiten, weisen Sie den Benutzern eine entsprechende Rolle für das Projekt zu.
 - a. Wählen Sie das Projekt **StandardizationRulesDesignerTutorial** aus und klicken Sie dann auf **Eigenschaften**.
 - b. Klicken Sie auf die Registerkarte **Berechtigungen**.
 - c. Wenn der Benutzer, für den Sie Rollen hinzufügen wollen, nicht in der Liste angezeigt wird, fügen Sie den Benutzer dem Projekt hinzu.
 - d. Wählen Sie den Benutzer, der das Lernprogramm durcharbeiten wird, in der Liste der Projektbenutzer aus.
 - e. Wählen Sie **DataStage- und QualityStage-Entwickler** oder **DataStage- und QualityStage-Produktionsmanager** in der Liste **Benutzerrollen** aus und klicken Sie dann auf **OK**.
6. Klicken Sie im Fenster **InfoSphere-DataStage-Verwaltung** auf **Schließen**.

Regelsatz für Lernprogramm importieren

Mithilfe von InfoSphere Information Server Manager können Sie den Regelsatz für das Lernprogramm in das Projekt importieren, das Sie für das Lernprogramm erstellt haben.

Vorbereitende Schritte

- „Lernprogrammdateien in ein Verzeichnis kopieren“ auf Seite 7
- „Projekt für Lernprogramm erstellen“ auf Seite 7

Vorgehensweise

1. Klicken Sie auf **Start > Programme > IBM InfoSphere Information Server > IBM InfoSphere Information Server Manager**.
2. Stellen Sie eine Verbindung zu dem Metadatenrepository her, das das Lernprogrammprojekt enthält:
 - a. Klicken Sie im Anwendungsfenster **Information Server Manager** mit der rechten Maustaste in die Ansicht **Repository** und klicken Sie anschließend auf **Add Domain**.
 - b. Geben Sie den Anwendungsserver an sowie den Namen und das Kennwort eines InfoSphere Information Server-Benutzers mit Administratorberechtigung.
 - c. Klicken Sie auf **OK**.
3. Klicken Sie in der Navigationsstruktur **Repository** mit der rechten Maustaste auf das Projekt **StandardizationRulesDesignerTutorial** und klicken Sie dann auf **Importieren**.
4. Importieren Sie die Datei `Address_Rules_for_Tutorial.isx`:
 - a. Wechseln Sie zu dem Verzeichnis, das die Datei enthält. Die Datei befindet sich standardmäßig in dem Verzeichnis `Standardization_Rules_Designer_address_tutorial`, in das Sie die Lernprogrammdateien kopiert haben.
 - b. Wählen Sie die Datei aus und klicken Sie dann auf **Öffnen**.
 - c. Klicken Sie auf **Importieren**.
 - d. Stellen Sie sicher, dass das Projekt **StandardizationRulesDesignerTutorial** ausgewählt ist und klicken Sie dann auf **OK**.

Modul 1: Überarbeitung für einen Regelsatz erstellen und Stichprobendaten importieren

In diesem Modul erstellen Sie in Standardization Rules Designer eine Überarbeitung für einen Standardisierungsregelsatz, der erweitert werden muss. Danach importieren Sie Stichprobendaten für die Verwendung in Standardization Rules Designer.

Lernziele

Nachdem Sie die Lerneinheiten dieses Moduls durchgearbeitet haben, verstehen Sie die für Regelsätze und Stichprobendaten relevanten Konzepte und Tasks:

- Erstellen einer Überarbeitung für einen Regelsatz in Standardization Rules Designer
- Importieren von Stichprobendaten

Erforderlicher Zeitaufwand

Für dieses Modul sind ungefähr 20 Minuten erforderlich.

Voraussetzungen

Stellen Sie sicher, dass Sie alle Schritte der Einrichtung vollständig ausgeführt haben, und dass die Lernprogrammdateien und Daten auf Ihr Computersystem geladen wurden.

Lerneinheit 1.1: Regelsätze anzeigen und Überarbeitung in Standardization Rules Designer erstellen

In dieser Lerneinheit finden Sie den Regelsatz USADDRTutorial.SET in Standardization Rules Designer und erstellen eine Überarbeitung für diesen Regelsatz. Wenn Sie eine Überarbeitung erstellen, erweitern Sie den Regelsatz durch mehrere Iterationen. Dabei arbeiten Sie zu verschiedenen Zeiten und mit anderen Benutzern zusammen.

Übersicht

Weil der Regelsatz USADDRTutorial.SET die Datenqualitätsanforderungen des fiktiven Unternehmens Sample Outdoor Company nicht erfüllt, muss er erweitert werden.

Weitere Informationen zu Regelsätzen:

Regelsätze überprüfen und normalisieren Eingabedaten. Mithilfe von Regeldatensätzen können Sie Inhalt und Struktur von Quelldaten verstehen, Daten standardisieren und für den Abgleich vorbereiten. Regelsätze enthalten die folgenden Elemente, die im Laufe des Lernprogramms genauer erläutert werden, wenn Sie damit arbeiten:

- Klassifikationen
- Lookup-Tabellen
- Regeln
- Ausgabespalten

Regelsätze werden in Jobs verwendet, die eine Stage **Investigate** (Untersuchung) oder eine Stage **Standardize** (Standardisieren) einschließen.

Bevor Sie den Regelsatz erweitern können, müssen Sie eine Überarbeitung für den Regelsatz in Standardization Rules Designer erstellen.

Weitere Informationen zu Überarbeitungen:

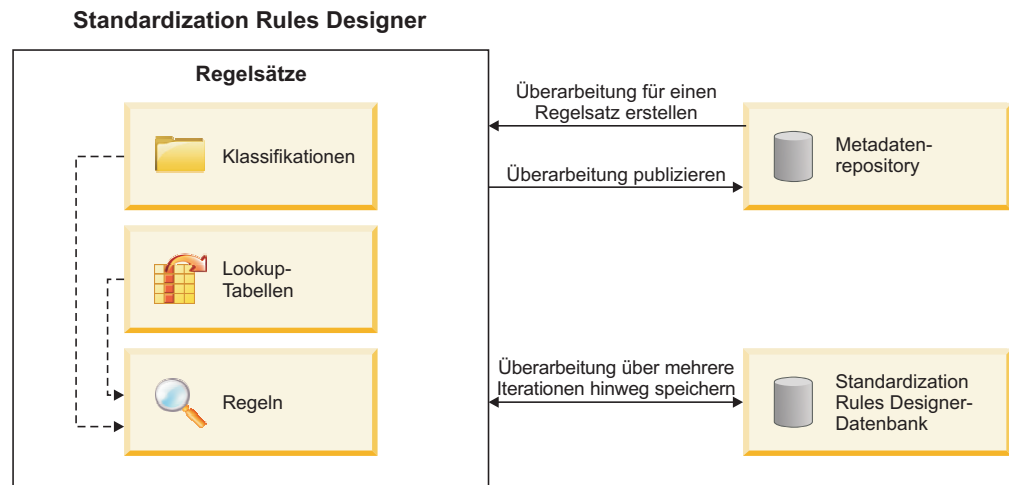
Eine Überarbeitung ist eine Zusammenstellung von Änderungen, die in Standardization Rules Designer an dem Regelsatz vorgenommen wurden. Für jeden Regelsatz kann eine Überarbeitung erstellt werden.

Wenn Sie einen Regelsatz in Standardization Rules Designer erweitern, arbeiten Sie mit einer Kopie des Regelsatzes, die in einer Datenbank für Standardization Rules Designer gespeichert wird. Sie können die Überarbeitung beliebig lange offen lassen und der Regelsatz kann von anderen Benutzern angezeigt und modifiziert werden. Wenn eine Überarbeitung offen ist, können Sie den Regelsatz gemeinsam über mehrere Iterationen erweitern, die sich nur auf die Kopie des Regelsatzes auswirken. Bei der Zusammenarbeit mit mehreren Benutzer kann jeweils immer nur ein Benutzer die Überarbeitung öffnen.

Wenn Sie den tatsächlichen Regelsatz aktualisieren wollen, können Sie die Überarbeitung publizieren. Beim Publizieren der Überarbeitung wird der im Metadatenrepository befindliche Regelsatz mit den in Standardization Rules Designer vorgenommenen Änderungen aktualisiert.

Wenn Sie die aktuellen Änderungen in der Überarbeitung verwerfen wollen, die Überarbeitung jedoch offen bleiben soll, können Sie die Überarbeitung zurücksetzen. Beim Zurücksetzen der Überarbeitung wird die in der Datenbank von Standardization Rules Designer befindliche Version des Regelsatzes durch die im Metadatenrepository befindliche Version ersetzt.

Das folgende Diagramm zeigt, wie Sie während des Lebenszyklus eines Regelsatzes in Standardization Rules Designer mit Überarbeitungen arbeiten können.



Vorgehensweise

1. Wenn Sie Standardization Rules Designer in Ihrem Browser öffnen wollen, wechseln Sie zu https://Hostname:sichere_Portnummer/ibm/iis/qs/StandardizationRulesDesigner/, wobei *Hostname* den Hostnamen des Servers angibt, auf dem Standardization Rules Designer installiert ist.
2. Melden Sie sich an Standardization Rules Designer an. Eine Liste der Server, auf denen sich IBM InfoSphere QualityStage-Projekte befinden, wird angezeigt.

3. Erweitern Sie das Pfeilsymbol des Servers, auf dem Sie das Projekt StandardizationRulesDesignerTutorial erstellt haben.
4. Erweitern Sie **StandardizationRulesDesignerTutorial**.
5. Wählen Sie den Regelsatz USADDRTutorial.SET aus und klicken Sie dann auf **Bearbeiten**.
6. Klicken Sie im Fenster für die Nachricht zur Bearbeitung des Regelsatzes auf **OK**. Eine Überarbeitung für den Regelsatz USADDRTutorial.SET wird erstellt und die Seite **Ausgangsposition** wird für den Regelsatz angezeigt.

Lerneinheit 1.2: Stichprobendaten importieren

In dieser Lerneinheit importieren Sie Stichprobendaten für die Verwendung in Standardization Rules Designer.

Übersicht

Die von Ihnen verwendeten Stichprobendaten geben Aufschluss über Änderungen, die Sie an einem Regelsatz in Standardization Rules Designer vornehmen. Wenn Sie einen Regelsatz in Standardization Rules Designer erweitern, können Sie sehen, wie sich die Änderungen auf die Bearbeitung der Werte und Datensätze in Ihren Stichprobendaten auswirken. Damit Sie mit Ihren Daten arbeiten können, müssen Sie sie in Standardization Rules Designer importieren.

Das fiktive Unternehmen Sample Outdoor Company hat ein Stichproben-Dataset für die Verwendung in Standardization Rules Designer vorbereitet. Die Stichprobendaten enthalten Kundenstammdaten von den Unternehmen, die das fiktive Unternehmen Sample Outdoor Company aufgekauft hat.

Vorgehensweise

1. Klicken Sie in Standardization Rules Designer auf die Registerkarte **Ausgangsposition**.
2. Klicken Sie im Navigationssteifenster auf **Stichprobendaten importieren**.
3. Klicken Sie neben dem Feld **Quelle auswählen** auf **Importieren**.
4. Importieren Sie die Datei `address_sample_data.csv`:
 - a. Wechseln Sie zu dem Verzeichnis, das die Datei enthält. Die Datei befindet sich standardmäßig in dem Verzeichnis `Standardization_Rules_Designer_address_tutorial`, in das Sie die Lernprogrammdateien kopiert haben.
 - b. Wählen Sie die Datei aus und klicken Sie dann auf **Öffnen**.

Standardization Rules Designer zeigt die unbearbeiteten Datensätze an, die importiert wurden, und Stichprobendatensätze, die mit Trennzeichen und Abschlusszeichen verarbeitet wurden.

Wie geht es weiter

In diesem Modul haben Sie die folgenden Tasks ausgeführt:

- Erstellen einer Überarbeitung für einen Regelsatz in Standardization Rules Designer
- Importieren von Stichprobendaten

Sie können sich von Standardization Rules Designer abmelden und das nächste Modul später durcharbeiten. Die Überarbeitung bleibt geöffnet.

Im nächsten Modul werden Sie die Kontextinformationen, die von den neuen Adressdaten bereitgestellt werden, stärken, indem Sie Teile der Daten kategorisieren.

Modul 2: Werte klassifizieren

In diesem Modul stärken Sie die Kontextinformationen, die von den neuen Daten bereitgestellt werden, indem Sie Teile der Daten kategorisieren. Sie werden Klassifikationsdefinitionen hinzufügen, die vorhandenen Klassen neue Werte zuweisen, und eine benutzerdefinierte Klasse hinzufügen, um die Daten zu kategorisieren.

Die Daten für die neuen Kundenadressen, die das fiktive Unternehmen Sample Outdoor Company erworben hat, enthalten die folgenden Werttypen:

- Neue Werte, die einer vorhandenen Klasse zugewiesen werden müssen
- Ein neuer Werttyp, für den das Unternehmen eine Klasse erstellen möchte

Weil einige der Gebäudetypen in den neuen Kundenadressdaten nicht der Gebäudetypenklasse zugewiesen sind, werden die neuen Daten von den Regeln, die Muster bearbeiten, die die Gebäudetypenklasse enthalten, nicht bearbeitet. Im ersten Teil des Moduls ermitteln Sie, wie die Gebäudetypen in den neuen Adressen zurzeit klassifiziert werden. Damit Sie sicher sein können, dass die Gebäudetypen von den entsprechenden vorhandenen Regeln bearbeitet werden, fügen Sie Klassifikationsdefinitionen hinzu, die der Gebäudetypklasse neue Gebäudetypen zuweisen.

Darüber hinaus enthalten einige der Adressen, die das fiktive Unternehmen Sample Outdoor Company kürzlich erworben hat, Informationen, die nicht Teil der Adresse sind, beispielsweise INVALID ADDRESS (Ungültige Adresse) oder BEWARE OF DOG (Warnung vor dem Hund). Damit diese Werte kategorisiert werden, fügen Sie eine benutzerdefinierte Klasse für Ausnahmedaten hinzu. Nachdem Sie die Klasse hinzugefügt haben, können Sie Regeln schreiben, die Ausnahmedaten ordnungsgemäß bearbeiten.

Lernziele

Nachdem Sie die Lerneinheiten dieses Moduls durchgearbeitet haben, verstehen Sie die Konzepte und Tasks zum Klassifizieren von Werten:

- Ermitteln der Klasse, zu der ein Wert gehört
- Hinzufügen von Klassifikationsdefinitionen
- Hinzufügen einer angepassten Klasse

Erforderlicher Zeitaufwand

Für dieses Modul sind ungefähr 30 Minuten erforderlich.

Voraussetzungen

Stellen Sie sicher, dass Sie die Schritte zum Einrichten des Lernprogramms ausgeführt und die Lernprogrammdateien und Stichprobendaten auf Ihren Computer geladen haben.

Lerneinheit 2.1: Werte durch Hinzufügen von Klassifikationsdefinitionen zu einer vorhandenen Klasse hinzufügen

In dieser Lerneinheit geben Sie an, wie Gebäudetypen in neuen Adressdaten zurzeit klassifiziert werden. Danach weisen Sie einer vorhandenen benutzerdefinierten Klasse Werte für Gebäudetypen zu, indem Sie Klassifikationsdefinitionen hinzufügen.

Übersicht

Das fiktive Unternehmen Sample Outdoor Company möchte sicherstellen, dass Datensätze, die die neuen Gebäudetypen enthalten, ordnungsgemäß bearbeitet werden. Sie können die Klasse anzeigen, zu der die neuen Gebäudetypwerte gehören, und anschließend diese Informationen verwenden, um zu sehen, welche Regeln, sofern vorhanden, die Werte bearbeiten. Danach können Sie einer Klasse Gebäudetypwerte zuweisen, indem Sie Klassifikationsdefinitionen hinzufügen.

Weitere Informationen zu Klassifikationen:

In IBM InfoSphere QualityStage werden Datensätze als Muster dargestellt. Ähnlich wie Datensätze, die aus mindestens einem Wert bestehen, bestehen Muster aus mindestens einem abstrakten Zeichen, das eine Klasse darstellt. Klassen geben an, dass der zugrundeliegende Werte zu einer bestimmten Kategorie gehört und erweitern so die von Mustern bereitgestellten Kontextinformationen. So hat das fiktive Unternehmen Sample Outdoor Company z. B. Datensätze wie 157 SUGARLOAD ROAD, die durch das Muster $\wedge+T$ dargestellt werden. Die folgende Tabelle zeigt die Kontextinformationen an, die jede Klasse des Musters $\wedge+T$ bereitstellt.

Tabelle 1. Jede aus einem Zeichen bestehende Klassenbezeichnung gibt an, dass der zugrundeliegende Wert zu einer bestimmten Kategorie gehört

\wedge	+	T
Nur aus Zahlen bestehender Wert	Nur aus Buchstaben bestehender Wert	Straßentypen

Informationen zu Werten werden durch die folgenden Klassentypen bereitgestellt:

- Standardklassen stellen Basisinformationen zum Typ des Werts bereit, beispielsweise, ob der Wert alphabetische Zeichen, numerische Zeichen oder eine Kombination aus beiden enthält. Standardklassen werden durch eine Reihe von Zeichen dargestellt, die in Standardization Rules Designer aufgelistet sind.
- Benutzerdefinierte Klassen stellen stärkere Kontextinformationen zum Typ des Werts bereit, beispielsweise ob ein alphabetischer Wert der Name eines Straßentyps oder eines Gebäudetyps ist. Die aus einem Zeichen bestehende Bezeichnung für angepasste Klassen können Buchstaben aus dem lateinischen Alphabet oder die Ziffer 0 sein. Letzteres gibt eine Nullklasse an.

Tasks in dieser Lerneinheit

1. „Klasse ermitteln, zu der ein Wert gehört“
2. „Einer Klasse Werte zuweisen“ auf Seite 15

Klasse ermitteln, zu der ein Wert gehört


Sie können die Standardklasse oder benutzerdefinierte Klasse ermitteln, der ein Gebäudetypwert zurzeit zugewiesen ist. Wenn die Gebäudetypwerte nicht der benutzerdefinierten Klasse **L Building Types** (Gebäudetypen) zugeordnet sind, werden Datensätze, die diese Werte enthalten, möglicherweise nicht von den entsprechenden Regeln bearbeitet.

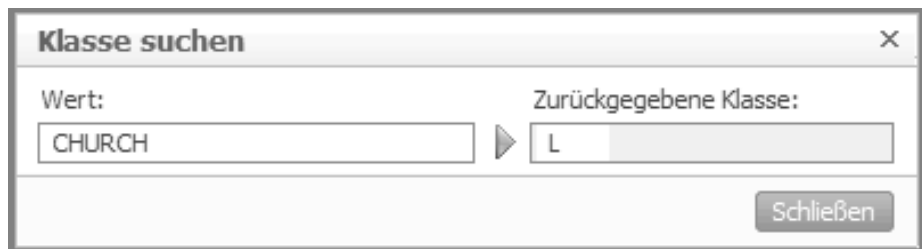
Zu den Kundenadressdaten, die das fiktive Unternehmen Sample Outdoor Company erworben hat, gehören die folgenden Gebäudetypen:

- CHURCH
- INSTITUTE
- HOSTEL

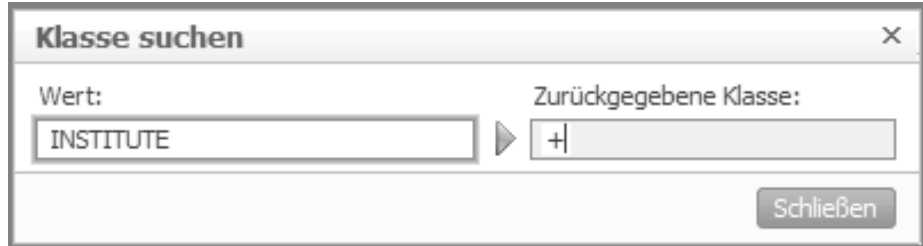
Um sicherzustellen, dass Datensätze, die diese Gebäudetypen enthalten, ordnungsgemäß bearbeitet werden, müssen Sie zunächst angeben, zu welcher Klasse die Werte gehören. Sie können auch anzeigen, welche Muster die Klasse **L Building Types** enthalten und Regeln anzeigen, die Datensätze mit diesem Muster bearbeiten.

Vorgehensweise

1. Klicken Sie in Standardization Rules Designer auf die Registerkarte **Klassifikationen**.
2. Klicken Sie auf **Klasse suchen**.
3. Ermitteln Sie die Klasse, zu der die Gebäudetyppwerte gehören:
 - a. Geben Sie den Wert CHURCH in das Feld **Wert** ein und klicken Sie dann auf . Das Feld **Zurückgegebene Klasse** zeigt an, dass dieser Wert der angepassten Klasse L für Gebäudetypen zugewiesen ist.



The screenshot shows a dialog box titled "Klasse suchen" with a close button (X) in the top right corner. It contains two input fields: "Wert" with the text "CHURCH" and "Zurückgegebene Klasse" with the text "L". A right-pointing arrow is between the two fields. A "Schließen" button is located at the bottom right of the dialog.

- b. Wiederholen Sie Schritt 3a für die Werte INSTITUTE und HOSTEL. Das Feld **Zurückgegebene Klasse** zeigt, dass diese Werte der Standardklasse + zugewiesen sind.
- 
- The screenshot shows a dialog box titled "Klasse suchen" with a close button (X) in the top right corner. It contains two input fields: "Wert" with the text "INSTITUTE" and "Zurückgegebene Klasse" with the text "+". A right-pointing arrow is between the two fields. A "Schließen" button is located at the bottom right of the dialog.
- c. Klicken Sie auf **Schließen**.
4. Öffnen Sie eine Regelgruppe, um auf die Muster in den Daten und auf die Regeln für diese Muster zuzugreifen:
 - a. Klicken Sie auf die Registerkarte **Regeln**.
 - b. Wählen Sie die Regelgruppe **Input_Overrides** aus und klicken Sie dann auf **Öffnen**.
 5. Blenden Sie **Musterregel** ein. Eine Liste der in den Daten enthaltenen Muster wird angezeigt.

▼ Musterregel		Muster hinzufügen...	
Regel	Häufigkeit	Unbearbeitet	
▶ ^+T	42.73%	0%	
▶ ^D+T	5.34%	5.34%	
▶ ^T+	4.59%	4.59%	
▶ ^++T	3.65%	3.65%	
▶ ^TT	3.37%	3.37%	

- Erweitern Sie **+L^+T** und erweitern Sie anschließend **Copy building details to building name output column**. Eine Liste der von dieser Regel bearbeiteten Beispieldatensätze wird angezeigt. Die Liste mit Datensätzen enthält Datensätze, die den Wert CHURCH enthalten. Da dieser Wert der Klasse **L Building Types** zugewiesen ist, bearbeitet die Regel für dieses Muster Datensätze, die den Wert enthalten.
- Erweitern Sie **++^+T** und anschließend **Unbearbeitete Datensätze**. Eine Liste mit diesem Muster entsprechenden Beispieldatensätzen wird angezeigt. Die Liste der Datensätze enthält Datensätze, die die Werte INSTITUTE und HOSTEL enthalten. Weil diese Werte der Klasse **L Building Types** nicht zugewiesen sind, werden Datensätze, die diese Werte enthalten, nicht von der entsprechenden Regel verarbeitet.
- Klicken Sie auf **Regelgruppe auswählen** oben auf der Seite **Regeln**, um die Regelgruppe **Input_Overrides** zu verlassen.

Wenn Sie sicherstellen wollen, dass Datensätze, die die Werte INSTITUTE und HOSTEL enthalten, von der entsprechenden Regel verarbeitet werden, weisen Sie diese Werte der Klasse **L Building Types** zu.

Einer Klasse Werte zuweisen

Sie können die neuen Gebäudetypwerte der Klasse der Gebäudetypen zuweisen, indem Sie Klassifikationsdefinitionen manuell zuweisen.

Weitere Informationen zu Klassifikationsdefinitionen:

Klassifikationsdefinitionen schließen die folgenden Elemente ein:

Wert Der definierte Eingabewert.

Standardwert

Eine standardisierte Schreibweise oder Darstellung des Werts, den eine Aktion oder Bedingung in einer Regel anstelle des Eingabewerts verwenden kann.

Klasse Die Klasse, der der Wert zugewiesen ist. Die Klasse wird durch eine aus einem Zeichen bestehende Klassenbezeichnung dargestellt.

Ähnlichkeitsschwelle

Der Variationsgrad, der in der Schreibweise oder Darstellung des Werts zulässig ist. Die Ähnlichkeitsschwelle muss im Bereich von 700 bis 900 liegen. Die Ähnlichkeitsschwelle 900 gibt an, dass die Definition sich nur auf genau mit der Definition übereinstimmende Werte auswirkt.

Liegt die Ähnlichkeitsschwelle unter 900, wirkt die Definition sich möglicherweise auf Werte aus, die dem Wert in der Definition ähneln. Wenn Sie z. B. eine Klassifikationsdefinition für den Wert GRAMS erstellen und die Ähnlichkeitsschwelle auf 800 setzen, wirkt sich die Definition auch auf den Wert GRAM aus.

Vorgehensweise

1. Klicken Sie in Standardization Rules Designer auf die Registerkarte **Klassifikationen**.
2. Klicken Sie auf **Werte definieren**.
3. Fügen Sie eine Definition für einen Gebäudetypwert hinzu:
 - a. Geben Sie den Wert INSTITUTE ein.
 - b. Wählen Sie **L** in der Liste **Klasse** aus.
 - c. Erweitern Sie **Definitionsstatus** und verifizieren Sie, dass **Aktive Definition** ausgewählt ist. Eine inaktive Definition wirkt sich nicht auf den Wert aus.
 - d. Klicken Sie auf **OK**.
4. Öffnen Sie eine Regelgruppe, um auf die Muster in den Daten und auf die Regeln für diese Muster zuzugreifen:
 - a. Klicken Sie auf die Registerkarte **Regeln**.
 - b. Wählen Sie die Regelgruppe **Input_Overrides** aus und klicken Sie dann auf **Öffnen**.
5. Blenden Sie **Musterregel** ein. Eine Liste der in den Daten enthaltenen Muster wird angezeigt.
6. Erweitern Sie **+L^+T** und erweitern Sie anschließend **Copy building details to building name output column**. Eine Liste der von dieser Regel bearbeiteten Beispieldatensätze wird angezeigt. Die Liste der Datensätze enthält nun Datensätze, die den Wert INSTITUTE enthalten.
7. Erweitern Sie **++^+T** und anschließend **Unbearbeitete Datensätze**. Eine Liste der Beispieldatensätze mit diesem Muster wird angezeigt. Die Liste der Datensätze enthält immer noch Datensätze, die den Wert HOSTEL enthalten.

Lerneinheit 2.2: Angepasste Klasse hinzufügen und füllen

In dieser Lerneinheit fügen Sie eine angepasste Klasse hinzu, um einen neuen Werttyp festzulegen. Sie fügen dann gleichzeitig mehrere Klassifikationen hinzu, um dieser Klasse eine Anfangsgruppe mit Werten zuzuweisen.

Übersicht

Zu den Adressdaten, die das fiktive Unternehmen Sample Outdoor Company erworben hat, gehören Werte, die nicht Teil der Adresse sind. So enthalten beispielsweise einige Datensätze Informationen wie INVALID ADDRESS (Ungültige Adresse) oder BEWARE OF DOG (Warnung vor dem Hund). Das Unternehmen betrachtet diesen Informationstyp als Ausnahmedaten. Damit Ausnahmedaten von Adressdaten unterschieden werden können, möchte das Unternehmen eine separate benutzerdefinierte Klasse für Werte hinzufügen, die Ausnahmedaten angeben. Das Unternehmen beschließt, die benutzerdefinierte Klasse X für Ausnahmedaten zu verwenden.

Weitere Informationen zu angepassten Klassifikationen:

Eine benutzerdefinierte Klassifikation stellt stärkere Kontextinformationen zu einem Wert bereit als eine Standardklassifikation. Wenn z. B. eine Definition einer

benutzerdefinierten Klasse nicht den Wert TRAILCHEF zuweist, wird der Wert der Standardklasse + zugewiesen. Diese Standardklasse gibt an, dass der Wert ein einzelnes Alphawort ist. Wenn eine Definition den Wert der benutzerdefinierten Klasse **Product Brand** (Produktmarke) zuweist, die durch das Zeichen B dargestellt wird, stellt der Wert mehr Kontextinformationen bereit. Wenn diese Informationen bereitgestellt werden, können Sie Regeln schreiben, die nur für eine bestimmte Untergruppe der Daten gelten und diese Daten daher effektiver bearbeiten.

Tasks in dieser Lerneinheit

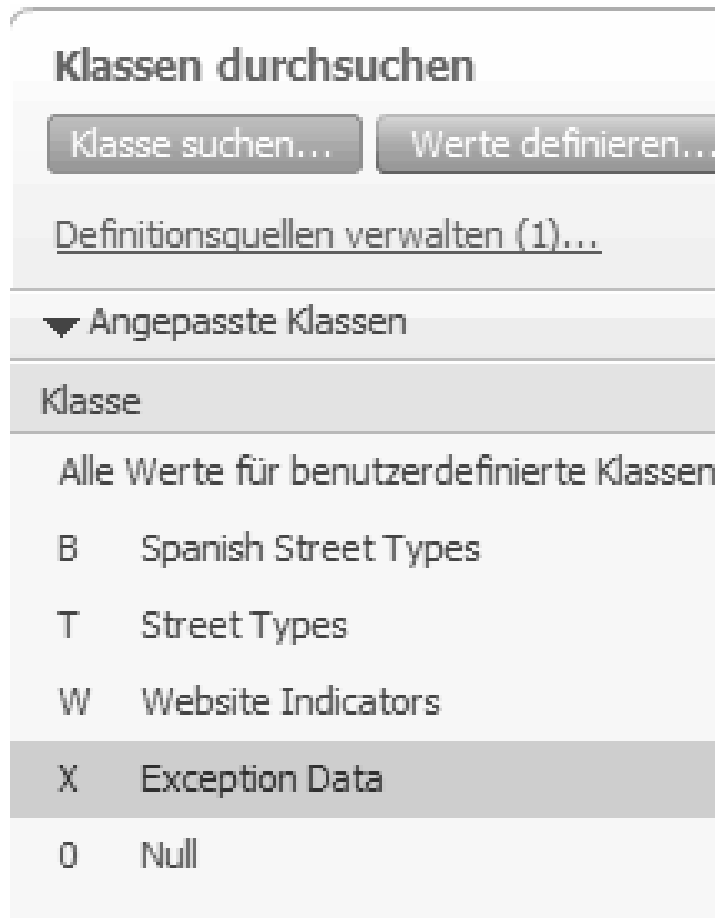
1. „Angepasste Klasse hinzufügen“
2. „Mehrere Klassifikationsdefinitionen hinzufügen“ auf Seite 18

Angepasste Klasse hinzufügen

Wenn Sie eine Klasse für Ausnahmedaten erstellen wollen, fügen Sie die benutzerdefinierte Klasse X aus der Liste verfügbarer Klassen hinzu.

Vorgehensweise

1. Klicken Sie in Standardization Rules Designer auf die Registerkarte **Klassifikationen**.
2. Klicken Sie im Abschnitt **Angepasste Klassen** auf **Angepasste Klasse hinzufügen**.
3. Wählen Sie X in der Liste **Verfügbare Klassen** aus.
4. Geben Sie Exception Data (Ausnahmedaten) in das Feld **Beschreibung** ein.
5. Klicken Sie auf **Hinzufügen**. In der Liste wird ein Eintrag für die neue Klasse angezeigt.



6. Klicken Sie auf **Schließen**.

Mehrere Klassifikationsdefinitionen hinzufügen

Sie können der neuen Klasse X mehrere Werte hinzufügen, indem Sie Mehrfachklassifikationen hinzufügen.

Wenn Sie in Standardization Rules Designer eine neue angepasste Klasse erstellen, werden dieser zunächst keine Werte zugewiesen. Zum Füllen der Klasse fügen Sie Klassifikationsdefinitionen hinzu, die der Klasse Werte zuweisen. Wenn Sie z. B. die benutzerdefinierte Klasse X füllen wollen, fügen Sie Klassifikationsdefinitionen für Werte hinzu, die Ausnahmedaten darstellen. Nachdem der Klasse Werte zugewiesen wurden, können Sie Regeln hinzufügen oder modifizieren, um die Werte für die Ausnahmedaten zu verarbeiten.

Das fiktive Unternehmen Sample Outdoor Company hat Werte ermittelt, die in seinen Firmendatensätzen Ausnahmedaten angeben. So ist beispielsweise der Wert USE nie Teil einer Adresse in seinen Daten, ist jedoch in Ausnahmedaten wie USE LEFT MAILBOX (Linken Briefkasten verwenden) und USE REAR ENTRANCE (Hintereingang verwenden) vorhanden. Im neuen Dataset für Adressen geben die folgenden Werte Ausnahmedaten an:

- INVALID
- BEWARE
- SLEEPS
- USE

Vorgehensweise

1. Klicken Sie in Standardization Rules Designer auf die Registerkarte **Klassifikationen**.
2. Klicken Sie auf **Werte definieren**.
3. Klicken Sie auf **Mehrere definieren** in der rechten unteren Ecke des Fensters **Werte definieren**.
4. Wählen Sie das Kontrollkästchen **Klasse auf alle Definitionen anwenden** aus.
5. Wählen Sie **X** in der Liste **Klasse auswählen** aus. In der Spalte **Klasse** ist **X** automatisch für jeden Wert ausgewählt, den Sie definieren.
6. Geben Sie **INVALID** in das erste Feld der Spalte **Wert** ein und klicken Sie dann außerhalb des Felds. Eine zweite Zeile wird aktiviert.
7. Geben Sie die folgenden Werte mithilfe der Methode in Schritt 6 ein:
 - BEWARE
 - SLEEPS
 - USEGeben Sie jeden Wert in einer separaten Zeile ein.
8. Erweitern Sie **Definitionsstatus** und verifizieren Sie, dass **Aktive Definition** ausgewählt ist. Eine inaktive Definition wirkt sich nicht auf den Wert aus.
9. Klicken Sie auf **OK**.

Wenn Sie die Werte anzeigen wollen, die der benutzerdefinierten Klasse **X** zugewiesen sind, erweitern Sie **Angepasste Klassen** und wählen **X** aus.

Wie geht es weiter

In diesem Modul haben Sie die folgenden Tasks ausgeführt:

- Ermitteln der Klasse, zu der ein Wert gehört
- Hinzufügen von Klassifikationsdefinitionen
- Hinzufügen einer angepassten Klasse

Sie können sich von Standardization Rules Designer abmelden und das nächste Modul später durcharbeiten. Die Überarbeitung bleibt geöffnet.

Im nächsten Modul werden Sie eine Lookup-Tabelle hinzufügen, die abgekürzte Straßennamen in vollständige Straßennamen konvertiert.

Modul 3: Lookup-Tabelle hinzufügen

In diesem Modul fügen Sie eine Lookup-Tabelle hinzu, die die abgekürzte Straßennamen in vollständige Straßennamen konvertiert.

Die Adressdaten, die das fiktive Unternehmen Sample Outdoor Company erworben hat, enthalten abgekürzte Straßennamen. In einigen Datensätzen bestehen die Abkürzungen aus Initialen. So wird z. B. der Wert MARTIN LUTHER KING durch MLK abgekürzt und DANIEL DAFOE durch DD.

Damit sichergestellt wird, dass die Straßennamen nicht mehrdeutig sind, hat das fiktive Unternehmen Sample Outdoor Company Lookup-Tabellendefinitionen für abgekürzte Straßennamen erstellt. Das Unternehmen kann die Lookup-Tabellendefinitionen in Regeln verwenden, die die abgekürzten Straßennamen in vollständige Straßennamen konvertieren.

In diesem Modul werden Sie eine neue Lookup-Tabelle hinzufügen und Lookup-Tabellendefinitionen in Standardization Rules Designer importieren. Danach werden Sie Lookup-Tabellendefinitionen für neue Abkürzungen hinzufügen, die in der aktuellen Lookup-Tabelle nicht enthalten sind.

Lernziele

Nachdem Sie die Lerneinheiten dieses Moduls durchgearbeitet haben, verstehen Sie die für Lookup-Tabellen relevanten Konzepte und Tasks:

- Hinzufügen einer Lookup-Tabelle
- Importieren von Lookup-Tabellendefinitionen
- Hinzufügen von Lookup-Tabellendefinitionen

Erforderlicher Zeitaufwand

Für dieses Modul sind ungefähr 20 Minuten erforderlich.

Voraussetzungen

Stellen Sie sicher, dass Sie die Schritte zum Einrichten des Lernprogramms ausgeführt und die Lernprogrammdateien und Stichprobendaten auf Ihren Computer geladen haben.

Lerneinheit 3.1: Lookup-Tabelle hinzufügen

In dieser Lerneinheit fügen Sie dem Regelsatz USADDRTutorial.SET eine Lookup-Tabelle hinzu, die abgekürzte Straßennamen in vollständige Straßennamen konvertiert.

Übersicht

Das fiktive Unternehmen Sample Outdoor Company möchte mithilfe einer Lookup-Tabelle abgekürzte Straßennamen in vollständige Straßennamen konvertieren. Sie können dem Regelsatz in Standardization Rules Designer eine Lookup-Tabelle hinzufügen.

Weitere Informationen zum Verwenden von Lookup-Tabellen:

Lookup-Tabellen enthalten Definitionen, die Regeln als Teil einer Aktion oder Bedingung referenzieren können. Aktionen oder Bedingungen können eine Lookup-Tabelle wie folgt verwenden:

- Eine Aktion oder Bedingung kann einen bestimmten Wert mit einem Wert in der Lookup-Tabelle vergleichen. So kann z. B. eine Bedingung festlegen, dass eine Regel einen Datensatz nur bearbeitet, wenn ein Wert, der sich an einer bestimmten Position im Datensatz befindet, auch in der Lookup-Tabelle vorhanden ist.
- Eine Aktion kann einen bestimmten Wert in einen Wert in der Lookup-Tabelle konvertieren. So kann eine Aktion z. B. eine Lookup-Tabelle verwenden, die geografische Informationen enthält, um numerische Bereichscodes in Bereichsnamen zu konvertieren.

Vorgehensweise

1. Klicken Sie in Standardization Rules Designer auf die Registerkarte **Lookup-Tabellen**.
2. Klicken Sie auf **Lookup-Tabelle hinzufügen**.
3. Geben Sie Informationen zur Lookup-Tabelle für Straßennamen ein:
 - a. Geben Sie `Street_Name` in das Feld **Name** ein.
 - b. Geben Sie Konvertiert abgekürzte Straßennamen in vollständige Straßennamen in das Feld **Beschreibung** ein.
 - c. Klicken Sie auf **OK**.

Lerneinheit 3.2: Lookup-Tabellendefinitionen importieren und hinzufügen

In dieser Lerneinheit importieren Sie Lookup-Tabellendefinitionen und fügen sie der Lookup-Tabelle **Street_Name** hinzu.

Übersicht

Das fiktive Unternehmen Sample Outdoor Company hat eine CSV-Datei erstellt, die Lookup-Tabellendefinitionen für die Lookup-Tabelle **Street_Name** enthält. Sie können diese Definitionen in die Lookup-Tabelle importieren, die Sie in Standardization Rules Designer hinzugefügt haben.

Vor Kurzem hat das Unternehmen eine neue Abkürzung in den Adressdaten ermittelt und einen Schreibfehler in einer Lookup-Tabellendefinition festgestellt. Sie können die Lookup-Tabelle aktualisieren, indem Sie Lookup-Tabellendefinitionen in Standardization Rules Designer hinzufügen.

Weitere Informationen zu Lookup-Tabellendefinitionen:

Lookup-Tabellendefinitionen enthalten die folgenden Elemente:

Wert Der definierte Eingabewert.

Zurückgegebener Wert

Ein Wert, den eine Aktion oder Bedingung in einer Regel anstelle des Eingabewerts verwenden kann.

Ähnlichkeitsschwelle

Der Variationsgrad, der in der Schreibweise oder Darstellung des Werts zulässig ist. Die Ähnlichkeitsschwelle muss im Bereich von 700 bis 900 liegen. Die Ähnlichkeitsschwelle 900 gibt an, dass die Definition sich nur auf genau mit der Definition übereinstimmende Werte auswirkt.

Liegt die Ähnlichkeitsschwelle unter 900, wirkt die Definition sich möglicherweise auf Werte aus, die dem Wert in der Definition ähneln. Wenn Sie z. B. eine Lookup-Tabellendefinition für den Wert GRAMS erstellen und die Ähnlichkeitsschwelle auf 800 setzen, wirkt sich die Lookup-Tabellendefinition auch auf den Wert GRAM aus.

Vorgehensweise

1. „Lookup-Tabellendefinitionen importieren“
2. „Lookup-Tabellendefinitionen hinzufügen“

Lookup-Tabellendefinitionen importieren

Sie können Lookup-Tabellendefinitionen aus mindestens einer Datei in eine Lookup-Tabelle in Standardization Rules Designer importieren. Sie können Lookup-Tabellendefinitionen in Standardization Rules Designer auch manuell hinzufügen.

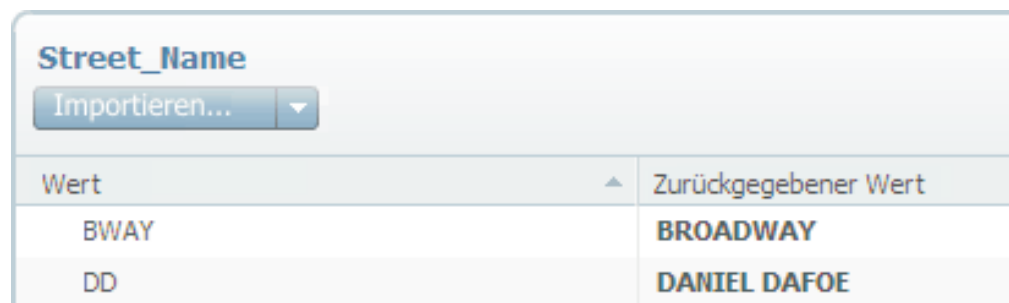
Vorgehensweise

1. Klicken Sie in Standardization Rules Designer auf die Registerkarte **Lookup-Tabellen**.
2. Wählen Sie die Tabelle **Street_Name** in der Liste der Lookup-Tabellen aus. Im rechten Teilfenster wird eine Liste der Werte in der Lookup-Tabelle angezeigt. Da für diese Lookup-Tabelle keine Werte definiert sind, ist diese Liste leer.



3. Wählen Sie **Importieren** in der Liste **Werte definieren** aus.
4. Importieren Sie die Datei `street_abbreviation_lookup.csv`:
 - a. Wechseln Sie zu dem Verzeichnis, das die Datei enthält. Die Datei befindet sich standardmäßig in dem Verzeichnis `Standardization_Rules_Designer_address_tutorial`, in das Sie die Lernprogrammdateien kopiert haben.
 - b. Wählen Sie die Datei aus und klicken Sie dann auf **Öffnen**.
 - c. Klicken Sie im Fenster **Definitionen importieren** auf **OK**.

Die Lookup-Tabellendefinitionen werden importiert und im rechten Fenster werden die in der Lookup-Tabelle definierten Werte angezeigt.



Lookup-Tabellendefinitionen hinzufügen

Sie können Lookup-Tabellendefinitionen in Standardization Rules Designer manuell hinzufügen.

Vorgehensweise

1. Wählen Sie für die Tabelle 'Street_Name' **Werte definieren** in der Liste **Importieren** aus.
2. Füllen Sie die Felder aus, um eine Lookup-Tabellendefinition für den neuen Wert hinzuzufügen:
 - a. Geben Sie die Abkürzung IB in das Feld **Wert** ein.
 - b. Geben Sie den Straßennamen IRL0 BRONSON in das Feld **Zurückgegebener Wert** ein.
 - c. Erweitern Sie **Wert definieren als** und stellen Sie sicher, dass **Aktive Definition** ausgewählt ist. Eine inaktive Definition wirkt sich nicht auf den Wert aus.
 - d. Klicken Sie auf **OK**.

Im rechten Teilfenster wird der Wert angezeigt, für den Sie eine Definition hinzugefügt haben.

Street_Name	
Werte definieren...	
Wert	Zurückgegebener Wert
DD	DANIEL DAFOE
IB	IRLO BRONSON

3. Erstellen Sie eine neue Lookup-Tabellendefinition für einen Wert, indem Sie Schritt 2 wiederholen, diesmal für den Wert jedoch SP und für den Rückgabewert SARA PELLETIER eingeben.
4. Verifizieren Sie, dass die neue Lookup-Tabellendefinition die aktive Definition ist:
 - a. Wählen Sie die Tabelle **Street_Name** in der Liste der Lookup-Tabellen aus. Im rechten Teilfenster wird eine Liste der Werte in der Lookup-Tabelle angezeigt.
 - b. Wählen Sie in der Liste der Werte SP aus. Auf der Seite **Wert definieren** wird eine Liste der Definitionen für den Wert angezeigt. Die von Ihnen hinzugefügte Definition ist ausgewählt und ist daher die aktive Definition.

Wert:	Zurückgegebener Wert:
<input checked="" type="radio"/> SP	SARA PELLETIER
<input type="radio"/> SP	SARA PELLETER

Wie geht es weiter

In diesem Modul haben Sie die folgenden Tasks ausgeführt:

- Hinzufügen einer Lookup-Tabelle
- Importieren von Lookup-Tabellendefinitionen
- Hinzufügen von Lookup-Tabellendefinitionen

Im nächsten Modul werden Sie Regeln erstellen und modifizieren, die die Adressen in den Kundendatensätzen für das fiktive Unternehmen Sample Outdoor Company bearbeiten.

Modul 4: Regeln hinzufügen und modifizieren

In diesem Modul fügen Sie Regeln hinzu und modifizieren sie, um sicherzustellen, dass Datensätze für neue Adressdaten ordnungsgemäß standardisiert werden.

Das fiktive Unternehmen Sample Outdoor Company möchte Regeln hinzufügen oder modifizieren, damit die Datensätze ihrer neuen Adressdaten richtig bearbeitet werden.

Zunächst betrachtet das Unternehmen die Regel, die das häufigste Muster in den Daten bearbeitet. Die Datensätze, die mit diesem Muster übereinstimmen, werden relativ gut bearbeitet, einige Straßennamen sind jedoch abgekürzt. Das Unternehmen möchte die in Modul 3 hinzugefügte Lookup-Tabelle verwenden, um abgekürzte Straßennamen in vollständige Straßennamen zu konvertieren.

Danach stellt das Unternehmen fest, dass Straßennamen, die führende Ziffern gefolgt von Buchstaben enthalten, nicht bearbeitet werden. So werden z. B. Straßennamen wie 45TH und 76E von keiner Regel bearbeitet. Das Unternehmen muss eine Regel hinzufügen, um diese Daten zu bearbeiten.

Bei der Betrachtung der übrigen unbearbeiteten Muster findet das Unternehmen ein Muster, in dem Hausnummer und Straßename in einem Wert verknüpft sind. So sind beispielsweise im folgenden Datensatz die Werte 637 und HEGENBERGER zu 637HEGENBERGER verknüpft: 673HEGENBERGER ROAD. Das Unternehmen muss für dieses Muster eine Regel hinzufügen, die die Daten auf die entsprechenden Ausgabespalten aufteilt.

Nachdem die Regeln hinzugefügt und modifiziert wurden, beschließt das Unternehmen, den tatsächlichen Regelsatz zu aktualisieren, indem die Überarbeitung publiziert wird.

Lernziele

Nachdem Sie die Lerneinheiten dieses Moduls durchgearbeitet haben, verstehen Sie die Konzepte und Tasks, die zum Hinzufügen und Modifizieren von Regeln sowie zum Publizieren von Überarbeitungen erforderlich sind:

- Modifizieren einer Regel
- Ermitteln unbearbeitbarer Muster und Anzeigen von Stichprobendatensätzen für diese Muster
- Anpassen der in Standardization Rules Designer verfügbaren Ausgabespalten
- Hinzufügen einer Basisregel durch Zuordnung von Werten in einem Datensatz zu entsprechenden Ausgabespalten
- Hinzufügen einer Regel, die einen verknüpften Wert in zwei unterschiedliche Werte aufteilt
- Publizieren einer Überarbeitung

Erforderlicher Zeitaufwand

Für dieses Modul sind ungefähr 60 Minuten erforderlich.

Voraussetzungen

Stellen Sie sicher, dass Sie alle Schritte der Einrichtung vollständig ausgeführt haben, und dass die Lernprogrammdateien und Stichprobendaten auf Ihr Computersystem geladen wurden.

Lerneinheit 4.1: Regel für die ordnungsgemäße Bearbeitung von Daten modifizieren

In dieser Lerneinheit stellen Sie sicher, dass alle Datensätze, die mit einem Muster übereinstimmen, ordnungsgemäß verarbeitet werden, indem Sie die Regel, die diese Datensätze verarbeitet, modifizieren.

Übersicht

Um Fehler bei der Zustellung zu vermeiden, möchte das fiktive Unternehmen Sample Outdoor Company sicherstellen, dass alle Straßennamen in den Adressdaten vollständig sind. Für das häufigste Muster in den Daten sind einige Straßennamen abgekürzt. Sie können die Regel, die diese Daten bearbeitet, so modifizieren, dass die in Modul 3 hinzugefügte Lookup-Tabelle verwendet wird. Die Regel verwendet die Lookup-Tabelle, um abgekürzte Straßennamen in vollständige Straßennamen zu konvertieren.

Weitere Informationen zu Regeln und Regelgruppen:

Regeln sind Prozesse, die Gruppen zusammengehöriger Datensätze standardisieren. Regeln können auf Datensätze angewendet werden, die mit demselben Muster oder mit exakten Textzeichenfolgen übereinstimmen. Wenn Sie eine Regel hinzufügen oder modifizieren, ordnen Sie Werte in Eingabedatensätzen Ausgabespalten zu, geben Aktionen an, die die Daten bearbeiten, und geben Bedingungen an, die sicherstellen, dass Regeln nur auf die entsprechenden Datensätze angewendet werden.

Eine Regelgruppe ist eine Sammlung von Regeln, die an demselben Punkt im Standardisierungsprozess auf Datensätze angewendet werden. Um sicherzustellen, dass Regeln in einer bestimmten Reihenfolge angewendet werden, können Sie die Regeln in Regelgruppen in Standardization Rules Designer zusammenfassen. Danach können Sie die Regelgruppen aus der Musteraktionsspezifikation (früher Musteraktionsdatei genannt) aufrufen.

Vorgehensweise

1. Klicken Sie auf die Registerkarte **Regeln**, wählen Sie die Regelgruppe **Input_Overrides** aus und klicken Sie dann auf **Öffnen**.
2. Blenden Sie **Musterregel** ein. Eine Liste der in den Daten enthaltenen Muster wird angezeigt.

▼ Musterregel		Muster hinzufügen...	
Regel	Häufigkeit	Unbearbeitet	
▶ ^+T	42.73%	0%	
▶ ^D+T	5.34%	5.34%	
▶ ^T+	4.59%	4.59%	
▶ ^++T	3.65%	3.65%	
▶ ^TT	3.37%	3.37%	

3. Erweitern Sie ^+T und erweitern Sie anschließend **Copy address data to output columns**.
4. Klicken Sie mit der rechten Maustaste auf den Wert in der Ausgabespalte StreetName und klicken Sie dann auf **Editieraktion**. Im Fenster **Aktionen** können Sie die Daten bearbeiten, die an die Ausgabespalte gesendet werden.
5. Klicken Sie im Abschnitt **Objekt suchen** auf **Ja**.
6. Verwenden Sie die Lookup-Tabelle **Street_Name**, um abgekürzte Straßennamen in vollständige Straßennamen zu konvertieren:
 - a. Wählen Sie **Lookup-Tabelle** in der Liste **Quelle** aus.
 - b. Wählen Sie **Street_Name** in der Liste **Lookup-Tabelle** aus.
 - c. Wählen Sie in der Liste **Wenn gefunden** die Option **In Rückgabewert konvertieren** aus. Ist der Wert eines Straßennamens in der Lookup-Tabelle **Street_Name** vorhanden, wird der abgekürzte Straßennamen in den vollständigen Straßennamen konvertiert.
 - d. Wählen Sie in der Liste **Wenn nicht gefunden** die Option **Wert verwenden** aus.
 - e. Klicken Sie auf **OK**.
7. Klicken Sie auf **Anwenden**. Die Seite **Regel definieren** enthält Stichprobendatensätze, die einen abgekürzten Straßennamen enthalten, dessen vollständiger Name in der Ausgabespalte StreetName angezeigt wird. Beispiel: Der Stichprobendatensatz enthält den abgekürzten Straßennamen SM und die Ausgabespalte StreetName enthält den Wert STUART MILL.

Beispieldatensatz

588 SM ST

588

HouseNumber HouseNumberSuffix StreetPrefixDirectional StreetPrefixType **STUART MILL**
StreetName

ST

StreetSuffixType StreetSuffixQualifier StreetSuffixDirectional RuralRouteType RuralRouteValue

Lerneinheit 4.2: Unbearbeitete Datensätze ermitteln und Ausgabespalten anpassen

In dieser Lerneinheit ermitteln Sie Datensätze, die von Regeln im Regelsatz nicht verarbeitet werden, sodass Sie eine Regel für diese Datensätze hinzufügen können.

Sie blenden dann die Ausgabespalten aus, die Sie nicht benötigen, wenn Sie eine Regel für unbearbeitete Datensätze hinzufügen.

Tasks in dieser Lerneinheit

1. „Unbearbeitete Datensätze ermitteln“
2. „Ausgabespalten anpassen“ auf Seite 28

Unbearbeitete Datensätze ermitteln

Ein Fachverantwortlicher des fiktiven Unternehmens Sample Outdoor Company führte einen SQA-Bericht für einen Job aus, der die erworbenen Kundendaten standardisierte. Der Job wendete den Regelsatz USADDRTutorial.SET auf die Daten an. Bei der Überprüfung des SQA-Berichts stellte der Fachverantwortliche unbearbeitete Datensätze fest und erkannte, welche Muster mit den unbearbeiteten Datensätzen übereinstimmen.

Der Fachverantwortliche stellte weiterhin fest, dass einige der unbearbeiteten Datensätze Straßennamen enthalten, die führende Ziffern gefolgt von Buchstaben enthalten. So wurden z. B. Datensätze, die Straßennamen wie 45TH und 76E enthalten, von keiner Regel bearbeitet. Der Fachverantwortliche stellte außerdem fest, dass Datensätze, die derartige Straßennamen enthalten, häufig mit dem Muster $\wedge > \mathbf{T}$ übereinstimmen. So stimmt z. B. der Datensatz 243 45E AV mit dem Muster $\wedge > \mathbf{T}$ überein.

Tabelle 2. Die Werte 243 45E AV stimmen mit dem Muster $\wedge > \mathbf{T}$ überein

\wedge	$>$	\mathbf{T}
Nur aus Zahlen bestehender Wert	Führende Ziffern gefolgt von Buchstaben	Straßentyp

Sie können die unbearbeiteten Datensätze angeben, die den Straßennamentyp enthalten, für die der Fachverantwortliche eine Regel hinzufügen will. Später können Sie eine Regel hinzufügen, die diese Datensätze bearbeitet.

Weitere Informationen zu unbearbeiteten Datensätzen:

Unbearbeitete Datensätze sind Datensätze, auf die keine Regel des Regelsatzes angewendet wird. Unbearbeitete Datensätze können Datensätze aus den Stichproben- daten oder vom System generierte Datensätze sein.

Über Standardization Rules Designer können Sie Ihre Standardisierungsprozesse am effektivsten verwalten, wenn Sie Regeln hinzufügen, die sich auf eine möglichst große Anzahl Datensätze in Ihren Daten auswirken. Nachdem Sie unbearbeitete Datensätze ermittelt haben, können Sie Regeln hinzufügen, die auf diese Datensätze angewendet werden.

Vorgehensweise

1. Klicken Sie auf die Registerkarte **Regeln**, wählen Sie die Regelgruppe **Input_Overrides** aus und klicken Sie dann auf **Öffnen**.
2. Blenden Sie **Musterregel** ein. Eine Liste der in den Daten enthaltenen Muster wird angezeigt.

▼ Musterregel		Muster hinzufügen...	
Regel	Häufigkeit	Unbearbeitet	
▶ ^+T	42.73%	0%	
▶ ^D+T	5.34%	5.34%	
▶ ^T+	4.59%	4.59%	
▶ ^++T	3.65%	3.65%	
▶ ^TT	3.37%	3.37%	

3. Filtern Sie die Liste mit den Mustern, dass nur das Muster ^>T angezeigt wird:
 - a. Klicken Sie in der rechten oberen Ecke des Teilfensters **Regeln durchsuchen** auf .
 - b. Stellen Sie sicher, dass das Kontrollkästchen **Filter aktivieren** ausgewählt ist.
 - c. Wählen Sie **Gleich** in der Liste **Bedingung** aus.
 - d. Geben Sie ^>T in das Feld **Muster** ein und klicken Sie dann auf **OK**.

Die Liste wird gefiltert und nur das Muster ^>T wird angezeigt.

▼ Musterregel		Muster hinzufügen...	
Regel	Häufigkeit	Unbearbeitet	
▶ ^>T	1.68%	1.68%	

4. Erweitern Sie ^>T. Für dieses Muster sind keine Regeln vorhanden. Alle Datensätze sind unbearbeitet.

Weil für dieses Muster keine Regeln erstellt wurden und ein Fachverantwortlicher festgestellt hat, dass das Muster wichtig ist, können Sie wichtige Datensätze in Ihren Daten bearbeiten, wenn Sie eine Regel für dieses Muster hinzufügen. In der nächsten Lerneinheit geben Sie eine Regel für dieses Muster ein.

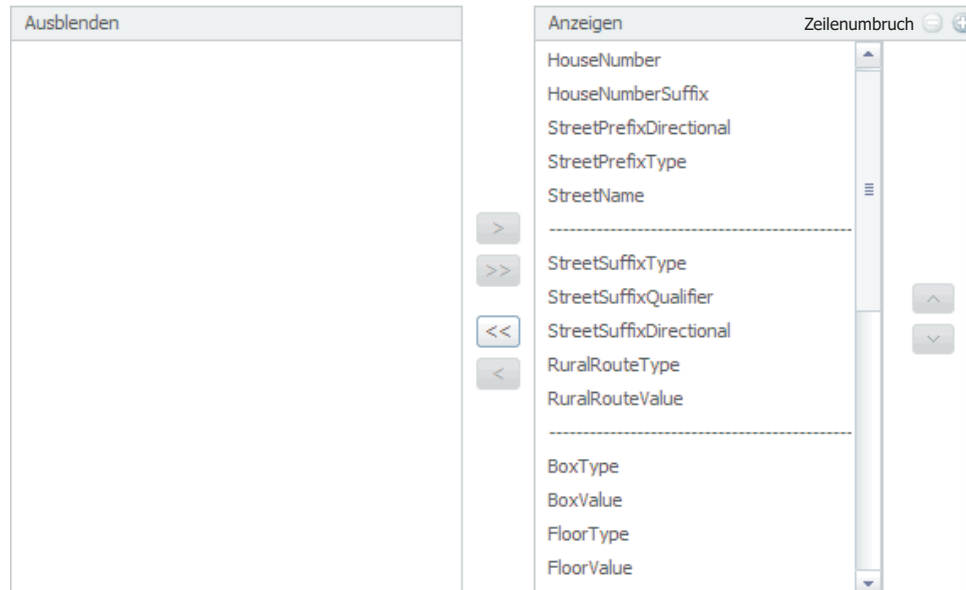
Ausgabespalten anpassen



Sie können Ausgabespalten ausblenden, die für eine Regel nicht erforderlich sind.

Für die Regel, die das fiktive Unternehmen Samples Outdoor Company für das Muster ^>T hinzufügen möchte, sind nicht viele der auf der Seite **Regel definieren** angezeigten Ausgabespalten erforderlich. Damit das Teilfenster, in dem die Regel hinzugefügt wird, vereinfacht wird, möchte das Unternehmen die Ausgabespalten, die nicht erforderlich sind, auf der Seite **Regel definieren** ausblenden.

Vorgehensweise

1. Klicken Sie im Muster ^>T auf **Unbearbeitete Datensätze**. Ein Beispieldatensatz und eine List mit Ausgabespalten wird angezeigt.
2. Klicken Sie im Teilfenster **Unbearbeitete Datensätze** auf **Ausgabespalten anpassen**. Eine Liste der ausgeblendeten und der eingeblendeten Spalten wird angezeigt.



3. Passen Sie die Ausgabespalten so an, dass nur eine Untergruppe der Ausgabespalten angezeigt wird:
 - a. Klicken Sie auf , um alle Ausgabespalten in die Liste mit den ausgeblendeten Spalten zu versetzen.
 - b. Wählen Sie die Ausgabespalte HouseNumber aus und klicken Sie dann auf , um die Ausgabespalte in die Liste mit den angezeigten Spalten zu versetzen.
 - c. Wiederholen Sie Schritt 3b für die Ausgabespalten StreetName, StreetSuffixType und AdditionalAddress.
 - d. Klicken Sie auf **OK**.

Auf der Seite **Regel definieren** werden die Ausgabespalten nicht angezeigt, die Sie soeben ausgeblendet haben.

Lerneinheit 4.3: Regel für unbearbeitete Datensätze hinzufügen

In dieser Lerneinheit fügen Sie eine Regel hinzu, die ein wichtiges Muster in den Daten bearbeitet.

Übersicht

Das fiktive Unternehmen Sample Outdoor Company möchte eine Regel für ein Muster hinzufügen, das von einem Fachverantwortlichen als wichtiges Muster eingestuft wurde.

Weitere Informationen zu Aktionen:

Eine Aktion ist ein Teil einer Regel, die angibt, wie die Regel einen Datensatz verarbeitet. Eine Aktion kann angeben, dass Informationen aus den Eingabedaten in die Ausgabespalten versetzt werden und wie diese Daten bearbeitet werden sollen. In Standardization Rules Designer können Sie eine Aktion verwenden, um Daten auf die folgenden Arten zu bearbeiten:

- Angeben des Teils der Eingabedaten, der bearbeitet wird. Dabei kann es sich um einen genauen Wert, einen Standardwert, mindestens ein Zeichen im Wert oder ein Literal handeln.
- Vergleichen des ausgewählten Teils der Eingabedaten mit Werten, die sich in einer Lookup-Tabelle oder Liste befinden, oder Konvertieren der Eingabedaten in diese Werte.
- Hinzufügen der Ausgabedaten zu einer Ausgabespalte und Angeben eines führenden Trennzeichens zwischen diesen Daten und anderen Daten in der Spalte.

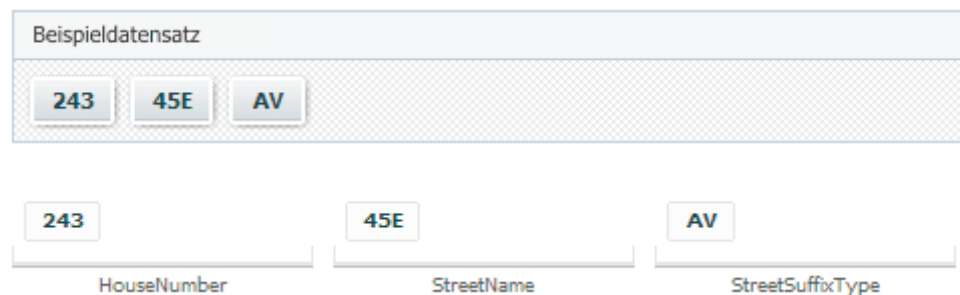
Vorgehensweise

1. Klicken Sie auf die Registerkarte **Regeln**, wählen Sie die Regelgruppe **Input_Overrides** aus und klicken Sie dann auf **Öffnen**.
2. Blenden Sie **Musterregel** ein. Eine Liste der in den Daten enthaltenen Muster wird angezeigt.
3. Erweitern Sie **^>T** und klicken Sie dann auf **Unbearbeitete Datensätze**. Ein Stichprobendatensatz und eine Liste mit Ausgabespalten wird angezeigt.
4. Ist der Stichprobendatensatz auf der Seite **Regel definieren** nicht 243 45E AV, wählen Sie 243 45E AV in der Liste der Stichprobendatensätze aus.
5. Ziehen Sie die Werte auf die Spalte, um die Adresswerte des Stichprobendatensatzes den entsprechenden Ausgabespalten zuzuordnen. Die folgende Tabelle zeigt die Ausgabespalte für die einzelnen Adresswerte an:

Tabelle 3. Ausgabespalten für die einzelnen Adresswerte im Stichprobendatensatz

Wert	Ausgabespalte
243	HouseNumber
45E	StreetName
AV	StreetSuffixType

Die Werte der Beispieldatensätze werden den Ausgabespalten zugeordnet.




6. Wenden Sie die Regel auf das Muster an:
 - a. Klicken Sie auf **Als neue Regel anwenden**.
 - b. Geben Sie Map address values to output columns in das Feld **Regelname**.
 - c. Klicken Sie auf **OK**.

Die Regel wird hinzugefügt und unter dem Muster **^>T** angezeigt.

▼ 📁 ^>T	1.68%
▶ 📌 Map address values to output columns	1.68%

7. Wenn die Musterliste noch gefiltert wird, entfernen Sie den Filter.

- a. Klicken Sie in der rechten oberen Ecke des Teilfensters **Regeln durchsuchen** auf .
- b. Wählen Sie das Kontrollkästchen **Filter aktivieren** ab und klicken Sie dann auf **OK**.

Lerneinheit 4.4: Regel hinzufügen, um Werte auf verschiedene Ausgabespalten aufzuteilen

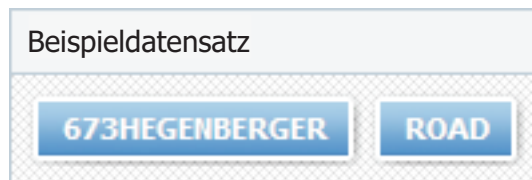
In dieser Lerneinheit fügen Sie eine Regel hinzu, die einen Wert aufteilt und jeden Teil des aufgeteilten Werts einer anderen Ausgabespalte zuordnet.

Übersicht

Datensätze in Eingabedaten können verknüpfte Informationen enthalten. Die Eingabedaten können z. B. den Wert 637HEGENBERGER enthalten. Dieser Wert verknüpft 637, eine Hausnummer, mit HEGENBERGER, einem Straßennamen. Das fiktive Unternehmen Sample Outdoor Company hat festgestellt, dass viele Datensätze, die verknüpfte Werte enthalten, mit dem Muster >T übereinstimmen. Das Unternehmen möchte in diesen Datensätzen die Hausnummer vom Straßennamen trennen. Damit die verknüpften Werte aufgeteilt werden, können Sie eine Regel hinzufügen, die jeden Teil des Eingabewerts einer anderen Ausgabespalte zuordnet.

Vorgehensweise

1. Klicken Sie auf die Registerkarte **Regeln**, wählen Sie die Regelgruppe **Input Overrides** aus und klicken Sie dann auf **Öffnen**.
2. Blenden Sie **Musterregel** ein. Eine Liste der in den Daten enthaltenen Muster wird angezeigt.
3. Erweitern Sie >T und klicken Sie dann auf **Unbearbeitete Datensätze**. Ein Stichprobendatensatz wird angezeigt. Einer der Werte des Datensatzes, der im Muster als > dargestellt wird, enthält führende Ziffern, die von Buchstaben gefolgt werden.




4. Ist der Stichprobendatensatz auf der Seite **Regel definieren** nicht 673HEGENBERGER ROAD, wählen Sie 673HEGENBERGER ROAD in der Liste der Stichprobendatensätze aus. Wird 673HEGENBERGER ROAD nicht in der Liste der Stichprobendatensätze angezeigt, erhöhen Sie die Anzahl der Datensätze in der Liste.
 - a. Klicken Sie in der rechten oberen Ecke des Teilfensters **Regeln durchsuchen** auf .
 - b. Wählen Sie **100** in der Liste **Anzuzeigende Datensätze** aus.
 - c. Klicken Sie auf **OK**. Die Liste der Stichprobendatensätze enthält 673HEGENBERGER ROAD. Sie können den Datensatz in der Liste auswählen.
5. Ziehen Sie die Werte auf die Spalte, um die Werte des Stichprobendatensatzes den entsprechenden Ausgabespalten zuzuordnen. Die folgende Tabelle zeigt die Ausgabespalte für die einzelnen Werte an:

Tabelle 4. Ausgabespalten für die einzelnen Adresswerte im Stichprobendatensatz

Wert	Ausgabespalte
673HEGENBERGER	HouseNumber
ROAD	StreetSuffixType

Die Werte der Beispieldatensätze werden den Ausgabespalten zugeordnet. Der Wert in der Ausgabespalte HouseNumber enthält jedoch einen Wert, der in die Ausgabespalte StreetName gehört.

The screenshot shows a 'Beispieldatensatz' (Sample Record) with two input fields: '673HEGENBERGER' and 'ROAD'. Below, the data is mapped to three output columns: 'HouseNumber' (containing '673HEGENBERGER'), 'StreetName' (containing 'HEGENBERGER'), and 'StreetSuffixType' (containing 'ROAD'). A red vertical line is drawn through the '673HEGENBERGER' field to indicate the split between the house number and the street name.

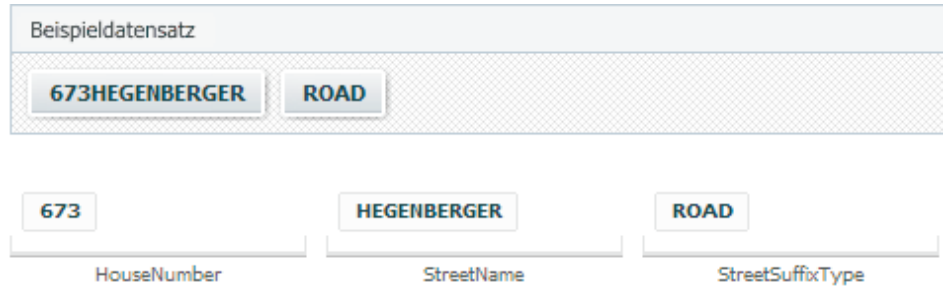
6. Klicken Sie in der Ausgabespalte HouseNumber mit der rechten Maustaste auf 673HEGENBERGER und klicken Sie dann auf **Editieraktion**. Im Fenster **Aktionen** können Sie die Daten bearbeiten, die an die Ausgabespalte gesendet werden.
7. Bearbeiten Sie die Aktion, damit nur die führenden Ziffern der Ausgabespalte HouseNumber hinzugefügt werden:
 - a. Wählen Sie **Präfix** in der Liste **Objekt** aus.
 - b. Klicken Sie auf **Alle führenden Ziffern**. Das Feld **Beispielergebnis** zeigt den Wert 673 an.
8. Klicken Sie auf **Aktion für die verbleibenden Zeichen angeben** und dann auf **OK**. Das Fenster **Aktionen** wird geöffnet. Optionen sind ausgewählt, die eine Aktion für die Zeichen angeben, die Sie nicht der Ausgabespalte HouseNumber zugeordnet haben.

The 'Aktionen' dialog box shows the following configuration:

- Klasse: > | Beispieldatensatz: 673HEGENBERGER ROAD
- Wählen Sie das Objekt aus, für das eine Aktion ausgeführt werden soll.
- *Objekt: Alles mit Ausnahme von... | *Objekt: Präfix... | Beispielergebnis: HEGENBERGER
- Alle führenden Buchstaben
- Alle führenden Ziffern
- Nur die folgende Anzahl führender Zeichen: 1

9. Ordnen Sie die verbleibenden Zeichen der Ausgabespalte StreetName zu:
 - a. Wählen Sie **StreetName** in der Liste **Ausgabespalte** aus.
 - b. Klicken Sie auf **OK**.

Die Werte werden in den richtigen Ausgabespalten angezeigt.



10. Wenden Sie die Regel auf das Muster an:
 - a. Klicken Sie auf **Als neue Regel anwenden**.
 - b. Geben Sie **Split house number and street name** in das Feld **Name**.
 - c. Klicken Sie auf **OK**.

Die Regel wird hinzugefügt und unter dem Muster >T angezeigt.

▼ 📁 >T	1.49%
▶ 📌 Split house number and street name	1.49%

Lerneinheit 4.5: Überarbeitung publizieren

In dieser Lerneinheit veröffentlichen Sie eine Überarbeitung im Metadatenrepository, damit der erweiterte Regelsatz in einem Standardisierungsjob verwendet werden kann.

Übersicht

Nachdem Sie in Standardization Rules Designer Änderungen an einem Regelsatz vorgenommen haben, können Sie diese Änderungen auf den Regelsatz im Metadatenrepository anwenden, indem Sie die Überarbeitung publizieren. Nach dem Publizieren der Überarbeitung muss der Regelsatz im Designer-Client bereitgestellt werden, bevor der Regelsatz in einem Standardisierungsjob angewendet werden kann.

Das fiktive Unternehmen Sample Outdoor Company hat die Klassifikationen, Lookup-Tabellen und Regeln in der Version des Regelsatzes USADDRTutorial.SET aktualisiert, die in der Datenbank von Standardization Rules Designer gespeichert ist. Um sicherzustellen, dass die Version des Regelsatzes, die im Standardisierungsjob verwendet wird, diese Änderungen enthält, publizieren Sie die Überarbeitung im Metadatenrepository.

Vorgehensweise

1. Klicken Sie in Standardization Rules Designer auf die Registerkarte **Ausgangsposition**.
2. Klicken Sie im Navigationsteilfenster auf **Überarbeitung anzeigen und publizieren**.
3. Optional: Geben Sie in das Teilfenster **Eigenschaften** Informationen zum Regelsatz ein:
 - a. Geben Sie in das Feld **Hinweise** eine Beschreibung der Änderungen ein, die Sie am Regelsatz vorgenommen haben, seit die Überarbeitung publiziert wurde.

- b. Klicken Sie auf **Anwenden**.
4. Klicken Sie auf **Publizieren**. Klicken Sie im Fenster für die Nachricht zum Publizieren von Änderungen auf **Ja**.
- Der Regelsatz im Metadatenrepository wird mit den Änderungen aktualisiert, die Sie in Standardization Rules Designer vorgenommen haben.

Wie geht es weiter

In diesem Modul haben Sie die folgenden Tasks ausgeführt:

- Modifizieren einer Regel
- Ermitteln unbearbeitbarer Muster und Anzeigen von Stichprobendatensätzen für diese Muster
- Verwalten der in Standardization Rules Designer verfügbaren Ausgabespalten
- Hinzufügen einer Basisregel durch Zuordnung von Werten in einem Datensatz zu entsprechenden Ausgabespalten
- Hinzufügen einer Regel, die einen verknüpften Wert in zwei unterschiedliche Werte aufteilt
- Publizieren einer Überarbeitung

Sie haben das Lernprogramm zum Erweitern eines Adressregelsatzes in IBM InfoSphere QualityStage Standardization Rules Designer durchgearbeitet.

Nun können Sie mindestens eine der folgenden Tasks ausführen:

- Vervollständigen Sie das Lernprogramm zum Erweitern eines Produktregelsatzes in Standardization Rules Designer.
- Importieren Sie Ihre eigenen Stichprobendaten und erweitern Sie einen Regelsatz in Standardization Rules Designer, der auf die Daten angewendet wird
- Weitere Informationen zu IBM InfoSphere QualityStage und Standardization Rules Designer finden Sie im *IBM InfoSphere QualityStage Benutzerhandbuch*.

Anhang A. Eingabehilfen in den Produkten

Sie erhalten Informationen zum Status der Eingabehilfen in IBM Produkten.

Die Produktmodule und Benutzerschnittstellen von IBM InfoSphere Information Server sind nicht uneingeschränkt für behindertengerechte Bedienung geeignet.

Informationen zum Status der Eingabehilfen in IBM Produkten finden Sie unter http://www.ibm.com/able/product_accessibility/index.html.

Dokumentation im behindertengerechten Format

Dokumentation im behindertengerechten Format für Produkt steht im IBM Knowledge Center zur Verfügung. Im IBM Knowledge Center wird zur Darstellung der Dokumentation das Format XHTML 1.0 verwendet, das mit den meisten Web-Browsern geöffnet werden kann. Da das IBM Knowledge Center XHTML verwendet, können Sie in Ihrem Browser Anzeigevorgaben festlegen. Darüber hinaus ist der Einsatz von Sprachausgabeprogrammen und anderen Unterstützungseinrichtungen für den Zugriff auf die Dokumentation möglich.

Die im IBM Knowledge Center zur Verfügung stehende Dokumentation wird auch in Form von PDF-Dateien bereitgestellt, die nicht uneingeschränkt für behindertengerechte Bedienung geeignet sind.

IBM und Eingabehilfen

Weitere Informationen zum Engagement von IBM hinsichtlich der Eingabehilfen finden Sie im IBM Human Ability and Accessibility Center.

Anhang B. Kontaktaufnahme mit IBM

Sie können sich an IBM wenden, um Unterstützung, Informationen zu Software-Services, Produktinformationen sowie allgemeine Informationen zu erhalten. Darüber hinaus können Sie Feedback zu den Produkten und zur Dokumentation an IBM senden.

In der folgenden Tabelle sind Ressourcen für die Kundenunterstützung, für Software-Services, für Schulungen sowie für Produkt- und Lösungsinformationen aufgeführt.

Tabelle 5. IBM Ressourcen

Ressource	Beschreibung und Position
IBM Support-Portal	Sie können die Unterstützungsinformationen je nach Bedarf anpassen, indem Sie die Produkte und Themen, die für Sie von Interesse sind, unter www.ibm.com/support/entry/portal/Software/Information_Management/InfoSphere_Information_Server auswählen.
Software-Services	Informationen zu Software-, IT- und Unternehmensberatungsservices erhalten Sie auf der Site 'Lösungen' unter www.ibm.com/businesssolutions/de .
Meine IBM	Auf der Site 'Meine IBM' unter www.ibm.com/account/de/de/ können Sie ein Konto einrichten und so Links auf IBM Websites und Informationen Ihren speziellen Anforderungen an die technische Unterstützung entsprechend verwalten.
Schulung und Zertifizierung	Informationen zu technischen Schulungs- und Ausbildungsservices, mit deren Hilfe Einzelpersonen sowie Mitarbeiter von Unternehmen und öffentlichen Organisationen IT-Kenntnisse erwerben, optimieren und auf dem neuesten Stand halten können, finden Sie unter http://www.ibm.com/training .
IBM Ansprechpartner	Einen IBM Ansprechpartner, bei dem Sie Informationen zu Lösungen erhalten, finden Sie unter www.ibm.com/connect/ibm/us/en/ bzw. www.ibm.com/contact/de/de/ .

Anhang C. Auf Produktdokumentation zugreifen

Die Dokumentation wird in einer Vielzahl von Formaten bereitgestellt: online im IBM Knowledge Center, optional in einem lokal installierten Information Center sowie in Form von Handbüchern im PDF-Format. Sie können direkt über die Produktclientschnittstelle auf die Onlinehilfe oder die lokal installierte Hilfe zugreifen.

IBM Knowledge Center ist die beste Methode, um aktuelle Informationen zu InfoSphere Information Server zu suchen. Das IBM Knowledge Center enthält Hilfe für die meisten Produktschnittstellen sowie die gesamte Dokumentation für alle Produktmodule der Suite. Sie können das IBM Knowledge Center über das installierte Produkt oder über einen Web-Browser öffnen.

Auf das IBM Knowledge Center zugreifen

Es gibt verschiedene Möglichkeiten, auf die Onlinedokumentation zuzugreifen:

- Klicken Sie auf den Link **Hilfe** rechts oben in der Clientschnittstelle.
- Drücken Sie die Taste F1. Mit der Taste F1 wird normalerweise das Thema aufgerufen, das eine Beschreibung des aktuellen Kontexts der Clientschnittstelle enthält.

Anmerkung: In Web-Clients kann die Taste F1 nicht verwendet werden.

- Geben Sie die Adresse in einem Web-Browser ein, beispielsweise, wenn Sie nicht am Produkt angemeldet sind.

Geben Sie die folgende Adresse ein, um auf alle Versionen der Dokumentation zu InfoSphere Information Server zuzugreifen:

<http://www.ibm.com/support/knowledgecenter/SSZJPZ/>

Wenn Sie auf ein bestimmtes Thema zugreifen wollen, geben Sie die Versionsnummer zusammen mit der Produkt-ID, dem Namen des Dokumentations-Plugins und dem Themenpfad in der URL an. Die URL für Version 11.3 dieses Themas lautet beispielsweise wie folgt, wobei das Symbol \Rightarrow eine Zeilenfortsetzung angibt:

http://www.ibm.com/support/knowledgecenter/SSZJPZ_11.3.0/com.ibm.swg.im.iis.common.doc/common/accessingiidoc.html

Tipp:

Für das Knowledge Center gibt es auch eine Kurz-URL:

<http://ibm.biz/knowctr>

Zur Angabe einer Kurz-URL zu einer bestimmten Produktseite oder Version oder zu einem bestimmten Thema geben Sie zwischen der Kurz-URL und der Produkt-ID ein Hashzeichen (#) an. Die Kurz-URL für die gesamte Dokumentation zu InfoSphere Information Server lautet beispielsweise wie folgt:

<http://ibm.biz/knowctr#SSZJPZ/>

Die folgende URL ist die leicht verkürzte URL zum obigen Thema (das Symbol \Rightarrow gibt eine Zeilenfortsetzung an):

http://ibm.biz/knowctr#SSZJPZ_11.3.0/com.ibm.swg.im.iis.common.doc/common/accessingiidoc.html

Hilfelinks so ändern, dass sie auf lokal installierte Dokumentation verweisen

Das IBM Knowledge Center enthält die aktuelle Version der Dokumentation. Sie können jedoch eine lokale Version der Dokumentation in Form eines Information Center installieren und Ihre Hilfelinks so konfigurieren, dass sie auf dieses Information Center verweisen. Ein lokales Information Center ist sinnvoll, wenn Ihr Unternehmen keinen Zugriff auf das Internet bereitstellt.

Befolgen Sie die Installationsanweisungen im Installationspaket für das Information Center, um das Information Center auf einem Computer Ihrer Wahl zu installieren. Nach der Installation und dem Start des Information Center können Sie die von der Taste F1 der Produkthilfe und den Hilfelinks verwendete Speicherposition für die Dokumentation mithilfe des Befehls **iisAdmin** auf der Serviceschicht ändern (das Symbol \Rightarrow gibt eine Zeilenfortsetzung an):

Windows

```
IS-Installationspfad\ASBServer\bin\iisAdmin.bat -set -key  $\Rightarrow$   
com.ibm.iis.infocenter.url -value http://<Host>:<Port>/help/topic/
```

AIX Linux

```
IS-Installationspfad/ASBServer/bin/iisAdmin.sh -set -key  $\Rightarrow$   
com.ibm.iis.infocenter.url -value http://<Host>:<Port>/help/topic/
```

Dabei ist <Host> der Name des Computers, auf dem das Information Center installiert ist, und <Port> ist die Portnummer für das Information Center. Die Standardportnummer lautet 8888. Für einen Computer mit dem Namen `server1.example.com`, der den Standardport verwendet, lautet der URL-Wert beispielsweise `http://server1.example.com:8888/help/topic/`.

PDF- und Hardcopy-Dokumentation abrufen

- Die PDF-Versionen der Handbücher sind online verfügbar und können über <https://www.ibm.com/support/docview.wss?uid=swg27008803&wv=1> aufgerufen werden.
- Sie können IBM Veröffentlichungen auch im Hardcopy-Format online oder über den zuständigen IBM Ansprechpartner bestellen. Wenn Sie Veröffentlichungen online bestellen möchten, rufen Sie das IBM Publications Center unter <http://www.ibm.com/e-business/linkweb/publications/servlet/pbi.wss> auf.

Anhang D. Feedback zur Produktdokumentation geben

Sie können hilfreiches Feedback zur IBM Dokumentation geben.

Ihr Feedback hilft IBM, Informationen von hoher Qualität bereitzustellen. Sie haben die folgenden Möglichkeiten, Kommentare zu senden:

- Wenn Sie einen Kommentar zu einem Thema im IBM Knowledge Center abgeben wollen, das sich auf der IBM Website befindet, melden Sie sich an und klicken Sie auf die Schaltfläche **Kommentar hinzufügen** am Ende des Themas. Auf diese Weise eingegebene Kommentare werden für alle Benutzer angezeigt.
- Wenn Sie einen Kommentar zu einem Thema im IBM Knowledge Center an IBM senden wollen, der nicht für alle Benutzer angezeigt werden soll, melden Sie sich an und klicken auf den Link **Feedback** unten im IBM Knowledge Center.
- Senden Sie Ihre Kommentare mithilfe des Onlineformulars für Leserkommentare unter www.ibm.com/software/awdtools/rcf/.
- Senden Sie Ihre Kommentare per E-Mail an comments@us.ibm.com. Geben Sie dabei den Namen des Produkts, die Versionsnummer des Produkts sowie den Namen und die Teilenummer der Informationen (falls vorhanden) an. Wenn sich Ihr Kommentar auf einen bestimmten Text bezieht, geben Sie die Position des Texts an (z. B. eine Überschrift, eine Tabellenummer oder eine Seitenzahl).

Bemerkungen und Marken

Die vorliegenden Informationen wurden für Produkte und Services entwickelt, die auf dem deutschen Markt angeboten werden. IBM stellt dieses Material möglicherweise auch in anderen Sprachen zur Verfügung. Für den Zugriff auf das Material in einer anderen Sprache ist eine Kopie des Produkts oder der Produktversion in der jeweiligen Sprache erforderlich.

Bemerkungen

Möglicherweise bietet IBM die in dieser Dokumentation beschriebenen Produkte, Services oder Funktionen in anderen Ländern nicht an. Informationen über die gegenwärtig im jeweiligen Land verfügbaren Produkte und Services sind beim zuständigen IBM Ansprechpartner erhältlich. Hinweise auf IBM Lizenzprogramme oder andere IBM Produkte bedeuten nicht, dass nur Programme, Produkte oder Services von IBM verwendet werden können. Anstelle der IBM Produkte, Programme oder Services können auch andere, ihnen äquivalente Produkte, Programme oder Services verwendet werden, solange diese keine gewerblichen oder anderen Schutzrechte von IBM verletzen. Die Verantwortung für den Betrieb von Produkten, Programmen und Services anderer Anbieter liegt beim Kunden.

Für in diesem Handbuch beschriebene Erzeugnisse und Verfahren kann es IBM Patente oder Patentanmeldungen geben. Mit der Auslieferung dieses Handbuchs ist keine Lizenzierung dieser Patente verbunden. Lizenzanforderungen sind schriftlich an folgende Adresse zu richten (Anfragen an diese Adresse müssen auf Englisch formuliert werden):

IBM Director of Licensing
IBM Europe, Middle East & Africa
Tour Descartes
2, avenue Gambetta
92066 Paris La Defense
France

Trotz sorgfältiger Bearbeitung können technische Ungenauigkeiten oder Druckfehler in dieser Veröffentlichung nicht ausgeschlossen werden. Die hier enthaltenen Informationen werden in regelmäßigen Zeitabständen aktualisiert und als Neuausgabe veröffentlicht. IBM kann ohne weitere Mitteilung jederzeit Verbesserungen und/oder Änderungen an den in dieser Veröffentlichung beschriebenen Produkten und/oder Programmen vornehmen.

Verweise in diesen Informationen auf Websites anderer Anbieter werden lediglich als Service für den Kunden bereitgestellt und stellen keinerlei Billigung des Inhalts dieser Websites dar. Das über diese Websites verfügbare Material ist nicht Bestandteil des Materials für dieses IBM Produkt. Die Verwendung dieser Websites geschieht auf eigene Verantwortung.

Werden an IBM Informationen eingesandt, können diese beliebig verwendet werden, ohne dass eine Verpflichtung gegenüber dem Einsender entsteht.

Lizenznehmer des Programms, die Informationen zu diesem Produkt wünschen mit der Zielsetzung: (i) den Austausch von Informationen zwischen unabhängig voneinander erstellten Programmen und anderen Programmen (einschließlich des

vorliegenden Programms) sowie (ii) die gemeinsame Nutzung der ausgetauschten Informationen zu ermöglichen, wenden sich an folgende Adresse:

IBM Corporation
J46A/G4
555 Bailey Avenue
San Jose, CA 95141-1003 USA

Die Bereitstellung dieser Informationen kann unter Umständen von bestimmten Bedingungen - in einigen Fällen auch von der Zahlung einer Gebühr - abhängig sein.

Die Lieferung des in diesem Dokument beschriebenen Lizenzprogramms sowie des zugehörigen Lizenzmaterials erfolgt auf der Basis der IBM Rahmenvereinbarung bzw. der Allgemeinen Geschäftsbedingungen von IBM, der IBM Internationalen Nutzungsbedingungen für Programmpakete oder einer äquivalenten Vereinbarung.

Alle in diesem Dokument enthaltenen Leistungsdaten stammen aus einer kontrollierten Umgebung. Die Ergebnisse, die in anderen Betriebsumgebungen erzielt werden, können daher erheblich von den hier erzielten Ergebnissen abweichen. Einige Daten stammen möglicherweise von Systemen, deren Entwicklung noch nicht abgeschlossen ist. Eine Gewährleistung, dass diese Daten auch in allgemein verfügbaren Systemen erzielt werden, kann nicht gegeben werden. Darüber hinaus wurden einige Daten unter Umständen durch Extrapolation berechnet. Die tatsächlichen Ergebnisse können davon abweichen. Benutzer dieses Dokuments sollten die entsprechenden Daten in ihrer spezifischen Umgebung prüfen.

Alle Informationen zu Produkten anderer Anbieter stammen von den Anbietern der aufgeführten Produkte, deren veröffentlichten Ankündigungen oder anderen allgemein verfügbaren Quellen. IBM hat diese Produkte nicht getestet und kann daher keine Aussagen zu Leistung, Kompatibilität oder anderen Merkmalen machen. Fragen zu den Leistungsmerkmalen von Produkten anderer Anbieter sind an den jeweiligen Anbieter zu richten.

Aussagen über Pläne und Absichten von IBM unterliegen Änderungen oder können zurückgenommen werden und repräsentieren nur die Ziele von IBM.

Diese Veröffentlichung dient nur zu Planungszwecken. Die in dieser Veröffentlichung enthaltenen Informationen können geändert werden, bevor die beschriebenen Produkte verfügbar sind.

Diese Veröffentlichung enthält Beispiele für Daten und Berichte des alltäglichen Geschäftsablaufs. Sie sollen nur die Funktionen des Lizenzprogramms illustrieren und können Namen von Personen, Firmen, Marken oder Produkten enthalten. Alle diese Namen sind frei erfunden; Ähnlichkeiten mit tatsächlichen Namen und Adressen sind rein zufällig.

COPYRIGHTLIZENZ:

Diese Veröffentlichung enthält Beispielanwendungsprogramme, die in Quellsprache geschrieben sind und Programmier Techniken in verschiedenen Betriebsumgebungen veranschaulichen. Sie dürfen diese Beispielprogramme kostenlos kopieren, ändern und verteilen, wenn dies zu dem Zweck geschieht, Anwendungsprogramme zu entwickeln, zu verwenden, zu vermarkten oder zu verteilen, die mit der Anwendungsprogrammierschnittstelle für die Betriebsumgebung konform sind, für die diese Beispielprogramme geschrieben werden. Diese Beispiele wurden nicht

unter allen denkbaren Bedingungen getestet. Daher kann IBM die Zuverlässigkeit, Wartungsfreundlichkeit oder Funktion dieser Programme weder zusagen noch gewährleisten. Die Beispielprogramme werden ohne Wartung (auf "as-is"-Basis) und ohne jegliche Gewährleistung zur Verfügung gestellt. IBM übernimmt keine Haftung für Schäden, die durch die Verwendung der Beispielprogramme entstehen.

Kopien oder Teile der Beispielprogramme bzw. daraus abgeleiteter Code müssen folgenden Copyrightvermerk beinhalten:

© (Name Ihrer Firma) (Jahr). Teile des vorliegenden Codes wurden aus Beispielprogrammen der IBM Corporation abgeleitet. © Copyright IBM Corp. _Jahr/Jahre angeben_. Alle Rechte vorbehalten.

Hinweise zur Datenschutzrichtlinie

IBM Softwareprodukte, einschließlich Software as a Service-Lösungen ("Softwareangebote"), können Cookies oder andere Technologien verwenden, um Informationen zur Produktnutzung zu erfassen, die Endbenutzererfahrung zu verbessern und Interaktionen mit dem Endbenutzer anzupassen oder zu anderen Zwecken. In vielen Fällen werden von den Softwareangeboten keine personenbezogenen Daten erfasst. Einige der IBM Softwareangebote können Sie jedoch bei der Erfassung personenbezogener Daten unterstützen. Wenn dieses Softwareangebot Cookies zur Erfassung personenbezogener Daten verwendet, sind nachfolgend nähere Informationen über die Verwendung von Cookies durch dieses Angebot zu finden.

Abhängig von den implementierten Konfigurationen kann dieses Softwareangebot Sitzungscookies oder persistente Cookies verwenden. Wird ein Produkt oder eine Komponente nicht aufgelistet, verwendet dieses Produkt bzw. diese Komponente keine Cookies.

Tabelle 6. Verwendung von Cookies durch Produkte und Komponenten von InfoSphere Information Server/InfoSphere Information Server

Produktmodul	Komponente oder Feature	Typ des verwendeten Cookies	Erfasste Daten	Zweck der Daten	Inaktivierung des Cookies
Beliebig (Bestandteil der InfoSphere Information Server-Installation)	InfoSphere Information Server-Webkonsole	<ul style="list-style-type: none"> • Sitzung • Persistent 	Benutzername	<ul style="list-style-type: none"> • Sitzungsmanagement • Authentifizierung 	Kann nicht inaktiviert werden
Beliebig (Bestandteil der InfoSphere Information Server-Installation)	InfoSphere Metadata Asset Manager	<ul style="list-style-type: none"> • Sitzung • Persistent 	Keine personenbezogenen Daten	<ul style="list-style-type: none"> • Sitzungsmanagement • Authentifizierung • Besserer Bedienungskomfort • SSO-Konfiguration 	Kann nicht inaktiviert werden
InfoSphere DataStage	Stage 'Big Data File'	<ul style="list-style-type: none"> • Sitzung • Persistent 	<ul style="list-style-type: none"> • Benutzername • Digitale Signatur • Sitzungs-ID 	<ul style="list-style-type: none"> • Sitzungsmanagement • Authentifizierung • SSO-Konfiguration 	Kann nicht inaktiviert werden
InfoSphere DataStage	Stage 'XML'	Sitzung	Interne IDs	<ul style="list-style-type: none"> • Sitzungsmanagement • Authentifizierung 	Kann nicht inaktiviert werden

Tabella 6. Verwendung von Cookies durch Produkte und Komponenten von InfoSphere Information Server/InfoSphere Information Server (Forts.)

Produktmodul	Komponente oder Feature	Typ des verwendeten Cookies	Erfasste Daten	Zweck der Daten	Inaktivierung des Cookies
InfoSphere DataStage	IBM InfoSphere DataStage and QualityStage Operations Console	Sitzung	Keine personbezogenen Daten	<ul style="list-style-type: none"> • Sitzungsmanagement • Authentifizierung 	Kann nicht inaktiviert werden
InfoSphere Data Click	InfoSphere Information Server-Webkonsole	<ul style="list-style-type: none"> • Sitzung • Persistent 	Benutzername	<ul style="list-style-type: none"> • Sitzungsmanagement • Authentifizierung 	Kann nicht inaktiviert werden
InfoSphere Data Quality Console		Sitzung	Keine personbezogenen Daten	<ul style="list-style-type: none"> • Sitzungsmanagement • Authentifizierung • SSO-Konfiguration 	Kann nicht inaktiviert werden
InfoSphere QualityStage Standardization Rules Designer	InfoSphere Information Server-Webkonsole	<ul style="list-style-type: none"> • Sitzung • Persistent 	Benutzername	<ul style="list-style-type: none"> • Sitzungsmanagement • Authentifizierung 	Kann nicht inaktiviert werden
InfoSphere Information Governance Catalog		<ul style="list-style-type: none"> • Sitzung • Persistent 	<ul style="list-style-type: none"> • Benutzername • Interne IDs • Status der Baumstruktur 	<ul style="list-style-type: none"> • Sitzungsmanagement • Authentifizierung • SSO-Konfiguration 	Kann nicht inaktiviert werden
InfoSphere Information Analyzer	Stage 'Data Rules' im InfoSphere DataStage and QualityStage Designer-Client	Sitzung	Sitzungs-ID	Sitzungsmanagement	Kann nicht inaktiviert werden

Wenn die für dieses Softwareangebot genutzten Konfigurationen Sie als Kunde in die Lage versetzen, personenbezogene Daten von Endbenutzern über Cookies und andere Technologien zu erfassen, müssen Sie sich zu allen gesetzlichen Bestimmungen in Bezug auf eine solche Datenerfassung, einschließlich aller Mitteilungspflichten und Zustimmungsanforderungen, rechtlich beraten lassen.

Weitere Informationen zur Nutzung verschiedener Technologien, einschließlich Cookies, für diese Zwecke finden Sie in der "IBM Online-Datenschutzerklärung, Schwerpunkte" unter <http://www.ibm.com/privacy>, in der "IBM Online-Datenschutzerklärung" unter <http://www.ibm.com/privacy/details> im Abschnitt "Cookies, Web-Beacons und sonstige Technologien" und in "IBM Software Products and Software-as-a-Service Privacy Privacy Statement" unter <http://www.ibm.com/software/info/product-privacy>.

Marken

IBM, the IBM Logo und [ibm.com](http://www.ibm.com) sind Marken oder eingetragene Marken der International Business Machines Corporation. Weitere Produkt- und Servicennamen können Marken von IBM oder anderen Unternehmen sein. Eine aktuelle Liste der IBM Marken finden Sie im Web unter <http://www.ibm.com/legal/copytrade.shtml>.

Die folgenden Namen sind Marken oder eingetragene Marken anderer Unternehmen:

Adobe ist eine eingetragene Marke der Adobe Systems Incorporated in den USA und/oder anderen Ländern.

Intel und Itanium sind Marken oder eingetragene Marken der Intel Corporation oder ihrer Tochtergesellschaften in den USA und/oder anderen Ländern.

Linux ist eine Marke von Linus Torvalds in den USA und/oder anderen Ländern.

Microsoft, Windows und Windows NT sind Marken der Microsoft Corporation in den USA und/oder anderen Ländern.

UNIX ist eine eingetragene Marke von The Open Group in den USA und anderen Ländern.

Java[™] und alle auf Java basierenden Marken und Logos sind Marken oder eingetragene Marken der Oracle Corporation und/oder ihrer verbundenen Unternehmen.

United States Postal Service ist Inhaber der folgenden Marken: CASS, CASS Certified, DPV, LACS^{Link}, ZIP, ZIP + 4, ZIP Code, Post Office, Postal Service, USPS und United States Postal Service. Die IBM Corporation ist ein nicht ausschließlicher Lizenznehmer für DPV und LACS^{Link}.

Weitere Unternehmens-, Produkt- oder Servicenamen können Marken anderer Hersteller sein.

Index

B

Bemerkungen 43

K

Kundenunterstützung
kontaktieren 37

M

Marken
Liste 43

P

Produktdokumentation
Zugriff auf 39
Produkteingabehilfen
Eingabehilfen 35

R

Regelsatzerweiterungen
Lernprogramm für Adressdaten 1

S

Software-Services
kontaktieren 37
Standardization Rules Designer
Lernprogramm für Adressdaten 1

U

Unterstützung
Kundenunterstützung 37



SC43-1124-00

