

IBM InfoSphere DataStage and QualityStage
Version 11 Release 3

Lernprogramm für parallele Jobs



IBM InfoSphere DataStage and QualityStage
Version 11 Release 3

Lernprogramm für parallele Jobs



Hinweis

Vor Verwendung dieser Informationen und des darin beschriebenen Produkts sollten die Informationen unter „Bemerkungen und Marken“ auf Seite 77 gelesen werden.

Diese Veröffentlichung ist eine Übersetzung des Handbuchs
IBM InfoSphere DataStage and QualityStage, Version 11 Release 3, DataStage Parallel Job Tutorial,
IBM Form SC19-4284-00,
herausgegeben von International Business Machines Corporation, USA

© Copyright International Business Machines Corporation 2006, 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 2006, 2014.

Inhaltsverzeichnis

Kapitel 1. Lernprogramm: Parallele Jobs erstellen 1

Kapitel 2. Umgebung des Lernprogramms für parallele Jobs einrichten . . . 3

Ordner für die Lernprogrammdateien erstellen	3
Projekt für das Lernprogramm für parallele Jobs erstellen	3
Quellendateien für das Lernprogramm für parallele Jobs kopieren	4
Komponenten für das Lernprogramm für parallele Jobs importieren	4
Datenbanktabelle für das Lernprogramm für parallele Jobs erstellen	5
Datenquellennamen für die Tabelle des Lernprogramms für parallele Jobs erstellen	6
Datenquellennamen für die Lernprogrammtabelle auf einem Microsoft Windows-Computer erstellen	6
Datenquellennamen für die Lernprogrammtabelle auf einem AIX-, HP-UX-, Solaris- oder Linux-Computer erstellen	6

Kapitel 3. Modul 1: Beispieljob öffnen und ausführen 13

Lerneinheit 1.1: Beispieljob öffnen	13
Vorgehensweise	14
Lerneinheit 1.2: Beispieljob anzeigen und kompilieren	15
Stage 'Sequential File' untersuchen	15
Stage 'Data Set' untersuchen	16
Beispieljob kompilieren	17
Lerneinheit 1.3: Beispieljob ausführen	17
Job ausführen	18
Ausgabe anzeigen	18
Weitere Schritte	18

Kapitel 4. Modul 2: Ihren ersten Job entwerfen 21

Lerneinheit 2.1: Job erstellen	21
Neuen Job erstellen	22
Stages hinzufügen und verlinken	22
Lerneinheit 2.2: Metadaten angeben	23
Spaltenmetadaten angeben	24
Lookup-Eigenschaften angeben und Job ausführen	25
Lerneinheit 2.3: Metadaten importieren	26
Metadaten in Ihr Repository importieren	27
Spaltenmetadaten aus dem Repository laden	27
Lerneinheit 2.4: Parameter hinzufügen	29
Jobparameter definieren	29
Ihrem Job Jobparameter hinzufügen	30
Job mit Parametern ausführen	31
Lerneinheit 2.5: Parametersatz erstellen	31
Vorgehensweise	32

Weitere Schritte	33
----------------------------	----

Kapitel 5. Modul 3: Daten umsetzen . . . 35

Lerneinheit 3.1: Umsetzungsjob entwerfen	35
Umsetzungsjob erstellen	36
'Data Set'-Stages konfigurieren	36
Stage 'Transformer' konfigurieren	37
Umsetzungsjob ausführen	40
Lerneinheit 3.2: Stage 'Lookup' im Umsetzungsjob verwenden	41
Lookup-Job erstellen	41
Stage 'Lookup File Set' konfigurieren	43
Stage 'Lookup' konfigurieren	43
Lerneinheit 3.3: Zurückgewiesene Daten erfassen	45
Vorgehensweise	45
Lerneinheit 3.4: Mehrere Umsetzungen in einem einzigen Job abschließen	47
Umsetzungs- und Lookup-Operationen hinzufügen	47
Umsetzungslogik zum Job hinzufügen	48
Lookup-Operation konfigurieren	52
Weitere Schritte	54

Kapitel 6. Modul 4: Daten in eine relationale Datenbank laden 55

Lerneinheit 4.1: Datenverbindungsobjekt erstellen	55
Übersicht	55
Datenverbindungsobjekt erstellen	56
Lerneinheit 4.2: Spaltenmetadaten aus einer Datenbanktabelle importieren	57
Übersicht	57
Vorgehensweise	57
Lerneinheit 4.3: Daten in eine Datenbank schreiben	58
Übersicht	58
Job erstellen	59
Ziel konfigurieren	59
ODBC-Connector konfigurieren	60
Weitere Schritte	61

Anhang A. Eingabehilfen in den Produkten	63
Anhang B. Kontaktaufnahme mit IBM	65
Anhang C. Auf Produktdokumentation zugreifen.	67
Anhang D. Eingabehilfen in den Produkten	69
Anhang E. Kontaktaufnahme mit IBM	71
Anhang F. Auf Produktdokumentation zugreifen.	73
Anhang G. Feedback zu Produktdokumentation geben	75
Bemerkungen und Marken	77
Index	83

Kapitel 1. Lernprogramm: Parallele Jobs erstellen

In diesem Lernprogramm entwickeln Sie mithilfe von InfoSphere DataStage Jobs zum Extrahieren, Umsetzen und Laden von Daten. Durch die Umsetzung und Bereinigung der Quelldaten und die Anwendung einer konsistenten Formatierung wird eine bessere Datenqualität erreicht.

In diesem Szenario gehen die weltweit vertretenen Unternehmen GlobalCo und WorldCo eine Fusion ein. Da sie ähnliche Geschäftstätigkeiten ausüben, haben die beiden Unternehmen einige gemeinsame Kunden. Das fusionierte Unternehmen GlobalCo Worldwide möchte ein Data-Warehouse für seine Lieferungs- und Rechnungsstellungsinformationen erstellen. Die Übungen in diesem Lernprogramm konzentrieren sich auf einen kleinen Teil der Arbeit, die zur Erreichung dieses Ziels notwendig ist.

Ihre Aufgabe bei diesem Projekt ist die Arbeit an den Daten (einschließlich der Rechnungsstellungsdatensätze für Kundendaten) des Unternehmens GlobalCo. Sie lesen diese Daten aus Dateien mit durch Kommas getrennten Werten (CSV-Dateien) und werden die Daten anschließend bereinigen und umsetzen, um sie für das Zusammenfassen mit den entsprechenden Daten von WorldCo vorzubereiten. Diese Daten bilden im fertig gestellten Data-Warehouse die Dimensionstabelle GlobalCo_billing. Ein anderer Entwickler fasst diese Dimensionstabelle mit der Dimensionstabelle WorldCo_billing zusammen, um die Rechnungsstellungsinformationen für GlobalCo Worldwide zu erstellen.

Dieses Lernprogramm führt Sie durch die Tasks, die Sie zum Extrahieren, Umsetzen und Laden der Rechnungsstellungsdaten für GlobalCo ausführen. In den nachfolgenden kurz zusammengefassten Schritten sind die Aktionen in der Reihenfolge aufgelistet, in der Sie sie ausführen:

1. In Modul 1 öffnen Sie den Job `samplejob` und untersuchen die einzelnen Stages, die Bestandteil des Jobs sind. Es ist wichtig, dass Sie mit der Funktionsweise der Stages vertraut sind, bevor Sie mit dem Entwurf Ihres eigenen Jobs beginnen. Außerdem erfahren Sie, wie Sie den Job kompilieren, den Job ausführen und die generierte Ausgabe anzeigen.
2. In Modul 2 erstellen Sie Ihren ersten Job durch Hinzufügen von Stages und Links zum InfoSphere DataStage und QualityStage Designer-Entwicklungsbereich. Anschließend erstellen Sie durch Importieren von Metadaten Tabellendefinitionen, die Sie im gesamten weiteren Verlauf des Lernprogramms verwenden. Darüber hinaus erstellen Sie Parameter und Parametersätze, die Sie wiederverwenden, um Ihren Jobentwurf zu vereinfachen und die jobübergreifende Wiederverwendung zu fördern.
3. In Modul 3 entwerfen Sie einen Umsetzungsjob, der die Rechnungsstellungsdaten von GlobalCo bereinigt, damit sie mit den Rechnungsstellungsdaten von WorldCo zusammengefasst werden können. Anschließend erweitern Sie diesen Job, indem Sie ihm mehrere Umsetzungen hinzufügen, mit dem Ziel, auf die Rechnungsstellungsdaten eine strengere Datentypisierung anzuwenden.
4. In Modul 4 laden Sie die bereinigten Rechnungsstellungsdaten von GlobalCo in eine relationale Datenbank, um die Daten anderen Entwicklern im Unternehmen GlobalCo Worldwide zugänglich zu machen. Sie erstellen ein Datenverbindungsobjekt, importieren Spaltenmetadaten aus einer Datenbanktabelle und schreiben die Ausgabe anschließend in eine vorhandene Tabelle in der Datenbank.

Lernziele

Während Sie das Lernprogramm durcharbeiten, lernen Sie, wie Sie die folgenden Tasks ausführen:

- Entwerfen von parallelen Jobs, mit denen Daten extrahiert, umgesetzt und geladen werden
- Erstellen wiederverwendbarer Objekte, die in andere Jobentwürfe aufgenommen werden können
- Modifizieren Ihres Jobentwurfs zur Implementierung einer strengeren Datentypisierung
- Ausführen der von Ihnen entworfenen Jobs und Anzeigen der Ergebnisse

Erforderlicher Zeitaufwand

Bevor Sie mit dem Lernprogramm beginnen, stellen Sie sicher, dass die in Kapitel 2, „Umgebung des Lernprogramms für parallele Jobs einrichten“, auf Seite 3 aufgeführten Schritte von Ihrem InfoSphere DataStage und QualityStage-Administrator ausgeführt wurden. Der zur Installation und Konfiguration des Lernprogramms erforderliche Zeitaufwand ist von Ihrer InfoSphere DataStage-Umgebung abhängig.

Für die Ausführung dieses Lernprogramms benötigen Sie etwa vier Stunden. Wenn Sie weitere Konzepte in Zusammenhang mit diesem Lernprogramm kennenlernen möchten, benötigen Sie möglicherweise etwas mehr Zeit.

Systemvoraussetzungen

Für dieses Lernprogramm ist die folgende Hardware und Software erforderlich:

- InfoSphere DataStage-Clients, die auf einer Windows-Plattform installiert sind
- Eine Verbindung zu einem InfoSphere DataStage-Server, der auf einer Windows- oder UNIX-Plattform ausgeführt wird

Tipp: Windows-Server können sich auf demselben Computer befinden wie die Clients

Voraussetzungen

Dieses Lernprogramm richtet sich an InfoSphere DataStage-Entwickler mit wenig Vorkenntnissen, die sich mit der Vorgehensweise zur Erstellung von parallelen Jobs vertraut machen wollen. Kenntnisse der grundlegenden Konzepte von InfoSphere DataStage wie Jobs, Stages und Links sind hilfreich, jedoch nicht erforderlich.

Bemerkungen: GlobalCo, WorldCo und GlobalCo Worldwide stellen fiktive Geschäftsvorgänge mit Beispieldaten dar, mit denen Beispielanwendungen für IBM® und IBM Kunden entwickelt 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 manuell oder vom System generierte fiktive Daten, aus wissenschaftlichen oder öffentlichen Quellen zusammengestellte Fakten sowie Daten enthalten, die mit Zustimmung der Copyrightinhaber als Beispieldaten zur Entwicklung von Beispielanwendungen genutzt werden dürfen. Referenzierte Produktnamen können Marken der jeweiligen Rechtsinhaber sein. Nicht autorisiertes Kopieren dieser Daten ist unzulässig.

Kapitel 2. Umgebung des Lernprogramms für parallele Jobs einrichten

Bevor Sie das Lernprogramm starten können, muss Ihr IBM InfoSphere DataStage und QualityStage Administrator Ordner erstellen, das Projekt für das Lernprogramm erstellen, Quellendateien importieren und andere Einrichtungstasks ausführen.

Vorbereitende Schritte

Zum Ausführen einiger der Installations- und Einrichtungstasks benötigen Sie InfoSphere DataStage und QualityStage-Zugriffsrechte und Windows-Administratorrechte. Für die Installations- und Einrichtungsschritte benötigen Sie umfangreichere Kenntnisse über das System und die Datenbank als für die Durchführung des Lernprogramms erforderlich sind.

Ordner für die Lernprogrammdateien erstellen

Erstellen Sie auf Ihrem InfoSphere DataStage-Client-Computer einen Ordner und kopieren Sie die Dateien von der Installations-CD in den Ordner.

Vorgehensweise

1. Navigieren Sie zum Ordner `\TutorialData\DataStage\parallel_tutorial`, der sich auf den Installationsmedien befindet. Der Ordner befindet sich möglicherweise in einem Verzeichnis `is-client` oder `is-suite`.
2. Erstellen Sie auf Ihrem Computer einen neuen Ordner, beispielsweise `C:\tutorial`).
3. Kopieren Sie den Ordner `\TutorialData\DataStage\parallel_tutorial` von den Installationsmedien in den Ordner, den Sie auf dem Client-Computer erstellt haben.

Projekt für das Lernprogramm für parallele Jobs erstellen

Erstellen Sie ein neues Projekt für das Lernprogramm, um die Übungen des Lernprogramms von anderen InfoSphere DataStage-Projekten getrennt zu halten.

Vorbereitende Schritte

Das Projekt für das Lernprogramm erstellen Sie mithilfe des IBM InfoSphere DataStage und QualityStage Administrator-Clients. Zum Erstellen des Projekts für das Lernprogramm benötigen Sie ausreichende Berechtigungen für dieses Tool.

Vorgehensweise

1. Wählen Sie **Start > Programme > IBM InfoSphere Information Server > IBM InfoSphere DataStage und QualityStage Administrator** aus.
2. Geben Sie Ihren Benutzernamen und das zugehörige Kennwort in das Fenster **Verbindung herstellen** ein.
3. Klicken Sie im Fenster **Administrator-Client** auf die Registerkarte **Projekte**.
4. Klicken Sie auf der Seite **Projekte** auf **Hinzufügen**.
5. Geben Sie in das Feld **Name** die Bezeichnung `Tutorial` als Namen des neuen Projekts ein und klicken Sie dann auf **OK**.

6. Klicken Sie auf **Schließen**, um den Administrator-Client zu schließen.

Quellendateien für das Lernprogramm für parallele Jobs kopieren

Sie haben Quellendateien für das Lernprogramm auf Ihren Computer heruntergeladen. Diese Dateien müssen jedoch für den Computer, der als Host für die Engineschicht fungiert, auch zugänglich sein. Kopieren Sie die Datendateien des Lernprogramms aus dem Verzeichnis `Tutorial` auf dem Computer, der als Host für die Clientschicht fungiert, in das Projektverzeichnis auf dem Computer, der als Host für die Engineschicht fungiert.

Informationen zu diesem Vorgang

Bei der Erstellung des Projekts `Tutorial` wurde auf dem Computer, der als Host für die Engineschicht fungiert, ein Verzeichnis generiert. Die Engineschicht befindet sich möglicherweise auf demselben Windows-Computer, der als Host für die Clientschicht fungiert, sie kann sich jedoch auch auf einem separaten Computer befinden. Stellen Sie fest, wo das Verzeichnis für das Projekt `Tutorial` erstellt wurde, damit Sie die Quellendateien in dieses Verzeichnis kopieren können.

Vorgehensweise

1. Öffnen Sie das Verzeichnis, in dem Sie die Quellendateien für das Lernprogramm extrahiert haben. Diese Datei enthält drei `.csv`-Dateien, drei `.ddl`-Dateien und eine `.dsx`-Datei.
2. Navigieren Sie zum Verzeichnis `Projects` und öffnen Sie das Verzeichnis `Tutorial` für Ihr Projekt. Die folgenden Projektverzeichnisse (Verzeichnisse `Projects`) sind Standardprojektverzeichnisse bei InfoSphere DataStage.

Option	Bezeichnung
Microsoft Windows	C:\IBM\InformationServer\Server\Projects\
AIX, HP-UX, Solaris, Linux	/opt/IBM/InformationServer/Server/Projects/

3. Erstellen Sie in diesem Verzeichnis ein neues Verzeichnis mit dem Namen `Tutorial_files`.
4. Kopieren Sie den Inhalt des Verzeichnisses `Tutorial` auf dem Computer, der als Host für die Clientschicht fungiert, in das Verzeichnis `Tutorial_files` auf dem Computer, der als Host für die Engineschicht fungiert.

Komponenten für das Lernprogramm für parallele Jobs importieren

Importieren Sie über den Designer-Client den Beispieljob und die Beispieltabellendefinitionen.

Vorgehensweise

1. Wählen Sie **Start > Programme > IBM InfoSphere Information Server > IBM InfoSphere DataStage und QualityStage Designer** aus.
2. Geben Sie Ihren Benutzernamen und das zugehörige Kennwort in das Fenster **Verbindung herstellen** ein.
3. Wählen Sie das Projekt für das Lernprogramm aus der Liste **Projekt** aus und klicken Sie anschließend auf **OK**.
4. Klicken Sie im Fenster **Neu** auf **Abbrechen**, um es zu schließen. Sie können dieses Fenster schließen, da Sie einen vorhandenen Job öffnen.

5. Klicken Sie in der Designer-Client-Menüleiste auf **Importieren > DataStage-Komponenten**.
6. Geben Sie *Tutorial_files\parallel_tutorial.dsx* in das Feld **Aus Datei importieren** des Fensters **DataStage-Repository importieren** ein. **Tutorial_files** ist der Name des Ordners, den Sie auf dem Computer erstellt haben, der als Host für die Engineschicht fungiert.
7. Klicken Sie auf das Optionsfeld **Alles importieren**.
8. Klicken Sie auf **OK**. Der Designer-Client importiert den Beispieljob und die Beispieltabellendefinitionen in das von Ihnen ausgewählte Verzeichnis.
9. Klicken Sie auf **Datei > Beenden**, um den Designer-Client zu schließen.

Datenbanktabelle für das Lernprogramm für parallele Jobs erstellen

Erstellen Sie eine Tabelle in der relationalen Datenbank, die als für das Metadatenrepository für InfoSphere DataStage fungiert.

Vorbereitende Schritte

Wenn Sie eine Zieldatenbanktabelle auf einem Microsoft Windows-Computer erstellen, müssen Sie über Administratorrechte verfügen oder Befehle als Administrator ausführen können.

Führen Sie den folgenden Befehl aus, um eine Administratorsitzung zu starten:

```
runas /User:Administrator cmd.exe
```

Informationen zu diesem Vorgang

InfoSphere DataStage verwendet ein Repository, für das eine relationale Datenbank als Host fungiert. Im Modul 4 des Lernprogramms werden Metadaten aus einer Tabelle in eine relationale Datenbank importiert und die Daten anschließend in die Tabelle geschrieben. Sie verwenden DDL-Skripts (DDL = Data Definition Language, Datendefinitionssprache) zum Erstellen der Datenbanktabelle.

Vorgehensweise

1. Erstellen Sie eine neue Datenbank mit dem Namen Tutorial. Diese Datenbank muss im codierten Zeichensatz UTF-8 erstellt werden. Die Datenbankerstellungsskripts schlagen fehl, wenn Sie sie für eine Datenbank ausführen, die nicht auf dem codierten Zeichensatz UTF-8 basiert.
2. Stellen Sie eine Verbindung zur neuen Datenbank her.
3. Navigieren Sie im Verzeichnis InfoSphere DataStage Projects zum Verzeichnis Tutorial_files.
4. Führen Sie das entsprechende DDL-Skript aus, um die Lernprogrammtabelle in der neuen Datenbank zu erstellen.

Datenbanktyp	Script
IBM DB2	DB2_table.ddl
Oracle	Oracle_table.ddl
Microsoft SQL Server	SQLserver_table.ddl

Datenquellennamen für die Tabelle des Lernprogramms für parallele Jobs erstellen

Erstellen Sie einen Datenquellennamen (DSN - Data Source Name) für die Datenbanktabelle, sodass Benutzer des Lernprogramms für parallele Jobs über eine ODBC-Verbindung auf die Tabelle zugreifen können.

Datenquellennamen für die Lernprogrammtabelle auf einem Microsoft Windows-Computer erstellen

Sie definieren den Datenquellennamen auf dem Computer definieren, auf dem die IBM InfoSphere Information Server-Engine installiert ist. Benutzer des Lernprogramms für parallele Jobs verwenden diesen Datenquellennamen im Designer-Client für die Verbindung zur Lernprogrammtabelle.

Vorbereitende Schritte

Zum Ausführen dieser Task benötigen Sie auf dem Windows-Computer Administratorrechte.

Vorgehensweise

1. Öffnen Sie die Systemsteuerung und wählen Sie **Verwaltung** aus.
2. Wählen Sie **Datenquellen (ODBC)** aus.
3. Klicken Sie auf die Registerkarte **System-DSN**.
4. Klicken Sie auf der Seite **System-DSN** auf **Hinzufügen**.
5. Wählen Sie im Fenster **Neue Datenquelle erstellen** einen Treiber für die Datenbank aus und klicken Sie dann auf **Fertig stellen**.

Wichtig: Wenn Ihr Datenbanktreiber nicht aufgelistet ist, müssen Sie ihn installieren.

6. Geben Sie im Fenster für den Datenbanktreiber, das geöffnet wird, die Details an, die der Treiber benötigt, um eine Verbindung zur Datenbank des Lernprogramms herzustellen. Wenn Sie für die Datenbank einen Benutzernamen und ein Kennwort angegeben haben, dann geben Sie auch diese Verbindungsdetails an.
7. Schließen Sie das Fenster für den Datenbanktreiber und klicken Sie dann auf **OK**, um das Fenster **ODBC-Datenquellen-Administrator** zu schließen.

Datenquellennamen für die Lernprogrammtabelle auf einem AIX-, HP-UX-, Solaris- oder Linux-Computer erstellen

Sie definieren den Datenquellennamen auf dem Computer definieren, auf dem die IBM InfoSphere Information Server-Engine installiert ist. Benutzer des Lernprogramms für parallele Jobs verwenden diesen Datenquellennamen im Designer-Client für die Verbindung zur Lernprogrammtabelle.

Vorbereitende Schritte

Beim Einrichten eines Datenquellennamens auf einem AIX-, HP-UX-, Solaris- oder Linux-Computer bearbeiten Sie drei Dateien:

- dsenv
- odbc.ini
- uvodbc.config

Welche Einträge Sie in den einzelnen Dateien vornehmen, hängt von Ihrem Betriebssystem und vom Datenbanktyp ab, zu dem eine Verbindung hergestellt werden soll.

Datei dsenv konfigurieren

Für einige ODBC-Verbindungen, Plug-ins und Connector sowie für Interaktionen mit externen Anwendungen wie IBM WebSphere MQ müssen Sie Umgebungsvariablen hinzufügen, um die interaktive Verwendung von ODBC-Treibern zur Herstellung einer Verbindung zu einer ODBC-Datenquelle hinzuzufügen.

Vorbereitende Schritte

Sie müssen auf dem Server für den InfoSphere DataStage-Administrator als InfoSphere DataStage-Administrator mit den Berechtigungsnachweisen für das Betriebssystem angemeldet sein.

Führen Sie ein Backup der Datei dsenv durch, bevor Sie die Datei bearbeiten.

Informationen zu diesem Vorgang

Die Datei dsenv enthält eine Reihe von Shellargumenten, die beim Starten der Engine verwendet werden. Die Datei dsenv kann von interaktiven Benutzern, anderen Programmen und Scripts verwendet werden. Für eine Verbindung, die einen Verbindungsprotokolltreiber verwendet, müssen Sie an der Datei dsenv keine Modifizierungen vornehmen.

Vorgehensweise

1. Bearbeiten Sie die Datei dsenv. Die Datei befindet sich im Verzeichnis \$DSHOME/DSEngine, wobei \$DSHOME für das InfoSphere DataStage-Installationsverzeichnis steht. Das Standardverzeichnis ist /opt/IBM/InformationServer/Server/DSEngine.
2. Geben Sie die folgenden Informationen in der Datei dsenv an:
 - Umgebungsvariablen für die Datenbank-Client-Software
 - Basisposition der Datenbank
 - Verzeichnis der Datenbankbibliothek

Tabelle 1. Namen der Umgebungsvariablen für den Bibliothekspfad nach Betriebssystem

Betriebssystem	Umgebungsvariable für Bibliothekspfad
AIX	LIBPATH
HP-UX	SHLIB_PATH
HP-UX auf Intel Itanium	LD_LIBRARY_PATH
Linux	LD_LIBRARY_PATH
Solaris	LD_LIBRARY_PATH

Die folgenden Beispiele stellen typische Einträge bei häufig verwendeten Datenbanken dar. Die Einträge variieren geringfügig je nach Betriebssystem. Weitere Informationen finden Sie in der Dokumentation zu den Datenquellen.

Sybase 11

```
LANG=  
export LANG  
SYBASE=/export/home/sybase/sybase  
export SYBASE  
LD_LIBRARY_PATH=$LD_LIBRARY_PATH:$SYBASE/lib:/usr/lib:/lib  
export LD_LIBRARY_PATH
```

Informix XPS 9.3

```
INFORMIXDIR=/space/informix  
export INFORMIXDIR  
INFORMIXSERVER=solnew2.1  
export INFORMIXSERVER  
ONCONFIG=onconfig.solnew2  
export ONCONFIG  
LD_LIBRARY_PATH=$LD_LIBRARY_PATH:$INFORMIXDIR/lib:$INFORMIXDIR/lib/  
esql:$INFORMIXDIR/lib/cli  
export LD_LIBRARY_PATH  
LANG=C  
export LANG
```

DB2 Version 10.5

```
DB2DIR=/opt/IBM/db2/V10.5  
export DB2DIR  
DB2INSTANCE=db2inst1  
export DB2INSTANCE  
INSTHOME=/export/home/db2inst1  
export INSTHOME  
PATH=$PATH:$INSTHOME/sql/lib/bin:$INSTHOME/sql/lib/adm:  
$INSTHOME/sql/lib/misc  
export PATH  
LD_LIBRARY_PATH=$LD_LIBRARY_PATH:$INSTHOME/sql/lib/lib  
export LD_LIBRARY_PATH  
THREADS_FLAG=native  
export THREADS_FLAG
```

3. Speichern Sie Ihre Änderungen.
4. Stoppen Sie die IBM InfoSphere Information Server-Engine und starten Sie sie erneut.

Datei odbc.ini konfigurieren

Die Datei `odbc.ini` enthält Informationen zum Herstellen einer Verbindung zu den von Ihnen verwendeten Datenbanken und Datenbank Anwendungen. Sie müssen diese Datei konfigurieren, damit InfoSphere DataStage auf ODBC-Datenquellen zugreifen kann.

Informationen zu diesem Vorgang

Im Verzeichnis `/opt/IBM/InformationServer/Server/branded_odbc/IBM_Tools` befinden sich `odbc.ini`-Beispieldateien für die einzelnen unterstützten Plattformen. Verwenden Sie beim Erstellen der Datei `odbc.ini` für das Verzeichnis `$DSHOME` den Inhalt dieser Dateien. Die Variable `$DSHOME` gibt das Installationsverzeichnis von InfoSphere DataStage an.

Vorgehensweise

1. Navigieren Sie zum Verzeichnis `<Installationsverzeichnis>/Server/DSEngine`, wobei `<Installationsverzeichnis>` für das Verzeichnis steht, in dem InfoSphere Information Server installiert wurde.
2. Führen Sie den folgenden Befehl aus, um für die Beispieldatei `.odbc.ini` einen auf die Datei `odbc.ini` verweisenden Softlink zu erstellen.

```
ln -s odbc.ini.odbc.ini
```

3. Modifizieren Sie die Datei `.odbc.ini` so, dass Sie Ihrer Datenbankumgebung entspricht.

Datei `uvodbc.config` konfigurieren

In der Datei `uvodbc.config` geben Sie den Namen der ODBC-Datenquelle für jede Datenbank an.

Informationen zu diesem Vorgang

Eine Kopie der Datei `uvodbc.config` befindet sich jeweils im Verzeichnis `$DSHOME` und in jedem Projektverzeichnis. Bei Projekten unter Linux beispielsweise wird `/opt/IBM/InformationServer/Server/Projects` als Standardpfad verwendet.

IBM InfoSphere DataStage sucht standardmäßig im aktuellen Projektverzeichnis nach einer Datei `uvodbc.config` und verwendet diese Datei, sofern vorhanden, anstelle der unter `$DSHOME` vorhandenen Datei. Diese Dateianordnung ermöglicht es Ihnen, Datenquellen zu konfigurieren, die einigen Projekten bekannt sind und anderen nicht.

Vorgehensweise

Verwenden Sie beim Hinzufügen von Einträgen zur Datei `uvodbc.config` die folgende Syntax:

```
<DSN-Name>  
DBMSTYPE = ODBC
```

DSN-Name ist der ODBC-Datenquellennamen (DSN - Data Source Name). Dieser Name muss mit dem in der Datei `.odbc.ini` angegebenen Namen übereinstimmen. Das Leerzeichen auf beiden Seiten des Gleichheitszeichens ist erforderlich.

In der Datei `uvodbc.config` beginnen Kommentarzeilen mit einem Nummernzeichen (`#`). Kommentare, die nicht im Kommentarblock am Anfang der Datei positioniert werden, führen zu Verarbeitungsfehlern. Fügen Sie Kommentare ausschließlich im Kommentarblock hinzu, der sich oben in der Datei `uvodbc.config` befindet. Nachfolgend ist ein typisches Beispiel einer Datei `uvodbc.config` dargestellt:

```
[ODBC DATA SOURCES]  
<localuv>  
DBMSTYPE = UNIVERSE  
network = TCP/IP  
service = uvserver  
host = 127.0.0.1  
<Sybase1>  
DBMSTYPE = ODBC  
<Sybase2>  
DBMSTYPE = ODBC  
<Oracle8>  
DBMSTYPE = ODBC  
<Informix>  
DBMSTYPE = ODBC  
<DB2>  
DBMSTYPE = ODBC
```

DB2-Paket erstellen und an einen Datenquellennamen binden

Damit ein DB2-Paket verwendet werden kann, müssen Sie dieses an die einzelnen DB2-Datenquellennamen (DSN - Data Source Name) binden.

Vorgehensweise

1. Öffnen Sie ein Terminalfenster und wechseln Sie in das Verzeichnis `$DSHOME/./branded_odbc/bin`. Die Variable `$DSHOME` gibt das Installationsverzeichnis von InfoSphere DataStage an. Das Standardverzeichnis ist `/opt/IBM/InformationServer/Server/DSEngine`.
2. Geben Sie den folgenden Befehl ein:

```
bind22 DSN
```

DSN ist der Name der ODBC-Datenquelle.
3. Geben Sie auf eine entsprechende Eingabeaufforderung hin Ihre Benutzer-ID und Ihr Kennwort an.

Andere ODBC-Treiber konfigurieren

Sie können die Verbindung zu Datenquellen mithilfe von ODBC-Treibern anderer Hersteller herstellen. Da allerdings nur ein einziger ODBC-Manager verwendet werden kann, können Treiber anderer Hersteller nicht verwendet werden, wenn die InfoSphere DataStage-Treiber verwendet werden.

Vorgehensweise

1. Installieren Sie den ODBC-Manager und die ODBC-Treiber. Verwenden Sie die vom Treiberhersteller bereitgestellte Dokumentation.
2. Konfigurieren Sie den Computer für die Verbindung zu dem Treiber.
 - a. Fügen Sie der Datei `dsenv` die erforderlichen Umgebungsvariablen hinzu.
 - b. Konfigurieren Sie in der Datei `odbc.ini` die Verbindung zur Datenbank oder zur Datenbankclientanwendung. Sie müssen den Datenquellennamen (DSN - Data Source Name), den Treibernamen und die Anmeldeinformationen angeben
 - c. Geben Sie in der Datei `uvodbc.config` den Datenquellennamen für die Verbindung zur Datenbank an.
3. Fügen Sie die gemeinsam genutzte Bibliothek `libodbc.sl` für den ODBC-Manager in den Pfad der InfoSphere DataStage-Ladebibliothek ein. (Unter HP-UX lautet der Suffix `so`.) Bearbeiten Sie die Datei `dsenv` und stellen Sie sicher, dass in der Umgebungsvariablen für den Pfad der gemeinsam genutzten Bibliothek der Pfad zum ODBC-Manager vor dem Pfad `$DSHOME/./branded_odbc/lib` steht.
4. Testen Sie nach der Installation und Konfiguration des Treibers die Konnektivität.

Ergebnisse

Sollen stattdessen wieder die DataDirect-Treiber verwendet werden, entfernen Sie den Pfad zum ODBC-Manager aus der Datei `dsenv` und stellen Sie sicher, dass die ODBC-Treiberbibliothek für InfoSphere DataStage vor allen ODBC-Treiberbibliotheken steht.

ODBC-Konnektivität testen

Nach dem Konfigurieren der für InfoSphere DataStage definierten ODBC-Konnektivität testen Sie die Konfiguration, um Ihre Verbindung zur Datenquelle zu verifizieren.

Vorgehensweise

1. Ändern Sie das Verzeichnis in `$DSHOME` und richten Sie die Umgebung durch Ausführen des folgenden Befehls ein:

```
./dsenv
```


2. Starten Sie die Engine-Shell für IBM InfoSphere Information Server, indem Sie den folgenden Befehl eingeben:

```
./bin/dssh
```

3. Melden Sie sich durch Eingabe des folgenden Befehls am Projekt an:

```
LOGTO Projektname
```

Bei dem Projektnamen muss die Groß-/Kleinschreibung beachtet werden.

4. Rufen Sie eine Liste der verfügbaren Datenquellennamen ab, indem Sie den folgenden Befehl eingeben:

```
DS_CONNECT
```

5. Testen Sie die Verbindung, indem Sie den folgenden Befehl eingeben:

```
DS_CONNECT Datenquellenname
```

Dabei steht *Datenquellenname* für den Datenquellennamen der zu testenden Verbindung.

6. Geben Sie Ihren Benutzernamen und Ihr Kennwort ein, um eine Verbindung zur angegebenen Datenbank herzustellen.
7. Geben Sie `.Q` ein, um die Verbindung zu beenden.

Kapitel 3. Modul 1: Beispieljob öffnen und ausführen

Das Lernprogramm enthält einen Beispieljob, den Sie untersuchen, um die grundlegenden Konzepte kennen zu lernen, auf denen Jobs basieren. Sie öffnen den Beispieljob, untersuchen die Stages, die Bestandteil des Jobs sind, kompilieren anschließend den Job und führen ihn aus.

Der Beispieljob besteht aus zwei Stages: eine Stage **Sequential File** (Sequenzielle Datei) und eine Stage **Data Set**. Die Stage **Sequential File** extrahiert Daten aus einer durch Kommas getrennten Wertedatei und schreibt die Daten in die Stage **Data Set**. Die Daten, die in die Stage **Data Set** geschrieben werden, werden von nachfolgenden Modulen im Lernprogramm verwendet.

Lernziele

Nach dem Durcharbeiten der Lerneinheiten in diesem Modul werden Sie in der Lage sein, folgende Tasks auszuführen:

- Starten von InfoSphere DataStage und QualityStage Designer (Designer-Client) und Herstellen einer Verbindung zu einem Projekt
- Öffnen eines vorhandenen Jobs
- Kompilieren eines Jobs, sodass er ausgeführt werden kann
- Öffnen von InfoSphere DataStage und QualityStage Director (Director-Client) und Ausführen eines Jobs
- Anzeigen der Ergebnisse des Jobs

Anmerkung: Im weiteren Verlauf des Lernprogramms wird über den Kurznamen auf die InfoSphere DataStage-Komponenten verwiesen, mit denen Sie interagieren. So wird InfoSphere DataStage und QualityStage Designer beispielsweise als Designer-Client bezeichnet.

Erforderlicher Zeitaufwand

Für dieses Modul sind ungefähr 30 Minuten erforderlich.

Voraussetzungen

Stellen Sie sicher, dass Sie über InfoSphere DataStage und QualityStage-Benutzerberechtigungen verfügen.

Lerneinheit 1.1: Beispieljob öffnen

In dieser Lerneinheit zeigen Sie einen abgeschlossenen Job an, um sich mit seiner Struktur und mit dem Designer-Client vertraut zu machen.

Als Einstieg in die Arbeit mit Designer-Client öffnen Sie zuerst den Job `samplejob`, der in diesem Lernprogramm enthalten ist.

Weitere Informationen zum Designer-Client:

Der Designer-Client ist Ihre Workbench und Toolbox zum Erstellen von Jobs. Die enthaltenen Tools verwenden Sie zum Erstellen von Jobs, die Daten extrahieren,

umsetzen und laden sowie die Qualität von Daten überprüfen. Die Palette enthält die Tools, die die Grundbausteine eines Jobs bilden. Zum Entwerfen von Jobs fügen Sie die folgenden Objekte aus der Palette Ihrem Entwicklungsbereich hinzu:

- *Stages* stellen eine Verbindung zu Datenquellen her, um Dateien zu lesen oder zu schreiben und Daten zu verarbeiten
- *Links* verbinden die Stages, durch die Ihre Daten fließen
- *Anmerkungen* stellen Informationen zu den von Ihnen erstellten Jobs bereit

Der Designer-Client verwendet ein Repository, in dem Sie die von Ihnen erstellten Objekte im Rahmen des Entwurfsprozesses speichern. Der Beispieljob ist ein Objekt im Repository, das in diesem Lernprogramm enthalten ist. Der Beispieljob verwendet eine Tabellendefinition, die ebenfalls ein Objekt im Repository ist. Sie speichern weitere Metadaten im Repository wie Jobparameter, Parametersätze und gemeinsam genutzte Container. Diese Objekte können von anderen Jobentwicklern wiederverwendet werden, wodurch die unternehmensweite Wiederverwendung gefördert wird.

Übersicht

Der Beispieljob liest Daten aus einer unstrukturierten Datei und schreibt die Ausgabe in ein Dataset. Parallele Jobs verwenden Datasets, um Daten zu speichern, während die Daten verarbeitet werden. Datasets können nicht resident und für Sie, den Entwickler, verborgen sein. Sie können jedoch auch persistente Datasets erstellen. Der Beispieljob schreibt Daten in ein persistentes Dataset, das einen internen Zwischenspeicherungsbereich bereitstellt, in dem die Daten aufbewahrt werden, bis sie in einem nachfolgenden Modul in ihren endgültigen Bestimmungsort geschrieben werden. Beim Entwerfen von Jobs ist es nicht zwingend erforderlich, dass Sie einen Zwischenspeicherungsbereich für Ihre Daten erstellen. Im Rahmen dieses Lernprogramms wird jedoch ein Dataset dazu verwendet, die Daten des einen Moduls im jeweils nächsten bereitzustellen.

Die in diesem Job verwendeten Daten sind die Rechnungsstellungsinformationen des Unternehmens GlobalCo. Diese Daten verwenden Sie im weiteren Verlauf dieses Lernprogramms als Eingabe für Ihren Umsetzungsjob. Die Daten enthalten Inkonsistenzen und Formatierungsfehler, die Sie mithilfe von Datentypisierungs- und Umsetzungsfunktionen beheben. Die bereinigten Daten bilden die Dimension GlobalCo_billing des Sternschemas 'GlobalCo Worldwide'. Dieses Sternschema wird von einem anderen Entwickler mit dem Ziel erstellt, die Rechnungsstellungsinformationen von GlobalCo und die von WorldCo zu kombinieren.

Vorgehensweise

1. Klicken Sie auf **Start > Programme > IBM InfoSphere Information Server > IBM InfoSphere DataStage und QualityStage Designer**.
2. Geben Sie Ihren Benutzernamen und das zugehörige Kennwort in das Fenster **Verbindung herstellen** ein.
3. Wählen Sie das Projekt für das Lernprogramm aus der Liste **Projekt** aus und klicken Sie anschließend auf **OK**.
4. Klicken Sie im Fenster **Neu** das nach dem Starten des Designer-Clients geöffnet wird, auf **Abbrechen**, um dieses zu schließen und einen vorhandenen Job zu öffnen.
5. Erweitern Sie in der Repository-Baumstruktur den Ordner **Tutorial** und doppelklicken Sie auf den Job **samplejob**. Alle für das Lernprogramm erforderli-

chen Objekte befinden sich in diesem Ordner. Im weiteren Verlauf des Lernprogramms werden Sie weitere Ordner zum Speichern Ihrer Jobs, Parametersätze und Tabellendefinitionen erstellen.

Der Job wird im Entwicklungsbereich des Designer-Clients geöffnet. Die folgende Abbildung zeigt den Job `samplejob`.

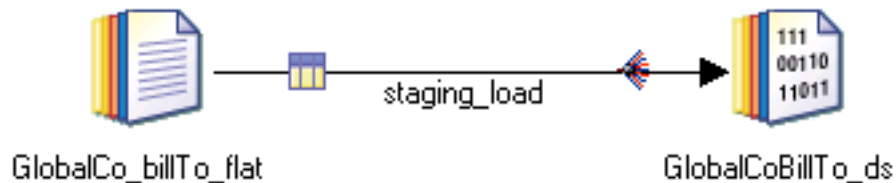


Abbildung 1. Der Job `samplejob` besteht aus zwei Stages, die über einen einzelnen Link miteinander verbunden sind.

In dieser Lerneinheit haben Sie Ihren ersten Job geöffnet. Nun, nachdem der Job offen ist, können Sie damit beginnen, die Stages zu untersuchen, die Bestandteil des Jobs sind.

Lerneinheit 1.2: Beispieljob anzeigen und kompilieren

In dieser Lerneinheit untersuchen Sie die einzelnen Stages im Job `samplejob`, um den ihm zugrunde liegenden Entwurf kennen zu lernen. Anschließend kompilieren Sie den Job, um ihn für die Ausführung auf Ihrem System vorzubereiten.

Übersicht

Der Job `samplejob` verwendet eine Stage **Sequential File** (Sequenzielle Datei), um Daten aus einer unstrukturierten Datei zu lesen, und eine Stage **Data Set**, um Daten in den Zwischenspeichersbereich zu schreiben. Die beiden Stages sind über einen Link miteinander verknüpft, über den die Daten von einer Stage an die andere übergeben werden. Die Metadaten, die zwischen den beiden Stages über den Link fließen, wurden beim Entwurf des Jobs definiert. Wenn Sie den Job ausführen, fließen die Metadaten den Link entlang.

Tasks in dieser Lerneinheit

1. „Stage 'Sequential File' untersuchen“
2. „Stage 'Data Set' untersuchen“ auf Seite 16
3. „Beispieljob kompilieren“ auf Seite 17

Stage 'Sequential File' untersuchen

Öffnen Sie die Stage **Sequential File**, um die Stage-Eigenschaften kennen zu lernen. Die **Sequential File**-Stage `GlobalCo_billTo_flat` bildet die Quelle. Sie gibt die Quellendatei für den Job `samplejob` an und definiert die Spalten, die an die **Data Set**-Stage `GlobalCoBillTo_ds` gesendet werden.

1. Doppelklicken Sie im Beispieljob auf die **Sequential File**-Stage 'GlobalCo_billTo_flat'.

Der Stage-Editor wird auf der Registerkarte **Eigenschaften** der Seite **Ausgabe** geöffnet. Alle Stages für parallele Jobs verfügen über eine Registerkarte **Eigenschaften**, auf der Sie die Aktionen angeben, die während der Ausführung des Jobs ausgeführt werden sollen.

2. Sehen Sie sich die Eigenschaft `Datei` unter der Kategorie **Quelle** an.

Über diese Eigenschaft geben Sie die Datei an, die von der Stage bei der Ausführung des Jobs gelesen wird. In dem Beispieljob verweist die Eigenschaft Datei auf eine Datei GlobalCo_BillTo.csv. Sie geben das Verzeichnis, das diese Datei enthält, während der Ausführung des Jobs an.

Der Name des Verzeichnisses wird als Jobparameter mit dem Namen #tutorial_directory# definiert. Jobparameter werden verwendet, damit variable Daten (z. B. Datei- oder Verzeichnisname) während der Ausführung angegeben werden können, was die Entwicklung und Verarbeitung vereinfacht. Jobparameter werden mithilfe von Nummernzeichen (#) angegeben.

3. Klicken Sie auf die Eigenschaft Spaltennamen in erster Zeile unter der Kategorie **Optionen**.

In dem Beispieljob ist diese Eigenschaft auf WAHR gesetzt, da die erste Zeile der Datei GlobalCo_BillTo.csv die Namen der Spalten in der Datei enthält. Für die übrigen Eigenschaften werden die Standardwerte verwendet.

4. Klicken Sie auf die Registerkarte **Format**.

Die Registerkarte **Format** ähnelt der Registerkarte **Eigenschaften**, jedoch beschreiben die von Ihnen festgelegten Eigenschaften das Format der unstrukturierten Datei, die von der Stage gelesen wird. Die Stage GlobalCo_billTo_flat verwendet eine CSV-Datei mit durch Kommas getrennten Werten (Begrenzer=Komma), d. h., dass jedes Feld innerhalb einer Zeile durch ein Komma getrennt ist. Der Wert Datensatzbegrenzerzeichenfolge=DOS-Format gibt an, dass die Datei auf AIX-, HP-UX- und Solaris-Systemen gelesen werden kann.

5. Klicken Sie auf die Registerkarte **Spalten**.

Auf der Registerkarte **Spalten** werden die Spaltenmetadaten für die Stage definiert. Die Spaltenmetadaten definieren die Daten, die den Link entlang zur **Data Set**-Stage GlobalCoBillTo_ds fließen. Die Datei GlobalCo_BillTo.csv enthält zahlreiche Spalten. Alle diese Spalten weisen den Datentyp VarChar auf. Während Sie das Lernprogramm durcharbeiten, wenden Sie eine strengere Datentypisierung auf diese Spalten an, um die Daten zu bereinigen.

6. Klicken Sie im Fenster des Stage-Editors oben rechts auf **Daten anzeigen**.

- a. Geben Sie im Feld **Wert** des Fensters für die Auflösung des Jobparameters den Namen des Verzeichnisses an, in das die Lernprogrammdateien importiert wurden, und klicken Sie auf **OK**. Sie müssen den Verzeichnispfad angeben, wenn Sie Daten in einer Stage anzeigen oder einen Job ausführen.

- b. Klicken Sie im Fenster **Datenbrowser** auf **OK**.

Es wird ein Fenster geöffnet, in dem die ersten 100 Zeilen der Daten aus der Datei GlobalCo_BillTo.csv angezeigt werden. Jede Zeile stellt einen Kundendatensatz dar, der eine Kundennummer, einen Namen, eine Adresse, Kontaktinformationen und wichtige Vertriebsdaten enthält.

Diese Datensätze schreiben Sie in die Zieldatei, die in der **Data Set**-Stage GlobalCoBillTo_ds angegeben ist.

- c. Klicken Sie auf **Schließen**, um das Fenster **Datenbrowser** zu schließen.

7. Klicken Sie auf **OK**, um den Editor der Stage **Sequential File** zu schließen.

Stage 'Data Set' untersuchen

Öffnen Sie die Stage **Data Set**, um die Stage-Eigenschaften kennen zu lernen. Die **Data Set**-Stage GlobalCoBillTo_ds bildet das Ziel und gibt die Datei an, in die der Datensatz geschrieben wird.

1. Doppelklicken Sie auf die **Data Set**-Stage GlobalCoBillTo_ds.

Der Stage-Editor wird auf der Registerkarte **Eigenschaften** der Seite **Eingabe** geöffnet. **Data Set**-Stages werden verwendet, um Daten zu speichern, die von

einem anderen Job verwendet werden. Datasets (gekennzeichnet durch die Erweiterung `.ds`) entsprechen dem internen Format für die Übertragung von Daten innerhalb paralleler Jobs.

2. Klicken Sie auf die Eigenschaft `Datei` unter der Kategorie **Ziel**.

Diese Eigenschaft verweist auf die Datei `GlobalCo_BillTo.ds`. Dies ist die Ausgabedatei, in die die Stage während der Ausführung Daten schreibt. Das Verzeichnis, in dem diese Datei gespeichert werden soll, geben Sie während der Ausführung des Jobs an. Wie die vorherige Stage verwendet auch diese Stage einen Parameter für das Verzeichnis, in dem sich die Lernprogrammdateien befinden.

3. Klicken Sie auf die Registerkarte **Spalten**.

Die Spaltenmetadaten für diese Stage und die Spaltenmetadaten für die Stage 'Sequential File' sind identisch, da Sie die Daten vor dem Schreiben keiner Umsetzung unterziehen. Die Registerkarte **Spalten** gibt an, dass die Daten bei der Ausführung des Jobs in die Datei `GlobalCo_BillTo.ds` geschrieben werden.

Der Editor der Stage **Data Set** verfügt nicht über die Registerkarte **Format**, da für das Dataset keine Formatierung erforderlich ist. Wenn Sie auf die Schaltfläche **Daten anzeigen** klicken, wird eine Nachricht angezeigt, die besagt, dass keine Daten vorhanden sind. Die Daten werden erstellt, nachdem Sie den Job kompiliert und ausgeführt haben.

4. Klicken Sie auf **OK**, um die Editor der Stage **Data Set** zu schließen.

Beispieljob kompilieren

Kompilieren Sie den Job `samplejob`, um diesen ausführen zu können.

1. Klicken Sie in der Designer-Client-Menüleiste auf **Datei > Kompilieren**.

Während der Jobkompilierung wird das Fenster **Kompilierungsjob** mit Nachrichten vom Compiler aktualisiert.

2. Wenn im Fenster **Kompilierungsjob** die Nachricht angezeigt wird, dass der Job kompiliert wurde, klicken Sie auf **OK**.

In dieser Lerneinheit haben Sie einen einfachen Datenextraktionsjob untersucht, der Daten aus einer unstrukturierten Datei liest und in ein Dataset schreibt. Nachdem der Job kompiliert wurde, können Sie ihn ausführen und die Ausgabe anzeigen.

Lerneinheit 1.3: Beispieljob ausführen

In der vorherigen Lerneinheit haben Sie den Beispieljob kompiliert. Nun, nachdem der Job kompiliert ist, können Sie mithilfe des Director-Clients den Job ausführen und das Jobprotokoll anzeigen.

Übersicht

Den Job führen Sie über den Director-Client aus. Der Director-Client ist eine Betriebskonsole, mit deren Hilfe Sie die in Designer-Client entwickelten Jobs ausführen und in diesen eine Fehlerbehebung durchführen. Darüber hinaus führen Sie über den Director-Client vollständig entwickelte Jobs in der Produktionsumgebung aus.

Nachdem Sie den Job ausgeführt haben, überprüfen Sie das Jobprotokoll auf Fehler und, wenn Fehler vorhanden sind, beheben Sie diese.

Tasks in dieser Lerneinheit

1. „Job ausführen“ auf Seite 18

2. „Ausgabe anzeigen“

Job ausführen

Bisher haben Sie mit dem Designer-Client gearbeitet. Zum Ausführen des Jobs `samplejob` verwenden Sie den Director-Client.

1. Klicken Sie in der Designer-Client-Menüleiste auf **Tools > Director ausführen**.
Da Sie über den Designer-Client am Projekt Tutorial angemeldet sind, können Sie den Director-Client starten, ohne Ihre InfoSphere DataStage-Berechtigungs-nachweise erneut eingeben zu müssen.
Im Director-Client hat der Beispieljob den Status **Kompiliert**, was bedeutet, dass der Job ausgeführt werden kann.
2. Wählen Sie den Job `samplejob` im rechten Teilfenster des Director-Clients aus und klicken Sie in der Menüleiste auf **Job > Jetzt ausführen**.
 - a. Geben Sie im Fenster **Jobausführungsoptionen** den Pfad Ihres Projekts an. Beispiel: `C:\IBM\InformationServer\Server\Projects\Tutorial`.
 - b. Klicken Sie auf **Ausführen**. Der Jobstatus ändert sich in **Aktiv**.
3. Wenn sich der Jobstatus in **Fertiggestellt** ändern, klicken Sie auf **Anzeigen > Protokoll**.
4. Sehen Sie sich im Jobprotokoll die Informationen an, die vom Director-Client während der Ausführung eines Jobs gemeldet werden.
Das Jobprotokoll enthält die Zeitmarke mit der Zeitangabe, zu der jedes Ereignis aufgetreten ist, den Ereignistyp und in welcher Stage oder welchem Operator das Ereignis aufgetreten ist. Da der Job erfolgreich ausgeführt wurde, handelt es sich bei den Nachrichten um Steuernachrichten oder Informationsnachrichten. Sind Fehler aufgetreten, enthält das Jobprotokoll möglicherweise Warnnachrichten mit einer Beschreibung der Fehler. Wenn ein Job fehlschlägt, wird zur Erläuterung des Fehlschlagens eine Abbruchnachricht in das Jobprotokoll geschrieben.
5. Klicken Sie auf **Projekt > Beenden**, um den Director-Client zu schließen.

Ausgabe anzeigen

Nachdem Sie den Job `samplejob` ausgeführt haben, zeigen Sie die Ausgabe an, um sicherzustellen, dass die Zieldaten in dem erwarteten Format vorliegen.

1. Öffnen Sie den Job `samplejob` im Designer-Client.
2. Öffnen Sie die Stage `GlobalCoBillTo_ds`.
3. Klicken Sie im Stage-Editor auf **Daten anzeigen**.
4. Klicken Sie im Fenster **Datenbrowser** auf **OK**, um die Standardeinstellungen zu akzeptieren. Es wird ein Fenster geöffnet, in dem bis zu 100 Zeilen der in das Dataset geschriebenen Daten angezeigt werden.
Die Daten im Dataset wurden im Verlaufe der Jobausführung nicht geändert. Sie haben die Daten in ein Dataset geschrieben, um diese Daten im weiteren Verlauf des Lernprogramms als Eingabe verwenden zu können. Die Daten enthalten Inkonsistenzen und Fehler, die Sie in Ihrem Umsetzungsjob beheben.
5. Klicken Sie auf **Schließen**, um das Fenster **Datenbrowser** zu schließen.
6. Klicken Sie auf **OK**, um die Stage `GlobalCoBillTo_ds` zu schließen.

Weitere Schritte

Sie haben den Job `samplejob` ausgeführt und die Ausgabedaten angezeigt. Nachdem Sie nun die Vorgehensweise beim Ausführen eines Datenextraktionsjobs kennen, können Sie jetzt mit dem Erstellen eines eigenen Jobs beginnen. Im nächsten

Modul werden Sie durch den Prozess der Erstellung eines weiteren Extraktionsjobs geführt. Sie erstellen zwei Lookup-Dateigruppen für die GlobalCo-Landescodeinformationen und Codes für Sonderbehandlung. Diese Lookup-Tabellen verwenden Sie in einem nachfolgenden Modul, um die GlobalCo-Kundendatensätze zu erweitern.

In diesem Modul haben Sie die folgenden Tasks ausgeführt:

- Starten des Designer-Clients
- Öffnen des Jobs `samplejob` und Lokalisieren von Objekten in der Repository-Baumstruktur
- Untersuchen von Stages und Anzeigen der Daten, die von den einzelnen Stages dargestellt werden
- Starten des Director-Clients über den Designer-Client
- Ausführen des Jobs `samplejob` und Anzeigen des Jobprotokolls
- Untersuchen der Ausgabedaten, die in das Dataset `GlobalCo_BillTo.ds` geschrieben wurden

Kapitel 4. Modul 2: Ihren ersten Job entwerfen

Sie haben gelernt, wie der Beispieljob geöffnet, kompiliert und ausgeführt wird. Da dieser Job jedoch bereits vordefiniert war, ist es nun an der Zeit, dass Sie lernen, Ihren eigenen Job zu entwerfen.

Der Job, den Sie in diesem Lernprogramm entwerfen, liest zwei unstrukturierte Dateien und füllt zwei Lookup-Tabellen. Das Unternehmen GlobalCo Worldwide möchte sich einen Überblick darüber verschaffen, wo die einzelnen Kunden leben und welche Sprache sie sprechen, um einen besseren Einblick in die Demografie seiner Kunden zu gewinnen. Daher erstellen Sie eine Lookup-Tabelle, die neben einem Code, der die Sprache der einzelnen Kunden angibt, auch einen Code enthält, der angibt, aus welchem Land ein Kunde stammt.

Ferner möchte GlobalCo Worldwide in Erfahrung bringen, in welchen Intervallen die Kunden Lieferungen erhalten. Erkenntnisse über Häufigkeit und Umfang der Lieferungen bieten GlobalCo Worldwide eine bessere Grundlage, auf der sie ihre Lieferoperationen optimieren und die Produktion an Zeiten mit hohem Lieferbedarf anpassen können. Diese Daten binden Sie ein, indem Sie eine zweite Lookup-Tabelle erstellen, die einen Code für Sonderbehandlung enthält und die in Modul 3 verwendet wird. Dieser Code entspricht einer Textbeschreibung, die das Intervall angibt, in dem die einzelnen Kunden eine Lieferung erhalten.

Lernziele

Nach dem Durcharbeiten der Lerneinheiten in diesem Modul werden Sie in der Lage sein, die folgenden Tasks auszuführen:

- Hinzufügen von Stages und Links zu einem Job
- Angeben von Eigenschaften für Stages und Links, um Ihre Funktionsweise beim Ausführen des Jobs festzulegen
- Importieren von Metadaten für Spalten
- Hinzufügen von Jobparametern und Erstellen von Parametersätzen
- Erweitern Ihrer Kenntnisse im Kompilieren und Ausführen von Jobs

Erforderlicher Zeitaufwand

Für dieses Modul sind ungefähr 90 Minuten erforderlich.

Lerneinheit 2.1: Job erstellen

Wenn Sie einen Job entwerfen, erstellen Sie zunächst einen Job und speichern ihn in einem Verzeichnis in Ihrem Projekt.

Übersicht

Sie erstellen einen parallelen Job und speichern ihn in einem neuen Ordner im Ordner Tutorial in der Repository-Baumstruktur. Ein Job besteht aus miteinander verlinkten Stages, die den Datenfluss von einer Datenquelle zu einem Datenziel beschreiben. Eine Stage ist eine grafische Darstellung der Daten selbst oder einer Umsetzung, die an den Daten durchgeführt wird.

Nachdem Sie Ihren Job erstellt haben, fügen Sie Stages hinzu und verlinken diese.

Tasks in dieser Lerneinheit

1. „Neuen Job erstellen“
2. „Stages hinzufügen und verlinken“

Neuen Job erstellen

Erstellen Sie im Designer-Client einen neuen Job und speichern Sie ihn in einem Verzeichnis in Ihrem Projekt.

1. Melden Sie sich beim Designer-Client an.
2. Klicken Sie in der Menüleiste auf **Datei > Neu**.
3. Wählen Sie im Fenster **Neu** im linken Teilfenster den Ordner **Jobs** aus und klicken Sie dann im rechten Teilfenster auf **Paralleler Job**.
4. Klicken Sie auf **OK**. Ein leerer Jobentwicklungsbereich wird geöffnet.
5. Klicken Sie auf **Datei > Speichern**.

- a. Klicken Sie im Fenster **Parallelen Job speichern unter** mit der rechten Maustaste auf den Ordner **Tutorial** und klicken Sie auf **Neu > Ordner** im Direktaufrufmenü.
- b. Geben Sie dem neuen Ordner den Namen **GCWW tutorial**, wobei **GCWW** für 'GlobalCo Worldwide' steht.
- c. Klicken Sie mit der rechten Maustaste auf den Ordner, den Sie erstellt haben, und klicken Sie auf **Neu > Ordner**.
- d. Geben Sie dem neuen Ordner den Namen **Jobs**. Sie speichern alle Jobs, die Sie erstellen, in diesem Ordner, um sie von der Arbeit anderer Entwickler getrennt zu halten.

Stellen Sie sicher, dass Ihre Ordnerstruktur im Repository der folgenden Hierarchie entspricht:

```
[ - ] Tutorial
    [ - ] GCWW tutorial
        [ + ] Jobs
```

- e. Geben Sie **Dev_CustData_Codes_Lookup** in das Feld **Elementname** ein. Dieser Name gibt an, dass es sich bei diesem Job um einen Entwicklungsjob handelt, der ein Lookup für Kundendaten enthält.

Tipp: Verwenden Sie für Ihre Jobs, Stages und Links immer aussagekräftige Namen. Die Verwendung einer konsistenten Namenskonvention mit aussagekräftigen Namen stellt sicher, dass Ihre Jobs leicht nachvollziehbar und in einem Verzeichnis einfach aufzufinden sind.

- f. Stellen Sie sicher, dass das Feld **Ordnerpfad** den Pfad **\Tutorial\Jobs** enthält, und klicken Sie dann auf **Speichern**.

Stages hinzufügen und verlinken

Fügen Sie Ihrem Job Stages hinzu und verlinken Sie diese. Die Stage **Lookup File Set** (Lookup-Dateigruppe), die Sie erstellen, wird im weiteren Verlauf dieses Lernprogramms als Eingabe verwendet.

Stellen Sie sicher, dass der von Ihnen erstellte Job **Dev_CustData_Codes_Lookup** im Designer-Client geöffnet ist.

1. Fügen Sie Ihrem Job eine Stage **Lookup File Set** hinzu.
 - a. Erweitern Sie in der Palette die Kategorie **Datei**.

- b. Klicken Sie auf **Lookup-Dateigruppe** und klicken Sie dann auf einen leeren Bereich rechts in Ihrem Jobentwicklungsbereich.
 - c. Klicken Sie mit der rechten Maustaste auf die Stage und wählen Sie **Umbenennen** aus.
 - d. Benennen Sie die Stage um in `CountryCode_LKP_FS` und drücken Sie die Eingabetaste. Wie bereits dargelegt, ist es wichtig, aussagekräftige Namen zu verwenden. Dieser Name gibt an, dass es sich bei dieser Stage um eine Lookup-Dateigruppe (LKP_FS - Lookup File Set) handelt, die Landescodedaten enthält.
2. Fügen Sie Ihrem Job eine Stage **Sequential File** hinzu.
 - a. Erweitern Sie in der Palette die Kategorie **Datei**.
 - b. Klicken Sie auf **Sequenzielle Datei** und klicken Sie dann auf einen leeren Bereich links neben der Stage `CountryCode_LKP_FS`.
 - c. Benennen Sie die Stage **Sequential File** in `CountryCode_In_SF` um. Dieser Name gibt an, dass es sich bei dieser Stage um eine sequenzielle Datei handelt, die Landescodes als Eingabedaten enthält.
 3. Verlinken Sie die Stages, die Sie Ihrem Job hinzugefügt haben.
 - a. Erweitern Sie in der Palette die Kategorie **Allgemein**.
 - b. Klicken Sie auf **Link**.
 - c. Positionieren Sie den Cursor auf der Stage `CountryCode_LKP_FS`, klicken Sie mit der linken Maustaste und ziehen Sie Ihren Cursor zur Stage `CountryCode_In_SF`.
 - d. Benennen Sie den Link in `lnk_CountryCode_In` um.
 4. Klicken Sie in der Symbolleiste auf **Datei > Speichern**, um Ihren Job zu speichern.

Ihr Job sollte so ähnlich wie der in der folgenden Abbildung aussehen:



Abbildung 2. Ihr erster Job umfasst eine Stage **Sequential File**, die in eine Stage **Lookup File Set** schreibt.

In dieser Lerneinheit haben Sie Ihren ersten Job erstellt, Ihrem Job Stages und Links hinzugefügt und Ihren Job im Repository gespeichert.

Lerneinheit 2.2: Metadaten angeben

In dieser Lerneinheit geben Sie die Spaltenmetadaten für den Job an, den Sie in Lerneinheit 2.1 erstellt haben. Wenn Sie nach und nach weitere Bestandteile für Ihren Job entwickeln, kompilieren Sie Ihren Job erneut und führen Sie ihn aus. Durch diese Vorgehensweise stellen Sie sicher, dass jeder einzelne Teil Ihres Jobs ordnungsgemäß funktioniert, bevor Sie den nächsten Teil Ihres Jobs entwerfen.

Übersicht

Im ersten Teil Ihres Jobs wird eine Datei mit durch Kommas getrennten Werten (CSV-Datei) gelesen, die eine Reihe von Kundennummern, einen entsprechenden Code für das Land, in dem die Kunden ansässig sind, und einen weiteren Code


für die Sprache der Kunden enthält. Ihr Job liest die CSV-Datei und schreibt den Inhalt in eine Lookup-Tabelle in einer Stage **Lookup File Set** (Lookup-Dateigruppe). Diese Lookup-Tabelle wird von einem nachfolgenden Job zum Füllen einer Dimensionstabelle verwendet.

Tasks in dieser Lerneinheit

1. „Spaltenmetadaten angeben“
2. „Lookup-Eigenschaften angeben und Job ausführen“ auf Seite 25

Spaltenmetadaten angeben

Zum Angeben der Aktionen, die während der Ausführung des Jobs ausgeführt werden sollen, bearbeiten Sie die Eigenschaften Ihrer Stages. Sie beginnen mit einer Beschreibung der Spaltenmetadaten, die von der Stage `CountryCode_In_SF` abgehen.

1. Doppelklicken Sie auf die Stage `CountryCode_In_SF`. Der Stage-Editor wird auf der Registerkarte **Eigenschaften** der Seite **Ausgabe** geöffnet.
 - a. Klicken Sie unter der Kategorie **Quelle** auf die Eigenschaft **Datei**.
 - b. Klicken Sie im Feld **Datei** auf das Symbol **Durchsuchen** () und klicken Sie dann auf **Nach Datei suchen**.
 - c. Navigieren Sie im Dateiauswahlfenster zu dem Verzeichnis, in dem sich die Dateien für die Lernprogramme befinden.
 - d. Wählen Sie die Datei `CustomerCountry.csv` aus und drücken Sie dann die Eingabetaste, um den Wert für die Eigenschaft **Datei** zu übergeben.
 - e. Klicken Sie unter der Kategorie **Optionen** auf die Eigenschaft **Spaltennamen in erster Zeile**.
 - f. Ändern Sie den Wert für diese Eigenschaft in **Wahr**. Dieser Wert gibt an, dass die Zeile, die die Spaltennamen enthält, gelöscht wird, wenn der Job die Quelldatei liest.
2. Klicken Sie auf die Registerkarte **Format**.
 - a. Klicken Sie auf **Datensatzebene**.
 - b. Klicken Sie im Abschnitt **Zum Hinzufügen verfügbare Eigenschaften** auf **Datensatzbegrenzerzeichenfolge**.
 - c. Klicken Sie im Dropdown-Menü **Datensatzbegrenzerzeichenfolge** auf **DOS-Format**. Diese Einstellung stellt sicher, dass die Quelldatei auf einem AIX-, HP-UX-, Solaris oder Linux-Computer gelesen werden kann.
3. Klicken Sie auf die Registerkarte **Spalten**.
 - a. Geben Sie die folgenden Spaltenmetadaten ein. Lassen Sie alle Spalten leer, deren Werte in der folgenden Tabelle nicht angegeben sind.

Wichtig: Bei Spaltennamen muss die Groß-/Kleinschreibung beachtet werden.

Spaltenname	Schlüssel	SQL-Typ	Länge	Kann Nullwerte enthalten	Beschreibung
CUSTOMER_NUMBER	✓	Char	7	Nein	Kundenidentifikationsnummer
COUNTRY		Char	2	Nein	Code, der das Land angibt, in dem der Kunde ansässig ist

Spaltenname	Schlüssel	SQL-Typ	Länge	Kann Nullwerte enthalten	Beschreibung
LANGUAGE		Char	2	Nein	Code, der die Sprache des Kunden angibt

Da die Datei CustomerCountry.csv nur drei Spalten enthält, geben Sie die Spaltendefinitionen manuell ein. Wenn eine Datei viele Spalten enthält, ist es zeitsparender und präziser, die Spaltendefinitionen direkt aus der Datenquelle zu importieren. Die Vorgehensweise zum Importieren von Tabellendefinitionen lernen Sie im weiteren Verlauf dieses Lernprogramms kennen.

- b. Klicken Sie auf **Speichern**, um die Spaltendefinitionen, die Sie als Tabellendefinitionsobjekt angegeben haben, im Repository zu speichern. Die Definitionen können dann in anderen Jobs wiederverwendet werden.
- c. Geben Sie im Fenster **Tabellendefinition speichern** die folgenden Informationen ein:

Tabelle 2. Details für die Tabellendefinition

Parameter	Wert
Datenquellentyp	CSV
Datenquellenname	CustomerCountry.csv
Tabellen-/Dateiname	lnk_CountryCode_In
Ausführliche Beschreibung	Tabellendefinition für die Landescode-Quellendatei

- d. Klicken Sie auf **OK**, um die Position anzugeben, an der die Tabellendefinition gespeichert werden soll.
 - e. Klicken Sie im Fenster **Tabellendefinition speichern unter** mit der rechten Maustaste auf die Datei Tutorial und wählen Sie **Neu > Ordner** aus.
 - f. Benennen Sie den Ordner in Table definitions um. In diesem Ordner speichern Sie alle Ihre Tabellendefinitionen.
 - g. Geben Sie Ihrer Tabellendefinition den Namen lnk_CountryCode_In und klicken Sie dann auf **Speichern**.
4. Klicken Sie im Fenster des Stage-Editors auf **Daten anzeigen** und anschließend im Fenster **Datenbrowser** auf **OK**, um die Standardeinstellungen zu verwenden.
- Die Daten aus der Datei CustomerCountry.csv werden im Fenster **Datenbrowser** angezeigt. Die Daten werden den von Ihnen angegebenen Spaltenmetadaten zugeordnet.
5. Klicken Sie auf **Schließen**, um das Fenster **Datenbrowser** zu schließen.
 6. Klicken Sie auf **OK**, um das Stage-Editor-Fenster zu schließen.
 7. Speichern Sie den Job.

Dem Link lnk_CountryCode_In wird ein Symbol **Tabellendefinition** () hinzugefügt. Dieses Symbol gibt an, dass dem Link Spaltenmetadaten zugeordnet sind.


Lookup-Eigenschaften angeben und Job ausführen

Sie haben die Metadaten definiert, die von der Stage CountryCode_In_SF an die Stage CountryCode_LKP_FS gesendet werden. Nun konfigurieren Sie die Lookup-Eigenschaften dieser Metadaten.



1. Öffnen Sie die Stage CountryCode_LKP_FS.

2. Klicken Sie auf **Lookup-Schlüssel** und wählen Sie dann die Eigenschaft **Schlüssel** aus der Liste **Zum Hinzufügen verfügbare Eigenschaften** aus.
3. Wählen Sie **CUSTOMER_NUMBER** im Dropdown-Menü **Schlüssel** aus und drücken Sie dann die Eingabetaste.

Sie haben die Spalte **CUSTOMER_NUMBER** als Lookup-Schlüssel für die Lookup-Tabelle angegeben, die Sie erstellen.

4. Klicken Sie auf die Eigenschaft **Lookup-Dateigruppe** unter der Kategorie **Ziel**.
 - a. Klicken Sie im Feld **Lookup-Dateigruppe** auf das Symbol **Durchsuchen** () und klicken Sie dann auf **Nach Datei suchen**. Das Dateibrowserfenster wird geöffnet.
 - b. Navigieren Sie im Dateibrowserfenster zu dem Verzeichnis, in dem Sie die Dateigruppe speichern möchten. Beispiel: `C:\IBM\InformationServer\Server\Projects\Tutorial\Tutorial_files`.
 - c. Geben Sie im Feld **Dateiname** einen Namen für die Dateigruppe ein, die Sie gerade erstellen. Beispiel: `CountryCode_lookup.fs`.

Wichtig: Achten Sie darauf, die Erweiterung `.fs` anzugeben, die angibt, dass es sich bei dieser Datei um eine Dateigruppe handelt.

- d. Drücken Sie die Eingabetaste, um den Pfadnamen für die Lookup-Dateigruppe zu übergeben.
5. Klicken Sie im Stage-Editor `CountryCode_LKP_FS` auf **OK**.
6. Speichern Sie Ihren Job und klicken Sie dann auf die Schaltfläche **Kompilieren** () , um Ihren Job zu kompilieren.
7. Nachdem der Job erfolgreich kompiliert wurde, öffnen Sie den Director-Client und führen Sie Ihren Job aus.
8. Wenn der Job erfolgreich ausgeführt wird, klicken Sie auf das Symbol **Jobprotokoll** () .

Das Jobprotokoll enthält eine Informationsnachricht, die in etwa wie die folgende Nachricht aussieht. Diese Nachricht gibt die Anzahl der Zeilen, die erfolgreich verarbeitet wurden, sowie die Anzahl der fehlgeschlagenen Datensätze an.

```
CountryCode_In_SF,0: Import abgeschlossen;
1998 Datensätze wurden erfolgreich importiert, 0 wurden zurückgewiesen.
```

Sie haben Ihren ersten Job entworfen und ausgeführt. Jetzt importieren Sie Metadaten und ordnen sie den Links zu, die die Stages in Ihrem Job miteinander verbinden.

Lerneinheit 2.3: Metadaten importieren

In Lerneinheit 2.2 haben Sie die Spaltenmetadaten manuell eingegeben, die Sie importieren möchten. In dieser Lerneinheit importieren Sie Spaltenmetadaten direkt aus der Datenquelle, die von Ihrem Job gelesen wird. Sie speichern die Metadaten im InfoSphere DataStage-Repository, wo sie in anderen Jobs verwendet werden können.

Übersicht

Das Importieren von Metadaten aus Quellendateien verringert die Fehlerrate und vereinfacht das Mapping von Metadaten. In dieser Lerneinheit fügen Sie dem Job, den Sie in Lerneinheit 2.2 entworfen haben, Stages hinzu. Die Stages, die Sie hinzu-



fügen, ähneln denen, die Sie zuvor hinzugefügt haben. Die Stages lesen eine Datei mit durch Kommas getrennten Werten (CSV-Datei), die Codenummern und entsprechende spezielle Lieferanweisungen enthält. Der Inhalt wird in eine Lookup-Tabelle geschrieben, die im weiteren Verlauf dieses Lernprogramms von einem Job verwendet wird. Der fertig gestellte Job enthält zwei separate Datenflüsse, die Daten jeweils in eine separate Stage **Lookup File Set** (Lookup-Dateigruppe) schreiben.

Tasks in dieser Lerneinheit

1. „Metadaten in Ihr Repository importieren“
2. „Spaltenmetadaten aus dem Repository laden“

Metadaten in Ihr Repository importieren

Sie importieren Spaltendefinitionen aus der CSV-Datei, die spezielle Lieferanweisungen enthält. Anschließend speichern Sie die Spaltendefinitionen als Tabellendefinition im Repository. Sie können Metadaten importieren, ohne einen Job im Designer-Client zu öffnen; es wird jedoch empfohlen, den Metadatenimport bei bereits geöffnetem Job durchzuführen, um sicherzustellen, dass Sie das Mapping der Metadaten ordnungsgemäß durchführen.

1. Klicken Sie im Designer-Client auf **Importieren > Tabellendefinitionen > Definitionen sequenzieller Dateien**.
2. Importieren Sie im Fenster **Metadaten importieren** die Datei `SpecialHandling.csv`.
 - a. Klicken Sie im Feld **Verzeichnis** auf das Symbol **Durchsuchen** () , um zu dem Verzeichnis zu navigieren, in dem sich die Quelldateien für das Lernprogramm befinden. Beispiel: `C:\IBM\InformationServer\Server\Projects\Tutorial\Tutorial_files`.
 - b. Wählen Sie **Durch Kommas getrennt (*.csv)** im Dropdown-Menü **Dateityp** aus.
 - c. Wählen Sie in der Liste **Dateien** die Datei `SpecialHandling.csv` aus.
 - d. Klicken Sie im Feld **Zielordner** auf das Symbol **Durchsuchen** () . Wählen Sie im Fenster **Ordner auswählen** den zuvor erstellten Ordner `Tutorial\Table definitions` aus und klicken Sie dann auf **OK**.
 - e. Klicken Sie im Fenster **Metadaten importieren** auf **Importieren**, um Metadaten aus der Datei `SpecialHandling.csv` zu importieren.
3. Wählen Sie im Fenster **Sequenzielle Metadaten definieren** das Kontrollkästchen **Spaltennamen in erster Zeile** aus.

Diese Auswahl gibt an, dass die Zeile, die die Spaltennamen enthält, gelöscht wird, wenn der Job die Quelldatei liest.
4. Klicken Sie auf die Registerkarte **Definieren**. Wählen Sie den Code `SPECIAL_HANDLING` als Lookup-Schlüssel aus.
5. Klicken Sie auf **OK**, um das Fenster **Sequenzielle Metadaten definieren** zu schließen.
6. Klicken Sie auf **Schließen**, um das Fenster **Metadaten importieren** zu schließen.

Spaltenmetadaten aus dem Repository laden

Sie geben die von einer Stage verwendeten Spaltenmetadaten an, indem Sie die Metadaten aus einer Tabellendefinition im Repository laden.

Öffnen Sie Ihren Job `Dev_CustData_Codes_Lookup`.

In diesem Teil der Lerneinheit vertiefen Sie Ihre Kenntnisse im Entwerfen von Jobs, die Sie bereits erworben haben, und laden die Spaltenmetadaten aus der importierten Tabellendefinition.


1. Fügen Sie Ihrem Job eine Stage **Sequential File** (Sequenzielle Datei) hinzu und benennen Sie die Stage in `SpecHandlingCode_In_SF` um. Platzieren Sie diese Stage direkt unter die Stage `CountryCode_In_SF`.
2. Fügen Sie rechts neben der Stage `SpecHandlingCode_In_SF` eine Stage **Lookup File Set** (Lookup-Dateigruppe) hinzu und benennen Sie die Stage in `SpecHandlingCode_LKP_FS` um.
3. Verlinken Sie die beiden Stages, die Sie Ihrem Job hinzugefügt haben. Benennen Sie den Link in `lnk_SpecHandling_In` um. Ihr Jobentwurf sollte so ähnlich wie der in der folgenden Abbildung aussehen:



Abbildung 3. Ihr aktualisierter Job enthält zwei Stages **Sequential File**, die in zwei Stages **Lookup File Set** schreiben

4. Konfigurieren Sie die Stage `SpecHandling_In_SF`.
 - a. Geben Sie die Datei `SpecialHandling.csv` als Quellendatei für die Stage an.
 - b. Klicken Sie auf die Registerkarte **Format**.
 - c. Wählen Sie die Eigenschaft **Datensatzebene** aus. Wählen Sie dann die Eigenschaft **Datensatzbegrenzerzeichenfolge** in der Liste **Zum Hinzufügen verfügbare Eigenschaften** aus.
 - d. Geben Sie **DOS-Format** aus dem Dropdown-Menü **Datensatzbegrenzerzeichenfolge** an. Diese Einstellung stellt sicher, dass die Datei gelesen werden kann, wenn die Engineschicht auf einem AIX-, HP-UX-, Linux- oder Solaris-Computer installiert wird.
 - e. Klicken Sie auf die Registerkarte **Spalten**.
 - f. Klicken Sie auf der Registerkarte **Spalten** auf **Laden**.
 - g. Wählen Sie im Fenster **Tabellendefinitionen** die zuvor erstellte Tabellendefinition `SpecialHandling.csv` aus und klicken Sie dann auf **OK**.
 - h. Stellen Sie im Fenster **Spalten auswählen** sicher, dass alle Spalten im Teilfenster „Ausgewählte Spalten“ aufgelistet werden und klicken Sie dann auf **OK**.

- i. Klicken Sie auf **OK**, um den Editor der Stage SpecHandling_In_SF zu schließen.

Ein Symbol **Tabellendefinition** () wird dem Link lnk_SpecHandling_In hinzugefügt, um anzugeben, dass Sie diesem Link Tabellendefinitionen hinzugefügt haben.

5. Öffnen Sie die Stage SpecHandlingCode_LKP_FS.
 - a. Geben Sie eine Lookup-Dateigruppe für die Zieldatei an. Geben Sie der Datei den Namen SpecHandlingCode.fs.
 - b. Wählen Sie die Eigenschaft **Lookup-Schlüssel** aus. Wählen Sie dann die Eigenschaft **Schlüssel** aus der Liste „Zum Hinzufügen verfügbare Eigenschaften“ aus.
 - c. Geben Sie den Code für Sonderbehandlung (SPECIAL_HANDLING_CODE) an.
 - d. Klicken Sie auf **OK**, um den Editor der Stage SpecHandlingCode_LKP_FS zu schließen.
6. Speichern und kompilieren Sie Ihren Job Dev_CustData_Codes_Lookup und führen Sie ihn aus.

Sie haben Ihren Job erweitert, eine weitere Lookup-Dateigruppe erstellt und Tabellendefinitionen für die Verwendung in Ihrem Job importiert. Jetzt erstellen Sie Parameter, um die Position von Quelldateien für jede Stage in Ihrem Job anzugeben. Diese Parameter werden Sie im Verlauf des gesamten restlichen Lernprogramm wiederverwenden.

Lerneinheit 2.4: Parameter hinzufügen

In der vorherigen Lerneinheit haben Sie die Position von Quelldateien für jede Stage in Ihrem Job angegeben. Sie haben angegeben, dass zwei Dateien mit durch Kommas getrennten Werten als Eingabe und zwei Lookup-Dateien als Zieldateien verwendet werden sollen. In dieser Lerneinheit ersetzen Sie die Pfadnamen für diese Dateien durch Jobparameter. Sie verwenden diese Parameter, um Ihrem Job während der Ausführung Werte für Variablen zu übergeben.

Übersicht

Indem der Wert des Parameters bei jeder Ausführung des Jobs angegeben wird, wird sichergestellt, dass Sie die richtigen Ressourcen (z. B. Quelldateien und Zieldateien) verwenden. Wenn Sie in Ihren Jobs Parameter verwenden, fördern Sie die Wiederverwendung und verbessern Ihren Jobentwurf.

Ähnlich wie bei anderen Objekten in Ihrem Job, speichern Sie jeden Parameter im Projektrepository.

Tasks in dieser Lerneinheit

1. „Jobparameter definieren“
2. „Ihrem Job Jobparameter hinzufügen“ auf Seite 30
3. „Job mit Parametern ausführen“ auf Seite 31

Jobparameter definieren

Definieren Sie die Parameter, die der Job verwendet, wenn er ausgeführt wird. Bei der Ausführung des Jobs geben Sie dann Werte für jeden dieser Parameter an.

Stellen Sie sicher, dass der von Ihnen erstellte Job Dev_CustData_Codes_Lookup im Designer-Client geöffnet ist.


1. Klicken Sie in der Menüleiste im Designer-Client auf **Bearbeiten > Jobeigenschaften**.
2. Klicken Sie im Fenster **Jobeigenschaften** auf die Registerkarte **Parameter**.
 - a. Doppelklicken Sie auf die erste Zelle des Rasters, um eine neue Zeile hinzuzufügen.
 - b. Fügen Sie die folgenden Informationen hinzu, um vier Jobparameter zu erstellen:

Parametername	Eingabe	Typ	Hilfetext
CountryCode_In_Source	Pfadname der Quellendatei für die Landescodes	Pfadname	Geben Sie den Pfadnamen der CSV-Datei an, die die Landescode-Definitionen enthält.
CountryCode_LKP_Target	Pfadname der Lookup-Dateigruppe für Landescodes	Pfadname	Geben Sie den Pfadnamen der Dateigruppe für die Lookup-Tabelle für Landescodes an.
SpecHandlingCode_In_Source	Pfadname der Quellendatei für die Codes für Sonderbehandlung	Pfadname	Geben Sie den Pfadnamen der CSV-Datei an, die die Definitionen für die Codes für Sonderbehandlung enthält.
SpecHandlingCode_LKP_Target	Pfadname der Lookup-Dateigruppe für die Codes für Sonderbehandlung	Pfadname	Geben Sie den Pfadnamen der Dateigruppe für die Lookup-Tabelle für die Sonderbehandlung an.

- c. Klicken Sie auf **OK**, um das Fenster **Jobeigenschaften** zu schließen.
3. Speichern Sie Ihren Job, um die erstellten Parameter zu speichern.

Ihrem Job Jobparameter hinzufügen

Nun, da Sie Parameter für Ihren Job definiert haben, binden Sie diese in Ihren Jobentwurf ein.

1. Öffnen Sie die Stage CountryCode_In_SF.
 - a. Klicken Sie auf die Eigenschaft **Datei** unter der Kategorie **Quelle**.
 - b. Löschen Sie den zuvor eingegebenen Pfadnamen.
 - c. Klicken Sie im Feld **Datei** auf das Symbol **Durchsuchen**  und klicken Sie dann auf **Jobparameter einfügen**.
 - d. Wählen Sie CountryCode_In_Source aus der Liste der verfügbaren Parameter aus und drücken Sie dann die Eingabetaste.

Der Name des Parameters wird zwischen Nummernzeichen (#) gesetzt. Diese Änderung gibt an, dass der Job den Namen der Datei bei der Ausführung des Jobs anfordert.


- e. Klicken Sie auf **OK**, um den Editor der Stage `CountryCode_In_SF` zu schließen.
2. Öffnen Sie die Stage `CountryCode_LKP_FS`.
 - a. Löschen Sie den Pfadnamen, den Sie in das Feld für die Lookup-Dateigruppe eingegeben haben.
 - b. Geben Sie im Feld für die Lookup-Dateigruppe den Parameter `CountryCode_LKP_Target` an.
 - c. Klicken Sie auf **OK**, um den Editor der Stage `CountryCode_LKP_Target` zu schließen.
3. Öffnen Sie die Stage `SpecHandlingCode_In_SF`.
 - a. Löschen Sie den Pfadnamen, den Sie in das Feld **Datei** eingegeben haben.
 - b. Geben Sie im Feld **Datei** den Parameter `SpecHandlingCode_In_Source` an.
 - c. Klicken Sie auf **OK**, um den Editor der Stage `CountryCode_In_SF` zu schließen.
4. Öffnen Sie die Stage `SpecHandlingCode_LKP_FS`.
 - a. Löschen Sie den Pfadnamen, den Sie in das Feld für die Lookup-Dateigruppe eingegeben haben.
 - b. Geben Sie im Feld für die Lookup-Dateigruppe den Parameter `SpecHandlingCode_LKP_Target` an.
 - c. Klicken Sie auf **OK**, um den Editor der Stage `SpecHandlingCode_LKP_FS` zu schließen.
5. Speichern Sie den Job und kompilieren Sie ihn.

Job mit Parametern ausführen

Sie führen Ihren Job mit den Parametern aus, die Sie erstellt haben. Bei der Ausführung des Jobs werden Sie dann aufgefordert, Werte für die Parameter anzugeben.

1. Öffnen Sie den Director-Client über den Designer-Client.
2. Wählen Sie im Director-Client Ihren Job aus und klicken Sie auf das Symbol

Ausführen ()

- a. Klicken Sie im Fenster **Jobausführungsoptionen** auf das Symbol **Durchsuchen** () , um nach den Quelldateien und Zieldateien zu suchen.
- b. Klicken Sie auf **Ausführen**, um Ihren Job auszuführen.

Sie haben Jobparameter definiert, die die Dateinamen in Ihrem Job darstellen, und die Werte für diese Parameter bei der Ausführung des Jobs angegeben. In der nächsten Lerneinheit kombinieren Sie diese Parameter in einem Parametersatz, um Ihren Jobentwurf weiter zu vereinfachen.

Lerneinheit 2.5: Parametersatz erstellen

In der vorherigen Lerneinheit haben Sie Parameter für Ihren Job erstellt. Mit dem Ziel, Ihren Jobentwurf und die Laufzeitvorgänge weiter zu vereinfachen, erstellen Sie einen Parametersatz und speichern ihn im Projektrepository. Der Parametersatz kann dann in anderen Jobentwürfen wiederverwendet werden.

Übersicht

Bei der Erstellung des Parametersatzes geben Sie eine Gruppe von Standardwerten für die Parameter ein, die während der Ausführung als Eingabe übergeben werden.

Mithilfe von Parametersätzen definieren Sie Jobparameter, die Sie möglicherweise in anderen Jobs wiederverwenden. Wenn Sie diese Parameter in einem Job verwenden möchten, fügen Sie die Parameter aus dem Parametersatz in die Jobeigenschaften ein. Sie können Parametersätze aus vorhandenen Jobparametern erstellen oder können die Jobparameter angeben, wenn Sie einen neuen Parametersatz erstellen.

Sie können auch verschiedene Gruppen von Werten für die einzelnen Parametersätze definieren. Parametersätze werden im Installationsverzeichnis des InfoSphere DataStage-Servers in Form von Dateien gespeichert und können in Ihren Jobentwürfen oder bei der Ausführung von Jobs verwendet werden, die diese Parametersätze verwenden. Wenn Sie an einem Parametersatzobjekt Änderungen vornehmen, finden die Änderungen in den Jobentwürfen Anwendung, die dieses Objekt verwenden, wenn der Job kompiliert wird.

Wichtig: Die Parameter, mit denen ein Job kompiliert wird, sind bei der Ausführung des Jobs verfügbar. Wenn Sie den Entwurf jedoch ändern, nachdem der Job kompiliert wurde, stellt der Job einen Link zur aktualisierten Version des Parametersatzes her.

Vorgehensweise

Stellen Sie sicher, dass der von Ihnen erstellte Job `Dev_CustData_Codes_Lookup` im Designer-Client geöffnet ist.

1. Erstellen Sie einen Parametersatz.
 - a. Klicken Sie im Menü des Designer-Clients auf **Bearbeiten** > **Jobeigenschaften**.
 - b. Klicken Sie im Fenster **Jobeigenschaften** auf die Registerkarte **Parameter**.
 - c. Halten Sie die Steuertaste (Taste Strg) gedrückt und klicken Sie dann auf jeden der Parameter, die Sie in der vorherigen Lerneinheit erstellt haben.
 - d. Klicken Sie auf **Parametersatz erstellen**.
 - e. Geben Sie im Fenster **Parametersatz** in das Feld **Parametersatzname** einen Namen für Ihren Parametersatz ein, z. B. `CustData_Codes_Lookup_parmset`.
 - f. Geben Sie eine kurze Beschreibung für Ihren Parametersatz ein, beispielsweise Dieser Parametersatz enthält Parameter für Quelldateien und Zieldateien eines Jobs für ein Kundendaten-Lookup.

Ihren Parametersätzen beigefügte kurze Beschreibungen ermöglichen es anderen Entwicklern schnell zu erfassen, welchen Zweck Ihre Parametersätze haben und wann sie verwendet werden.
 - g. Klicken Sie auf die Registerkarte **Werte**.
 - h. Geben Sie in das Feld **Name der Wertedatei** einen Namen für Ihre Parametersatzwertedatei ein (z. B. `CustData_Codes_Lookup_1`).
 - i. Geben Sie für jeden Ihrer Jobparameter den Standardpfadnamen für die Datei an, die der Parameter darstellt. Beispiel: Der Parameter `CountryCode_In_Source` könnte den folgenden Wert haben: `C:\IBM\InformationServer\Server\Projects\Tutorial\Tutorial_files\CustomerCountry.csv`.
 - j. Klicken Sie auf **OK**.

- k. Klicken Sie im Fenster zum Speichern von Parametersätzen unter einem bestimmten Namen mit der rechten Maustaste auf den Ordner Tutorial und erstellen Sie einen neuen Ordner mit dem Namen Parameter sets.
- l. Klicken Sie auf **Speichern**, um Ihren Parametersatz in dem erstellten Ordner zu speichern.

Eine Nachricht wird angezeigt und fordert Sie auf anzugeben, ob Sie die ausgewählten Parameter durch den erstellten Parametersatz ersetzen möchten. Klicken Sie auf **Nein**.

Vorgehensweise beim Modifizieren des Parametersatzes: Wenn Sie Parameter in Ihrem Parametersatz modifizieren, erweitern Sie den erstellten Ordner Tutorial\Parameter sets und doppelklicken Sie auf Ihren Parametersatz. Sie können neue Parameter hinzufügen, vorhandene Parameter modifizieren und Werte für Parameter ändern, die in den Parametersatz aufgenommen wurden.

Wenn ein anderer Entwickler Ihren Job importiert und andere Quelldateien verwenden möchte, öffnet der Entwickler Ihren Parametersatz, modifiziert die Werte der einzelnen Parameter und führt den Job anschließend mit den aktualisierten Quelldaten aus.

2. Klicken Sie auf **OK**, um das Fenster **Jobparameter** zu schließen.
3. Speichern Sie den Job und kompilieren Sie ihn.

Weitere Schritte

Sie haben einen Parametersatz erstellt, der für einen anderen Job verfügbar ist, den Sie im weiteren Verlauf dieses Lernprogramms erstellen werden. Der aktuelle Job verwendet weiterhin die einzelnen Parameter anstelle des Parametersatzes.

Nun da Sie mit der Vorgehensweise beim Entwickeln eines einfachen Jobs vertraut sind, können Sie an die Entwicklung eines komplexeren Jobs herangehen. Im nächsten Modul entwickeln Sie einen Umsetzungsjob, der eine Datentypisierung auf die Rechnungsstellungsdaten von GlobalCo anwendet, um Inkonsistenzen zu beheben und die Formatierung der Ausgabedaten zu verbessern.

In dieser Lerneinheit haben Sie die folgenden Tasks ausgeführt:

- Erstellen des ersten Jobs und Hinzufügen von Stages und Links zum Job
- Festlegen von Stage-Eigenschaften, die bestimmen, welche Aktionen beim Ausführen des Jobs ausgeführt werden sollen
- Importieren von im gesamten Job zu verwendenden Metadaten
- Erstellen eines Parametersatzes aus einer Gruppe von vorhandenen Jobparametern
- Angeben einer Gruppe von Standardwerten für die Parameter im Parametersatz
- Erstellen eines in einem nachfolgenden Job zu verwendenden Parametersatzes

Kapitel 5. Modul 3: Daten umsetzen

Sie haben einen Job entwickelt, der Daten aus einer Quelldatei in eine Zieldatei schreibt. Nun, da Sie mit den Grundlagen des Jobentwurfs vertraut sind, erfahren Sie, wie Sie einen Job entwerfen, der Daten umsetzt.

Ihr Umsetzungsjob liest das Dataset `GlobalCoBillTo_ds`, das im Beispieljob in Modul 1 dieses Lernprogramms das Ziel bildete. Ihr Umsetzungsjob bereinigt diese Daten, setzt die Daten durch Löschen nicht benötigter Spalten um und schneidet Fremddaten ab.

Lernziele

Nach dem Durcharbeiten der Lerneinheiten in diesem Modul werden Sie in der Lage sein, die folgenden Tasks auszuführen:

- Verwenden der Stage **Transformer** (Umsetzung) zur Umsetzung von Daten
- Umgang mit zurückgewiesenen Daten
- Kombinieren von Daten mithilfe einer Stage **Lookup**
- Abschließen mehrerer Umsetzungen in einem einzigen Job

Erforderlicher Zeitaufwand

Für dieses Modul sind ungefähr 60 Minuten erforderlich.

Lerneinheit 3.1: Umsetzungsjob entwerfen

Sie entwerfen einen Job, der Teile der `GlobalCoBillTo_ds`-Daten umsetzt und die Ergebnisse dann in eine Stage **Data Set** schreibt, und führen diesen Job aus.

Übersicht

Die Quelldaten für den Beispieljob `GlobalCo_BillTo.csv` enthalten eine große Anzahl Spalten. Für die Dimensionstabelle, die Sie im weiteren Verlauf dieses Lernprogramms erstellen, ist nur eine Untergruppe dieser Spalten erforderlich, so dass Sie mithilfe der Stage **Transformer** (Umsetzung) einige der nicht benötigten Spalten entfernen. Darüber hinaus fügen Sie eine strengere Datentypisierung für die übrigen Spalten hinzu, um eine Qualitätskontrolle der Daten zu implementieren. Zum Abschneiden nicht benötigter Zeichen geben Sie eine Funktion in die Stage **Transformer** ein. Mithilfe dieser Funktionalität und der Funktionen zum Bereinigen von Daten wird sichergestellt, dass die Ausgabe in dem von Ihnen gewünschten Format vorliegt.

Weitere Informationen zu verarbeitenden Stages: Der Umsetzungsjob, den Sie entwerfen, verwendet eine Stage **Transformer**, der Designer-Client enthält jedoch andere verarbeitende Stages, mit denen Daten umgesetzt werden können. Sie können z. B. eine Stage **Modify** (Modifizieren) in Ihrem Job verwenden, wenn Sie nur die Datentypen in einem Dataset ändern möchten. Einige der verarbeitenden Stages können Datenspalten bei ihrer Verarbeitung löschen. Sie verwenden jedoch die Stage **Transformer**, da Sie eine Umsetzungsfunktion benötigen, die Sie anpassen können.

Tasks in dieser Lerneinheit

1. „Umsetzungsjob erstellen“
2. „'Data Set'-Stages konfigurieren“
3. „Stage 'Transformer' konfigurieren“ auf Seite 37
4. „Umsetzungsjob ausführen“ auf Seite 40

Umsetzungsjob erstellen

Zum Erstellen eines Umsetzungsjobs fügen Sie zwischen Ihrer Quelle und Ihrem Ziel eine Stage **Transformer** hinzu.

1. Erstellen Sie einen parallelen Job mit dem Namen Dev_CustData_TrimAndStrip.
2. Speichern Sie den Job im Ordner \Tutorial\Jobs, den Sie in einer vorherigen Lerneinheit erstellt haben.
3. Fügen Sie Ihrem Job eine Stage **Data Set** hinzu.
 - a. Klicken Sie in der Designer-Client-Palette auf **Datei**.
 - b. Klicken Sie auf **Dataset** und klicken Sie dann auf einen leeren Bereich in Ihrem Jobentwicklungsbereich.
 - c. Benennen Sie die Stage **Data Set** in GlobalCo_BillTo_In_DS um.
4. Fügen Sie Ihrem Job eine Stage **Transformer** hinzu.
 - a. Klicken Sie in der Designer-Client-Palette auf **Verarbeitung**.
 - b. Klicken Sie auf **Umsetzung** und klicken Sie dann rechts neben der Stage GlobalCo_BillTo_In_DS auf einen leeren Bereich in Ihrem Jobentwicklungsbereich.
 - c. Benennen Sie die Stage in GlobalCo_BillTo_TFM um.
5. Klicken Sie mit der rechten Maustaste auf die Stage GlobalCo_BillTo_In_DS, ziehen Sie den Cursor zur Stage GlobalCo_BillTo_TFM und lassen Sie die Maustaste dann los, um die Stages zu verlinken. Mit dieser Methode lassen sich Stages schneller und einfacher verlinken als bei der Auswahl eines Links in der Palette.
6. Benennen Sie den Link in lnk_Full_BillTo_In um.
7. Fügen Sie rechts neben der Stage GlobalCo_BillTo_TFM eine weitere Stage **Data Set** hinzu.
8. Benennen Sie die Stage **Data Set** in GlobalCo_BillTo_Out_DS um.
9. Verlinken Sie die Stage GlobalCo_BillTo_TFM mit der Stage GlobalCo_BillTo_Out_DS.
10. Benennen Sie den Link in lnk_Stripped_BillTo_Out um.
Ihr Job sollte ähnlich wie der in der folgenden Abbildung aussehen:

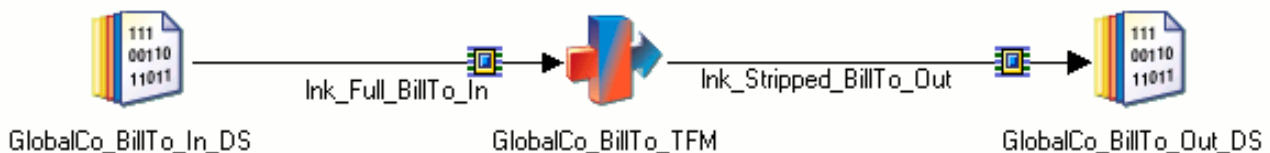



Abbildung 4. Ihr Job umfasst eine Stage **Transformer**, die eine Datentypisierung auf Ihre Daten anwendet, bevor diese in ein Dataset geschrieben werden.

'Data Set'-Stages konfigurieren

Konfigurieren Sie beide **Data Set**-Stages und laden Sie dann mithilfe einer Tabelendefinition die Spaltenmetadaten für jede der Stages.

1. Konfigurieren Sie die erste Stage **Data Set**.

- a. Öffnen Sie die Stage GlobalCo_BillTo_In_DS und legen Sie die Eigenschaft **Datei** so fest, dass sie auf das Dataset GlobalCoBillTo.ds verweist, auf das vom Beispieljob in Modul 1 geschrieben wurde.
 - b. Klicken Sie auf **OK**, um den Stage-Editor zu schließen.
2. Fügen Sie Ihrem Job eine Tabellendefinition hinzu.
 - a. Erweitern Sie im Fenster **Repository** den Ordner \Tutorial.
 - b. Wählen Sie die Tabellendefinition GlobalCoBillToSource aus und ziehen Sie sie zu dem Link lnk_Full_BillTo_In. Ein Symbol **Tabellendefinition** () wird dem Link hinzugefügt, um anzuzeigen, dass die Tabellendefinition nun dem Link zugeordnet ist.
Diese Tabellendefinition haben Sie beim Import des Jobs samplejob importiert. Diese Methode ermöglicht ein schnelleres und einfacheres Zuordnen von Tabellendefinitionen zu Jobs als das Laden von Tabellendefinitionen aus dem Stage-Editor.
3. Zeigen Sie die Daten in der Stage **Data Set** an.
 - a. Öffnen Sie die Stage GlobalCo_BillTo_In_DS.
 - b. Klicken Sie im Stage-Editor auf **Daten anzeigen** und klicken Sie anschließend auf **OK**, um das Fenster **Datenbrowser** zu öffnen.
Die Kundendatensätze im Dataset werden angezeigt. Nachdem Sie Ihren Jobs Tabellendefinitionen hinzugefügt haben, prüfen Sie die Daten, um zu verifizieren, dass die Spaltenmetadaten korrekt sind.
 - c. Klicken Sie auf **Schließen**, um das Fenster **Datenbrowser** zu schließen.
 - d. Klicken Sie auf **OK**, um den Stage-Editor zu schließen.
 4. Öffnen Sie die Stage GlobalCo_BillTo_Out_DS.
 - a. Legen Sie die Eigenschaft **Datei** so fest, dass sie auf das Dataset verweist, das Sie erstellen. Wenn Sie zum Beispiel Ihrem Dataset den Namen GlobalCo_BillTo_Out.ds zuweisen, geben Sie den vollständigen Pfadnamen zu dieser Datei an: C:\IBM\InformationServer\Server\Projects\Tutorial\Tutorial_files\GlobalCo_BillTo_Out.ds.
 - b. Klicken Sie auf **OK**, um den Stage-Editor zu schließen.
 5. Speichern Sie den Job.

Stage 'Transformer' konfigurieren

Definieren Sie die Umsetzungsoperationen, die vorgenommen werden sollen, wenn Sie Ihren Job ausführen.

1. Öffnen Sie die Stage GlobalCo_BillTo_TFM.
2. Geben Sie die Spalten an, die während der Ausführung durch die Stage **Transformer** fließen sollen.
 - a. Wählen Sie im linken Teilfenster die folgenden Spalten im Link lnk_Full_BillTo_In aus.

CUSTOMER_NUMBER
 CUST_NAME
 ADDR_1
 ADDR_2
 CITY
 REGION_CODE
 ZIP
 TEL_NUM

REVIEW_MONTH
 SETUP_DATE
 STATUS_CODE

- b. Ziehen Sie diese Spalten in den Link `Ink_Stripped_BillTo_Out` im rechten Teilfenster. Durch die Auswahl einer Untergruppe von Spalten geben Sie die Spalten an, die während der Ausführung durch die Stage **Transformer** fließen sollen. Die übrigen Spalten werden gelöscht.

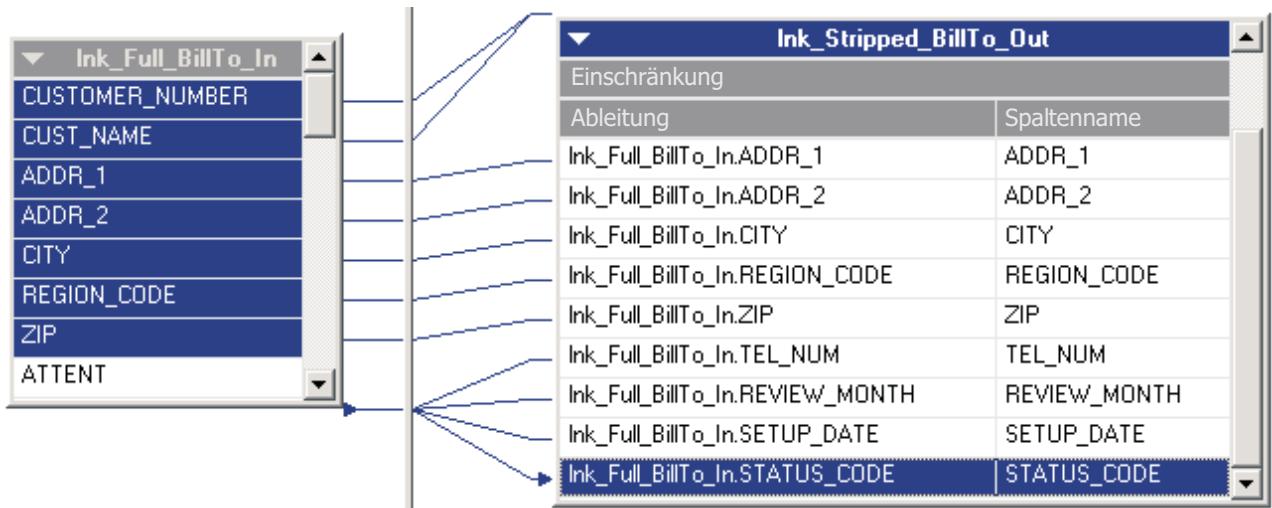




Abbildung 5. Sie ordnen die Spalten des Eingabelinks dem Ausgabelink zu. Durch dieses Mapping werden die Spaltenmetadaten definiert, die in die Stage **Transformer** fließen bzw. von dieser abgehen.

3. Geben Sie im Teilfenster unten rechts die folgenden Werte für die Spaltendefinitionen **SQL-Typ**, **Länge** und **Beschreibung** des Links `Ink_Stripped_BillTo_Out` ein.

Durch die Angabe einer strengeren Datentypisierung können Sie Inkonsistenzen in Ihren Quelldaten bei der Ausführung des Jobs besser diagnostizieren. Verwenden Sie die Standardwerte für alle Felder, die in der folgenden Tabelle nicht angegeben sind.

Spaltenname	SQL-Typ	Länge	Beschreibung
CUSTOMER_NUMBER	Char	7	Eindeutige Kundennummer
CUST_NAME	VarChar	30	Vollständiger Name des Kunden
ADDR_1	VarChar	30	Straßenname und Hausnummer der Kundenadresse
ADDR_2	VarChar	30	Zusätzliche Straßeninformationen wie die Nummer der Einheit
CITY	VarChar	30	Wohnort des Kunden
REGION_CODE	Char	2	Region in der der Kunde ansässig ist, sofern zutreffend
ZIP	VarChar	10	Postleitzahl des Wohnorts des Kunden

Spaltenname	SQL-Typ	Länge	Beschreibung
TEL_NUM	VarChar	10	Telefonnummer des Kunden
REVIEW_MONTH	VarChar	2	Monat, in dem der Kundendatensatz intern überprüft wird
SETUP_DATE	VarChar	12	Datum, an dem der Kundendatensatz im Dataset erstellt wurde
STATUS_CODE	Char	1	Code, der angibt, ob der Kundendatensatz aktiv, angehalten oder inaktiv ist

4. Fügen Sie der Spalte CUSTOMER_NUMBER eine Ableitung hinzu.
Der Ausdruckseditor wird geöffnet.
 - a. Doppelklicken Sie im Link lnk_Stripped_BillTo_Out auf das Feld **Ableitung** für die Spalte CUSTOMER_NUMBER und löschen Sie die Ableitung.
 - b. Klicken Sie auf das Symbol **Durchsuchen** () und klicken Sie dann auf **Funktion**.
 - c. Erweitern Sie die Kategorie **Zeichenfolge** ein und wählen Sie **Trim** aus. Das Ableitungsfeld wird mit einer einfachen Trimfunktion gefüllt.
 - d. Klicken Sie auf das Symbol **Durchsuchen** (), klicken Sie auf **Eingabespalte** und wählen Sie dann die Spalte CUSTOMER_NUMBER aus.
 - e. Löschen Sie die Parameter nach dem Spaltennamen und geben Sie anschließend die folgende Syntax ein.
' ', 'A'
 - f. Drücken Sie die Eingabetaste, um die Ableitung zu übergeben. Die Ableitung für die Spalte CUSTOMER_NUMBER sieht ähnlich wie die folgende Ableitung aus:
Trim(lnk_Full_BillTo_In.CUSTOMER_NUMBER, ' ', 'A')
Diese Ableitung gibt an, dass irrelevante Leerzeichen aus der Spalte CUSTOMER_NUMBER des Links lnk_Full_BillTo_In gelöscht werden. Die Leerzeichen werden entfernt, bevor die Daten in die Spalte CUSTOMER_NUMBER des Links lnk_Stripped_BillTo_Out geschrieben werden. Die Ableitung kann manuell eingegeben werden, die Verwendung der integrierten Funktionen hilft jedoch Fehler zu vermeiden und darüber hinaus stellen die Funktionen die Struktur der einzelnen Funktionen bereit.
5. Klicken Sie auf **OK**, um den Stage-Editor zu schließen.
6. Öffnen Sie die Stage GlobalCo_BillTo_Out_DS und klicken Sie dann auf die Registerkarte **Spalten**.
Das Dataset hat die Spalten übernommen, die Sie der Stage **Transformer** als Eingabe übergeben haben.
7. Speichern Sie den Job und kompilieren Sie ihn anschließend.

Umsetzungsjob ausführen

Bis jetzt haben Sie die Jobs über den Director-Client ausgeführt. In dieser Lerneinheit haben Sie Ihren Job über den Designer-Client ausgeführt. Diese Methode des Ausführens eines Jobs ist bei der Entwicklung von Jobs sinnvoll, da Sie den Designer-Client nicht verlassen müssen. Klicken Sie auf **Anzeigen > Jobprotokoll**, um die Jobprotokolldatei anzuzeigen.

1. Klicken Sie im Designer-Client-Menü auf **Diagramm > Leistungsstatistik anzeigen**. Während der Jobausführung stellt diese Ansicht zusätzliche Informationen bereit, z. B. die Anzahl der übertragenen Zeilen und die Anzahl der pro Sekunde verarbeiteten Zeilen.
2. Klicken Sie auf **Datei > Ausführen** und klicken Sie dann im Fenster **Jobausführungsoptionen** auf **Ausführen**.

Während der Ausführung des Jobs wird die Leistungsstatistik für den Link aktualisiert; die Links selbst ändern ihre Farbe je nach Status. Wenn der Job erfolgreich ausgeführt wird, ändert die Farbe der Links sich in Grün. Die Anzahl der verarbeiteten Zeilen wird unter den Linknamen angezeigt.

Anmerkung: Wenn Sie eine Fehlermeldung erhalten, die besagt, dass der Job aufgrund eines Fehlers in der Stage **Transformer** fehlgeschlagen ist, ist auf Ihrem Computer möglicherweise kein C++-Compiler installiert. Weitere Informationen zur Installation eines Compilers finden Sie im Abschnitt zum „Einrichten eines C++-Compilers“ in *IBM InfoSphere Information Server - Planung, Installation und Konfiguration*.

3. Nachdem der Job beendet ist, öffnen Sie den Director-Client. Sie können das Jobprotokoll im Director-Client anzeigen, auch wenn der Job über den Designer-Client ausgeführt wird.
 - a. Wählen Sie den Job `Dev_CustData_TrimAndStrip` aus.

- b. Klicken Sie auf das Symbol **Jobprotokoll** () , um das Jobprotokoll anzuzeigen.

- c. Doppelklicken Sie im Jobprotokoll auf das Ereignis mit dem folgenden Wortlaut: 'Hauptprogramm: Dieser Schritt umfasst 2 Datasets: (...)'

Das Feld **Nachricht** enthält die folgenden Informationen, die Aufschluss über die Komponenten geben, aus denen sich Ihr Job zusammensetzt.

Dieser Abschnitt befasst sich mit den Datasets, die in Ihrem Job verwendet werden.
Hauptprogramm: Dieser Schritt umfasst 2 Datasets:

Das erste Dataset (ds0) wird mit der Quelldatei GlobalCo_BillTo.ds verlinkt.
ds0: {C:/IBM/InformationServer/Server/Projects/Tutorial/Tutorial_files/GlobalCo_BillTo.ds
eAny=>eCollectAny
op0[2p] (parallel APT_CombinedOperatorController:GlobalCo_BillTo_In_DS)}

Das zweite Dataset (ds1) wird mit der Zieldatei GlobalCo_BillTo_Out.ds verlinkt.
ds1: {op0[2p] (parallel APT_CombinedOperatorController:GlobalCo_BillTo_Out_DS)
=>
C:/IBM/InformationServer/Server/Projects/Tutorial/Tutorial_files/GlobalCo_BillTo_Out.ds}

Ihr Job verwendet einen Operator, die Transformer-Stage GlobalCo_BillTo_TFM.
1 Operator wird verwendet:
op0[2p] {(parallel APT_CombinedOperatorController:
(GlobalCo_BillTo_In_DS)
(APT_TransformOperatorImplV0S1_Dev_CustData_TrimAndStrip_GlobalCo_BillTo_TFM
in GlobalCo_BillTo_TFM)
(GlobalCo_BillTo_Out_DS)
) on nodes (
node1[op0,p0]
node2[op0,p1]
))}

Ihr Job wird auf zwei Knoten ausgeführt: node1 und node2.
Er führt 2 Prozesse auf 2 Knoten aus.

- d. Klicken Sie auf **Schließen**, um das Fenster für die Ereignisdetails zu schließen.

Sie haben Ihren ersten Umsetzungsjob entworfen und konfiguriert. Sie haben die Spaltenmetadaten gefiltert, um die Daten für die Verarbeitung zu begrenzen und haben dann eine Funktion zur Bereinigung Ihrer Daten eingefügt.

In der nächsten Lerneinheit fügen Sie eine Stage **Lookup** hinzu, die ein zuvor erstelltes Dataset referenziert. Durch Hinzufügen dieses Lookups wird sichergestellt, dass nur übereinstimmende Datensätze in das Zieldataset geschrieben werden.

Lerneinheit 3.2: Stage 'Lookup' im Umsetzungsjob verwenden

In der vorherigen Lerneinheit haben Sie einen einfachen Umsetzungsjob entworfen. In dieser Lerneinheit erstellen Sie einen komplexeren Job, der Daten aus zwei verschiedenen Tabellen kombiniert.

Übersicht

Als Basis für den Job verwenden Sie den Umsetzungsjob, den Sie in der vorherigen Lerneinheit erstellt haben. Sie fügen eine Stage **Lookup** hinzu, die Daten einer Lookup-Tabelle referenziert, die Sie in Lerneinheit 2.2 erstellt haben. Sie haben die Schlüsselspalten für die Lookup-Tabellen bei der Erstellung der Lookup-Dateigruppen definiert.

Sie können **Lookup**-Stages so konfigurieren, dass sie in einer Lookup-Dateigruppe oder in einer relationalen Datenbank nach Daten suchen. Die Stage **Lookup** hat eine Datenstromeingabe und eine Referenzeingabe. Die Stage verwendet mindestens eine Schlüsselspalte in der Datenstromeingabe, um nach Daten in einer Referenztabelle zu suchen. Die Daten aus der Referenztabelle werden dann der Datenstromausgabe hinzugefügt. Wenn Sie Lookup-Dateigruppen verwenden, müssen Sie beim Definieren der Dateigruppe die Lookup-Schlüsselspalte angeben.

Eine mögliche Methode zum Kombinieren von Daten besteht darin, mithilfe einer Stage **Lookup** ein Lookup zu erstellen. Sie können Daten in einem parallelen Job auch über eine Stage **Join** kombinieren. Wenn Ihr Job eine große Referenztabelle verwendet, lassen sich die Daten über eine Stage **Join** schneller als über eine Stage **Lookup** kombinieren. Da Sie in Ihrem Job eine kleine Referenztabelle verwenden, ist eine Stage **Lookup** die bevorzugte Wahl. Die Stage **Lookup** ist am effizientesten, wenn die gesuchten Daten in den physischen Hauptspeicher passen.

Tasks in dieser Lerneinheit

1. „Lookup-Job erstellen“
2. „Stage 'Lookup File Set' konfigurieren“ auf Seite 43
3. „Stage 'Lookup' konfigurieren“ auf Seite 43

Lookup-Job erstellen

Erstellen Sie einen neuen Job, der die Stages und Links des Jobs `Dev_CustData_TrimAndStrip` enthält.

1. Erstellen Sie einen parallelen Job mit dem Namen `Dev_CustData_CleansePrepare`.
2. Speichern Sie den Job im Ordner `\Tutorial\Jobs`, den Sie in einer vorherigen Lerneinheit erstellt haben.
3. Öffnen Sie den Job `Dev_CustData_TrimAndStrip` und ziehen Sie anschließend den Cursor um die Stages und Links.
 - a. Klicken Sie im Menü auf **Bearbeiten** > **Kopieren**.

- b. Klicken Sie im Job Dev_CustData_CleansePrepare auf **Bearbeiten** > **Einfügen**, um die kopierten Stages und Links einzufügen.
Da die Arbeiten am Job Dev_CustData_TrimAndStrip abgeschlossen sind, können Sie ihn nun schließen.
4. Löschen Sie im Job Dev_CustData_CleansePrepare die Stage GlobalCo_BillTo_Out_DS. Sie ersetzen diese Stage durch eine andere.
5. Fügen Sie dem Job Dev_CustData_CleansePrepare eine Stage **Lookup** hinzu.
 - a. Klicken Sie in der Palette auf **Verarbeitung**.
 - b. Klicken Sie auf **Lookup** und legen Sie dann die Stage auf Ihrem Entwicklungsbereich rechts neben der Stage GlobalCo_BillTo_TFM ab.
 - c. Benennen Sie die Stage in GlobalCo_Country_LKP um.
 - d. Ziehen Sie den Link lnk_Stripped_BillTo_Out zur Stage GlobalCo_Country_LKP. Wenn Sie den Link verbinden, werden die zugehörigen Spaltenmetadaten mit dem Link versetzt. Diese Spaltenmetadaten können zwischen der Stage GlobalCo_BillTo_TFM und der Stage GlobalCo_Country_LKP fließen.
6. Fügen Sie Ihrem Job eine Stage **Lookup File Set** hinzu.
 - a. Klicken Sie in der Palette auf **Datei**.
 - b. Klicken Sie auf **Lookup-Dateigruppe** und legen Sie die Stage auf Ihrem Entwicklungsbereich über der Stage GlobalCo_Country_LKP ab.
 - c. Benennen Sie die Stage in GlobalCo_CountryCodes_FS um.
 - d. Verlinken Sie die Stage GlobalCo_CountryCodes_FS mit der Stage GlobalCo_Country_LKP.
 - e. Benennen Sie den Link in lnk_CountryCode_Ref um. Der Link wird als gepunktete Linie angezeigt. Dies weist daraufhin, dass es sich hierbei um einen Referenzlink handelt.
7. Fügen Sie rechts neben der Stage GlobalCo_Country_LKP eine Stage **Data Set** hinzu.
 - a. Benennen Sie die Stage in GlobalCo_CountryLookup_DS um.
 - b. Verlinken Sie die Stage GlobalCo_Country_LKP mit der Stage GlobalCo_CountryLookup_DS.
 - c. Benennen Sie den Link in lnk_CountryCode_Out um.

Ihr Job sieht wie der folgende Job aus:

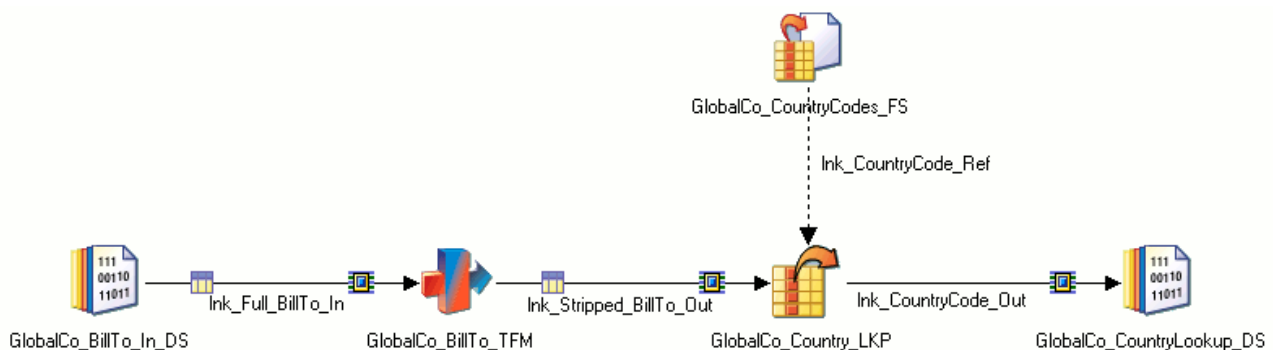



Abbildung 6. Ihr Job enthält eine Stage **Lookup**, die zwischen Landesdaten und von der Stage **Transformer** (Umsetzung) ausgehenden Kundendaten eine Referenz herstellt.

Stage 'Lookup File Set' konfigurieren

Im vorherigen Abschnitt dieser Lerneinheit haben Sie Stages des Jobs Dev_CustData_TrimAndStrip in den Job Dev_CustData_CleansePrepare kopiert. Diese Stages sind bereits konfiguriert; daher müssen Sie nur die Stages konfigurieren, die Sie hinzugefügt haben. Als erste Stage konfigurieren Sie die Stage GlobalCo_CountryCodes_FS. Diese Stage verwendet die Lookup-Dateigruppe, die Sie in Lerneinheit 2.2 erstellt haben. Darüber hinaus verwenden Sie den Parametersatz, den Sie in Lerneinheit 2.5 erstellt haben, um die Parameter zu füllen, die vom Job verwendet werden.

1. Fügen Sie Ihrem Job einen Parametersatz hinzu.
 - a. Öffnen Sie das Fenster **Jobeigenschaften** für den Job Dev_CustData_CleansePrepare und klicken Sie dann auf die Registerkarte **Parameter**.
 - b. Klicken Sie auf **Parametersatz hinzufügen**.
 - c. Erweitern Sie im Fenster **Öffnen** den Ordner Tutorial\Parameter sets.
 - d. Wählen Sie den Parametersatz CustData_Codes_Lookup_parmset aus, den Sie in Lerneinheit 2.5 erstellt haben, und klicken Sie dann auf **OK**.
 - e. Klicken Sie auf **OK**, um das Fenster **Jobeigenschaften** zu schließen.
2. Konfigurieren Sie die Stage GlobalCo_CountryCodes_FS.
 - a. Öffnen Sie die Stage GlobalCo_CountryCodes_FS. Klicken Sie unter der Kategorie **Quelle** auf die Eigenschaft **Lookup-Dateigruppe**.
 - b. Klicken Sie im Feld **Lookup-Dateigruppe** auf das Symbol **Durchsuchen** () und klicken Sie dann auf **Jobparameter einfügen**.
 - c. Wählen Sie den Parameter countryCode_LKP_Target aus und drücken Sie dann die Eingabetaste. Dieser Parameter ordnet der Stage die Dateigruppe SpecialHandlingCode.fs zu.
 - d. Klicken Sie auf die Registerkarte **Spalten**.
 - e. Laden Sie die Tabellendefinition lnk_CountryCode_In, die Sie in der vorherigen Lerneinheit erstellt haben.
 - f. Schließen Sie die Stage GlobalCo_CountryCodes_FS.

Stage 'Lookup' konfigurieren

In der Stage **Lookup** geben Sie die Daten an, die kombiniert werden sollen. Bei der Erstellung der Tabellendefinition lnk_CountryCode_In in Lerneinheit 2.2 haben Sie die Spalte CUSTOMER_NUMBER als Schlüssel für das Lookup angegeben. Mit diesem Schlüssel vergleichen Sie die Quelldaten mit den Daten in Ihrer Lookup-Dateigruppe.

1. Ordnen Sie die Lookup-Metadaten aus den Eingabespalten den Ausgabespalten zu.
 - a. Öffnen Sie die Stage GlobalCo_Country_LKP. Klicken Sie im linken Teilfenster auf die Titelleiste des Links lnk_Stripped_BillTo_Out und ziehen Sie dann den Cursor zum Link lnk_CountryCode_Out.
Die Spaltenmetadaten des Eingabelinks werden dem Ausgabelink zugeordnet.
 - b. Klicken Sie im linken Teilfenster auf die Spalte COUNTRY im Link lnk_CountryCode_Ref und ziehen Sie dann den Cursor zum Link lnk_CountryCode_Out.
Wenn der Wert der ankommenden Spalte CUSTOMER_NUMBER mit dem Wert der Spalte CUSTOMER_NUMBER der Lookup-Tabelle übereinstimmt, wird die

entsprechende Spalte COUNTRY dieser Datenzeile hinzugefügt. Im Stage-Editor werden die Spalten des Links `lnk_Stripped_BillTo_Out` dem Link `lnk_CountryCode_Out` zugeordnet. Die Spalte `CUSTOMER_NUMBER` wird in der Stage **Lookup** als Schlüssel zum Referenzieren von Kundendaten verwendet.

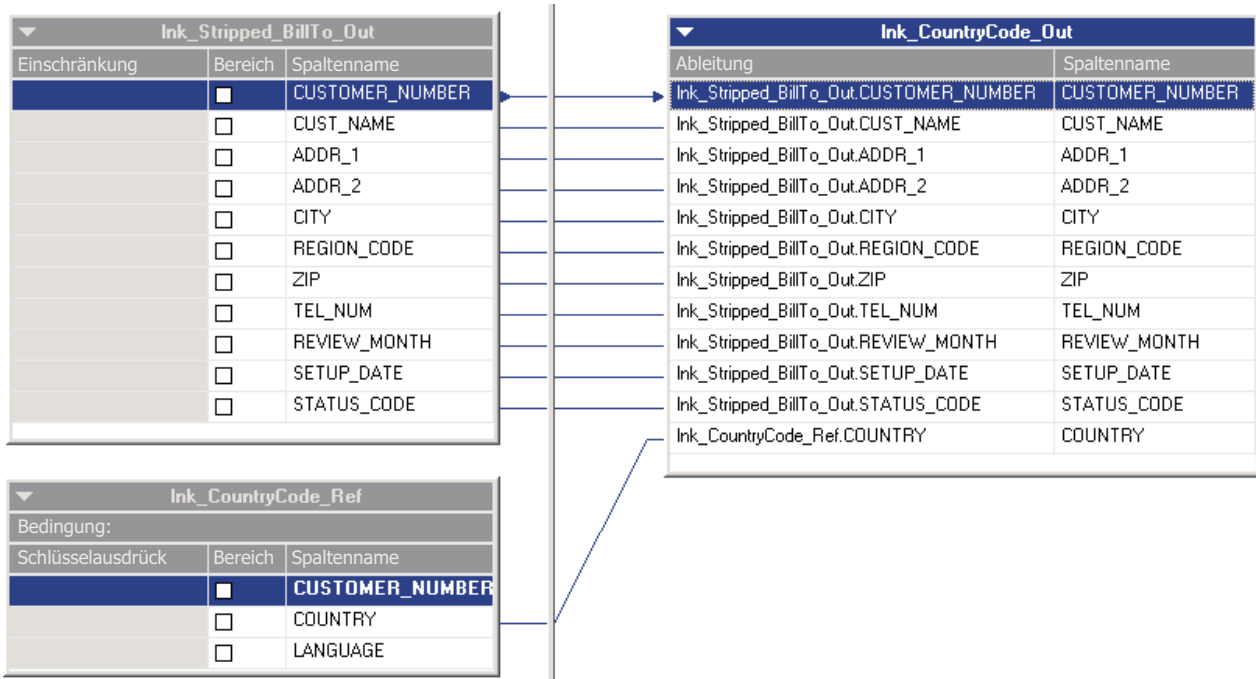


Abbildung 7. Die Spalten des Eingabelinks werden dem Ausgabelink zugeordnet.

- c. Doppelklicken Sie im Fenster **Stage 'Lookup' - Bedingungen** im Link `lnk_CountryCode_Ref` auf **Bedingung**.
- d. Wählen Sie **Fortfahren** in der Spalte **Lookup-Fehler** aus.
Diese Option gibt an, dass der Job, wenn ein Wert der Spalte `CUSTOMER_NUMBER` im Link `lnk_Stripped_BillTo_Out` mit keinem Wert in der Spalte `CUSTOMER_NUMBER` im Link `lnk_CountryCode_Ref` übereinstimmt, mit der Verarbeitung der Spalte `CUSTOMER_NUMBER` fortfährt.
- e. Klicken Sie auf **OK**, um das Fenster **Stage 'Lookup' - Bedingungen** zu schließen.
- f. Schließen Sie die Stage `GlobalCo_Country_LKP`.
2. Öffnen Sie die Stage `GlobalCo_CountryLookup_DS` und geben Sie einen Dateinamen für das Datasetziel an. Geben Sie der Datei beispielsweise den Namen `GlobalCo_CustData_Cleansed.ds`.
3. Speichern und kompilieren Sie den Job und führen Sie ihn aus. Wenn Sie den Job ausführen, werden im Fenster **Jobausführungsoptionen** alle Parameter im Parametersatz angezeigt.
 - a. Wählen Sie im Fenster **Jobausführungsoptionen** die Option `CustData_Codes_Lookup_1` aus der Liste neben dem Namen des Parametersatzes aus.
Die Parameterwerte werden mit den Pfadnamen gefüllt, die Sie bei der Erstellung des Parametersatzes angegeben haben.
 - b. Klicken Sie auf **Ausführen**, um den Job auszuführen.

4. Wenn der Job erfolgreich ausgeführt wird, öffnen Sie die Stage `GlobalCo_CountryLookup_DS` und klicken Sie dann auf **Daten anzeigen**. Die Kundendaten werden im Fenster **Datenbrowser** angezeigt.

In dieser Lerneinheit haben Sie irrelevante Zeichen aus den Kundendaten von GlobalCo entfernt und ein Lookup hinzugefügt, sodass nur übereinstimmende Datensätze in das Zieldataset geschrieben wurden. Durch die Bereinigung der Kundendaten haben Sie die Informationen für das Zusammenfassen mit den Kundendaten von WorldCo vorbereitet.

Lerneinheit 3.3: Zurückgewiesene Daten erfassen

In dieser Lerneinheit fügen Sie Ihrem Job einen Zurückweisungslink hinzu, um während der Ausführung des Jobs zurückgewiesene Daten zu erfassen.

Übersicht

In Lerneinheit 3.2 haben Sie in der Stage **Lookup** eine Option angegeben, die festlegt, dass mit der Verarbeitung der Kundendaten fortgefahren werden soll, wenn die Lookup-Operation fehlschlägt. Alle Zeilen, deren Spalte `CUSTOMER_NUMBER` Werte enthält, für die es in der Lookup-Tabelle keine Übereinstimmung gibt, werden umgangen, und die Spalte `COUNTRY` für diese Zeile wird auf `NULL` gesetzt. GlobalCo Worldwide möchte diese zurückgewiesenen Kundendatensätze erfassen, damit sie von ihrem Kundenberatungsteam analysiert und aktualisiert werden können.

Kommen Sie diesem Anliegen nach, indem Sie angeben, dass alle Kundennummern, die mit keinem Eintrag in der Lookup-Tabelle übereinstimmen, in eine Zurückweisungsdatei geschrieben werden sollen.

Vorgehensweise

1. Öffnen Sie den Job `Dev_CustData_CleansePrepare`.
2. Fügen Sie Ihrem Job unter der Stage `GlobalCo_Country_LKP` eine Stage **Sequential File** (Sequenzielle Datei) hinzu.
 - a. Benennen Sie die Stage **Sequential File** in `GlobalCo_CountryCode_Rej` um.
 - b. Verlinken Sie die Stage `GlobalCo_Country_LKP` mit der Stage `GlobalCo_BillTo_Rejects`.

Da die Stage **GlobalCo_Country_LKP** bereits über einen Ausgabelink verfügt, wird der Zurückweisungslink als gestrichelte Linie angezeigt.
 - c. Benennen Sie den Link in `lnk_CountryCode_Rej` um.Ihr Job sieht wie der folgende Job aus.

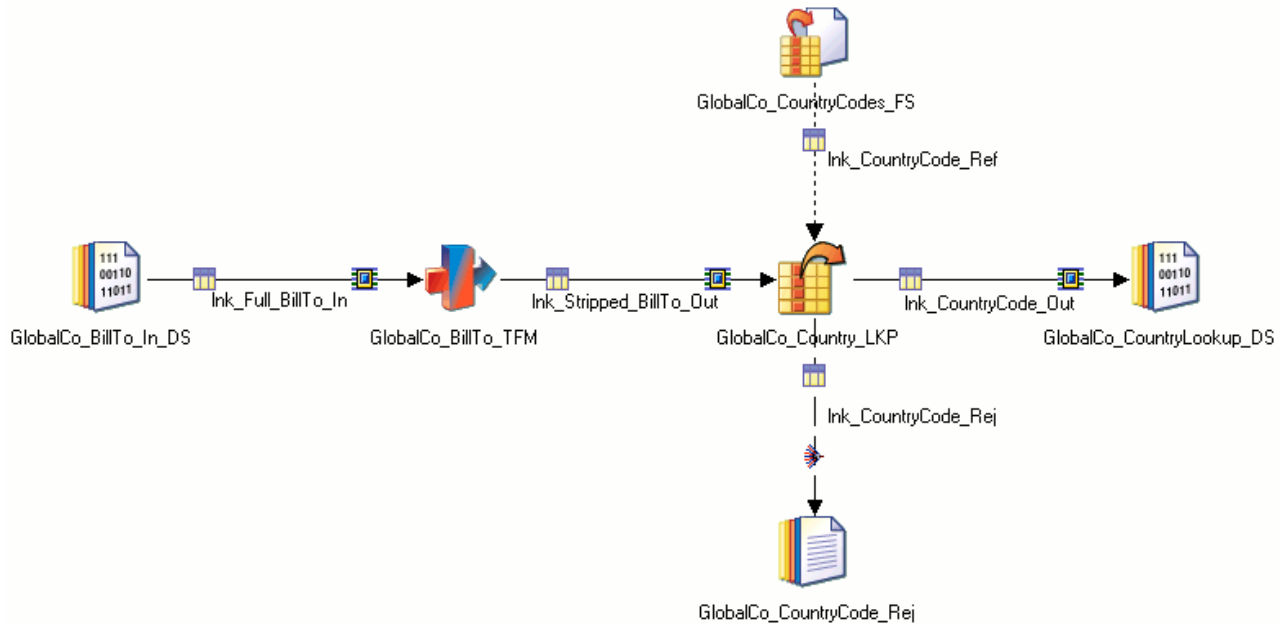


Abbildung 8. Ihr Job enthält einen Zurückweisungslink zum Erfassen von Kundendatensätzen, die mit keinem der Einträge in der Lookup-Tabelle übereinstimmen.

3. Öffnen Sie die Stage `GlobalCo_Country_LKP`.
 - a. Doppelklicken Sie im Link `Ink_CountryCode_Ref` auf **Bedingung**.
 - b. Wählen Sie **Zurückweisen** in der Spalte **Lookup-Fehler** des Fensters **Stage 'Lookup' - Bedingungen** aus.
Diese Option gibt an, dass die Zeile, wenn ein Wert der Spalte `CUSTOMER_NUMBER` im Link `Ink_Stripped_BillTo_Out` mit keinem Wert in der Spalte `CUSTOMER_NUMBER` im Link `Ink_CountryCode_Ref` übereinstimmt, an die Stage `GlobalCo_CountryCode_Rej` gesendet werden soll.
 - c. Schließen Sie das Fenster **Stage 'Lookup' - Bedingungen** und den Editor der Stage `GlobalCo_Country_LKP`.
4. Öffnen Sie die Stage `GlobalCo_CountryCode_Rej` und geben Sie einen Dateinamen für die Textdatei an, in die die zurückgewiesenen Datensätze geschrieben werden sollen. Beispiel: `C:\IBM\InformationServer\Server\Projects\Tutorial\Tutorial_files\GlobalCo_CountryCode_Rej.txt`.
5. Speichern und kompilieren Sie den Job `Dev_CustData_CleansePrepare` und führen Sie ihn aus.
6. Wenn der Job erfolgreich ausgeführt wird, klicken Sie mit der rechten Maustaste auf die Stage `GlobalCo_CountryCode_Rej` und klicken Sie auf **View Ink_CountryCode_Rej data**.
Das Fenster **Datenbrowser** wird geöffnet. Darin werden die Kundendatensätze angezeigt, die die Lookup-Kriterien nicht erfüllt haben.
7. Klicken Sie mit der rechten Maustaste auf die Stage `GlobalCo_CountryLookup_DS` und klicken Sie auf **View Ink_CountryCode_Out data**.
Das Fenster **Datenbrowser** wird geöffnet. Darin werden die Kundendatensätze angezeigt, die in das Dataset geschrieben wurden.

Sie haben die zurückgewiesenen Kundendatensätze erfasst, die die Lookup-Kriterien nicht erfüllt haben. Diese zurückgewiesenen Datensätze können mit Datensätzen von WorldCo kombiniert werden, damit das Kundenberatungsteam von GlobalCo Worldwide die kombinierten Daten einer Analyse unterziehen kann. Das

Team kann Ungenauigkeiten in den Daten untersuchen, um festzustellen, warum die Lookup-Tabelle keinen übereinstimmenden Landescode für die Kundennummern enthält. Durch Korrigieren dieser Inkonsistenzen wird sichergestellt, dass die Rechnungsstellungsdatensätze von GlobalCo Worldwide korrekt und vollständig sind.

Einige der Datumsangaben in der Spalte `SETUP_DATE` enthalten eine zu große Menge an Fremddaten, die entfernt werden müssen. Darüber hinaus enthält die Spalte `STATUS_CODE` nicht die zugeordnete Beschreibung der Bedeutung der einzelnen Codes. Diese Probleme müssen behoben werden, bevor die Rechnungsstellungsdaten der Kunden in die Datenbank von GlobalCo Worldwide geschrieben werden.

In der nächsten Lerneinheit unterziehen Sie die Rechnungsstellungsdaten der Kunden einer weiteren Bereinigung. Sie verwenden eine weitere Stage **Transformer** und Stage **Lookup**, um die Fremddaten aus der Spalte `SETUP_DATE` zu entfernen und fügen die Beschreibungen für die Spalte `STATUS_CODE` hinzu.

Lerneinheit 3.4: Mehrere Umsetzungen in einem einzigen Job abschließen

In dieser Lerneinheit nehmen Sie weitere Umsetzungen an Ihren Daten vor, um Geschäftsregeln anzuwenden, und führen anschließend ein Lookup durch, bei dem eine weitere Referenztable verwendet wird.

Übersicht

Nach dem Durcharbeiten von Lerneinheit 3.3 haben Sie die Ausgabedaten angezeigt, um Ihre Ergebnisse zu prüfen. Sie haben festgestellt, dass einer der Spalten im Dataset `GlobalCo_BillTo.ds` Fremddaten enthält und dass diese Fremddaten nun in Ihrem Ausgabedataset vorhanden sind. Einige Werte in der Spalte `SETUP_DATE` enthalten neben dem Datum, an dem das Account eingerichtet wurde, noch einen Code für Sonderbehandlung. Diese Werte müssen separiert, das heißt, in die entsprechenden Spalten aufgeteilt werden, bevor die Daten mit den Kundendaten von WorldCo kombiniert werden.

Zum Separieren der Werte fügen Sie eine Umsetzungslogik hinzu, die den Code für Sonderbehandlung aus der Spalte `SETUP_DATE` in eine separate Spalte extrahiert. Anschließend fügen Sie eine Lookup-Operation hinzu, um die Textbeschreibung zu suchen, die diesem Code in der Lookup-Tabelle entspricht, die Sie in Lerneinheit 2.3 gefüllt haben. Diese Beschreibung schreiben Sie in die Ausgabedaten und fügen darüber hinaus eine Zeilenanzahl hinzu.

Tasks in dieser Lerneinheit

1. „Umsetzungs- und Lookup-Operationen hinzufügen“
2. „Umsetzungslogik zum Job hinzufügen“ auf Seite 48
3. „Lookup-Operation konfigurieren“ auf Seite 52

Umsetzungs- und Lookup-Operationen hinzufügen

Damit in Ihrem Job mehrere Umsetzungen vorgenommen werden, fügen Sie eine weitere Stage **Transformer** (Umsetzung) und weitere Stages **Lookup** hinzu.

1. Öffnen Sie den Job `Dev_CustData_CleansePrepare`.
2. Löschen Sie die Stage `GlobalCo_CountryLookup_DS`. Sie erstellen ein neues Dataset für die Ausgabe.

3. Fügen Sie rechts neben der Stage GlobalCo_Country_LKP eine Stage **Transformer** hinzu.
 - a. Benennen Sie die Stage in SpecHandling_Extract_TFM um.
 - b. Ziehen Sie den Link lnk_CountryCode_Out zur Stage SpecHandling_Extract_TFM.
 Da die Logik jeweils bei dem Link verbleibt, bei dem sie definiert wurde, fließen die von Ihnen definierten Spalten von der ersten Stage **Lookup** zur neuen Stage **Transformer**.
4. Fügen Sie rechts neben der Stage SpecHandling_Extract_TFM eine Stage **Lookup** hinzu.
 - a. Benennen Sie die Stage in SpecHandling_Code_LKP um. Diese Stage führt auf der Suche nach der entsprechenden Textbeschreibung ein Lookup nach dem Code für Sonderbehandlung durch.
 - b. Verlinken Sie die Stage SpecHandling_Extract_TFM mit der Stage SpecHandling_Code_LKP.
 - c. Benennen Sie den Link in lnk_SpecHandlingCode_Out um.
5. Fügen Sie oberhalb der Stage SpecHandling_Code_LKP eine Stage **Lookup File Set** (Lookup-Dateigruppe) hinzu.
 - a. Benennen Sie die Stage in SpecHandling_Desc_FS um.
 - b. Verlinken Sie die Stage SpecHandling_Desc_FS mit der Stage SpecHandling_Code_LKP.
 - c. Benennen Sie den Link in lnk_SpecHandlingDesc_Ref um.
6. Fügen Sie unter der Stage SpecHandling_Code_LKP eine Stage **Data Set** hinzu.
 - a. Benennen Sie die Stage in GlobalCo_CustData_DS um.
 - b. Verlinken Sie die Stage GlobalCo_CustData_DS mit der Stage SpecHandling_Code_LKP.
 - c. Benennen Sie den Link in lnk_CustData_Out um.

Ihr Job sieht wie der folgende Job aus:

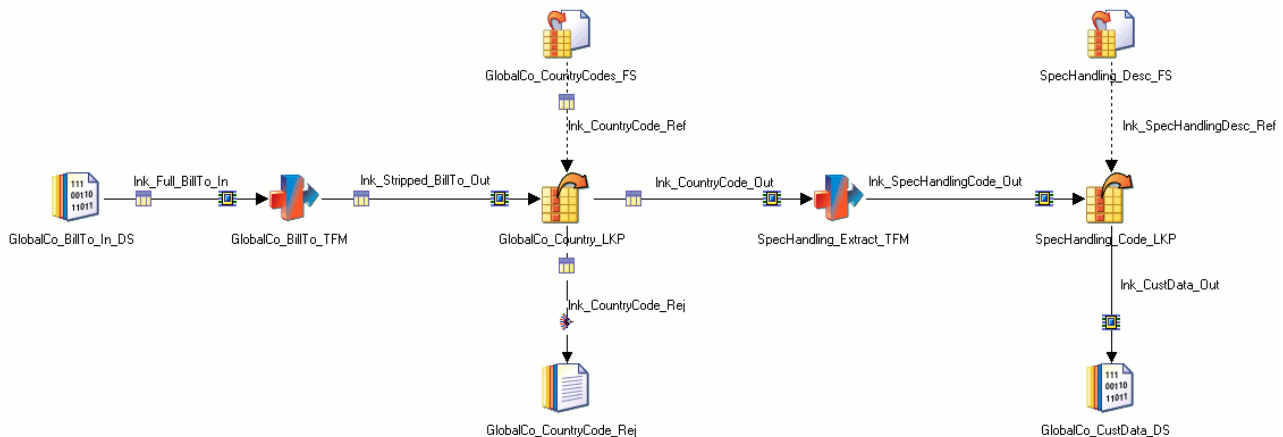


Abbildung 9. Ihr Job enthält zur weiteren Bereinigung der Kundendaten von GlobalCo eine zweite Stage **Transformer**.

Umsetzungslogik zum Job hinzufügen

Fügen Sie eine Umsetzungslogik hinzu, um den Code für Sonderbehandlung zu extrahieren, und fügen Sie dann den Ausgabedaten eine Zeilenanzahl hinzu.

1. Öffnen Sie die Stage SpecHandling_Extract_TFM.

2. Wählen Sie im linken Teilfenster des Fensters **Stage 'Transformer'** die folgenden Spalten im Link `lnk_CountryCode_Out` aus und ziehen Sie sie zum Link `lnk_SpecHandlingCode_Out` im rechten Teilfenster.

CUSTOMER_NUMBER
 CUST_NAME
 ADDR_1
 ADDR_2
 CITY
 REGION_CODE
 ZIP
 TEL_NUM

3. Fügen Sie im Teilfenster rechts unten die folgenden Spalten für den Link `lnk_SpecHandlingCode_Out` hinzu. Verwenden Sie die Standardwerte für alle Felder, die in der folgenden Tabelle nicht angegeben sind.

Wichtig: Die Spalten, die Sie hinzufügen, werden im Link `lnk_SpecHandlingCode_Out` rot hervorgehoben, da sie über keine gültigen Ableitungen verfügen. Sie fügen die Ableitungen hinzu, um anzugeben, wie die einzelnen Spalten von der Stage **Transformer** verarbeitet werden sollen.

Spaltenname	SQL-Typ	Länge	Kann Nullwerte enthalten	Beschreibung
SOURCE	Char	10	Nein	Gibt das Unternehmen an, von dem die Kundendaten stammen.
RECNUM	Char	10	Nein	Gibt die Anzahl Zeilen für die einzelnen Datensätze an.
SETUP_DATE	Char	10	Ja	Überprüft, ob die Spalte SETUP_DATE ein Datum enthält.
SPECIAL_HANDLING_CODE	Integer	10	Ja	Gibt an, ob die Spalte SETUP_DATE einen Code für Sonderbehandlung enthält.

4. Doppelklicken Sie im Link-Container `lnk_SpecHandlingCode_Out` auf das Feld **Ableitung** für die Spalte SOURCE.
 - a. Geben Sie im Ausdruckeditor 'GlobalCo': ein.
 - b. Positionieren Sie den Cursor rechts neben den eingegebenen Text, klicken Sie mit der rechten Maustaste und klicken Sie auf **Eingabespalte**.
 - c. Wählen Sie die Spalte COUNTRY aus und drücken Sie dann die Eingabetaste, um die Ableitung zu übergeben.


Diese Ableitung fügt dem aus zwei Buchstaben bestehenden Landescode den Präfix „GlobalCo“ hinzu. Beispielsweise wird, wenn Sie Ihren Job ausführen, ein Kundendatensatz mit dem Landescode US zu GlobalCoUS. Der folgende Code veranschaulicht die Ableitung für die Spalte SOURCE:

```
'GlobalCo': 1nk_CountryCode_Out.COUNTRY
```

5. Führen Sie für die Spalte REC_NUM einen Doppelklick auf das Feld **Ableitung** aus.
 - a. Geben Sie im Ausdruckseditor 'GC': ein.
 - b. Positionieren Sie den Cursor rechts neben den eingegebenen Text, klicken Sie mit der rechten Maustaste und klicken Sie auf **Systemvariable**.
 - c. Wählen Sie @OUTROWNUM aus und drücken Sie dann die Eingabetaste, um die Ableitung zu übergeben.

Diese Ableitung fügt Ihrer Ausgabe Zeilennummern hinzu, um eine leichter lesbare Formatierung zu erhalten. Darüber hinaus können andere Entwickler auf die Kundendatensätze über ihre Zeilennummer verweisen, anstatt die Ausgabe nach dem Zufallsprinzip durchsuchen zu müssen. Der folgende Code veranschaulicht die Ableitung für die Spalte REC_NUM:

```
'GC': @OUTROWNUM
```

6. Klicken Sie in der Symbolleiste des Editors für die Stage **Transformer** auf das Symbol **Stage-Eigenschaften** ().
 - a. Klicken Sie im Fenster **Eigenschaften der Stage 'Transformer'** auf die Registerkarte **Stagevariablen**.
 - b. Fügen Sie Stagevariablen hinzu. Verwenden Sie die Standardwerte für alle Felder, die in der folgenden Tabelle nicht angegeben sind.

Name	SQL-Typ	Genauigkeit
SpecHandlingExtract	Char	1
TrimSetupDate	VarChar	10

- c. Klicken Sie auf **OK**, um das Fenster **Eigenschaften der Stage 'Transformer'** zu schließen.
Die von Ihnen erstellten Stagevariablen werden im Container **Stagevariablen** aufgelistet.
7. Doppelklicken Sie im Container **Stagevariablen** auf das Feld **Ableitung** für die Variable SpecHandlingExtract.
 - a. Geben Sie den folgenden Ausdruck ein und drücken Sie die Eingabetaste. Sie können im Ausdruckseditor mit der rechten Maustaste klicken, um wie in Schritt 4 Spalten hinzuzufügen.

```
IF LEN(1nk_CountryCode_Out.SETUP_DATE) < 2 THEN 1nk_CountryCode_Out.SETUP_DATE  
ELSE FIELD(1nk_CountryCode_Out.SETUP_DATE, ' ',2)
```

Dieser Ausdruck gibt an, dass in der Spalte SPEC_HANDLING_CODE der Wert aus der Spalte SETUP_DATE verwendet wird, wenn die Anzahl Zeichen in der Spalte SPEC_HANDLING_CODE kleiner als 2 ist. Andernfalls wird der Wert der Spalte SETUP_DATE zurückgegeben und das zweistellige Zeichen wird aus dieser Spalte extrahiert und in die Spalte SPEC_HANDLING_CODE versetzt.
 - b. Ziehen Sie die Ableitung für die Variable SpecHandlingExtract zum Feld **Ableitung** für die Spalte SPEC_HANDLING_CODE im Link-Container 1nk_SpecHandlingCode_Out.

Zwischen der Stagevariablen und der Spalte wird eine Linie dargestellt und im Feld **Ableitung** wird der Name SpecHandlingExtract aufgelistet. Für

jede Zeile, die verarbeitet wird, wird der Wert der Variablen SpecHandlingExtract in die Spalte SPECIAL_HANDLING_CODE geschrieben.

8. Doppelklicken Sie im Container **Stagevariablen** auf das Feld **Ableitung** für die Variable TrimSetupDate.
 - a. Geben Sie den folgenden Ausdruck ein und drücken Sie die Eingabetaste. Sie können im Ausdruckseditor mit der rechten Maustaste klicken, um wie in Schritt 4 und 7 Spalten hinzuzufügen.

```
IF LEN(1nk_CountryCode_Out.SETUP_DATE) < 3 THEN '01/01/0001'  
ELSE FIELD(1nk_CountryCode_Out.SETUP_DATE, ' ',1)
```

Dieser Ausdruck gibt an, dass der Wert der Variablen TrimSetupDate auf '01/01/0001' gesetzt werden soll, wenn die Spalte SETUP_DATE kein Datum enthält. Andernfalls wird der Wert der Spalte SETUP_DATE in der Variablen TrimSetupDate verwendet.
 - b. Ziehen Sie die Ableitung für die Variable TrimSetupDate zum Feld **Ableitung** für die Spalte SETUP_DATE im Link-Container 1nk_SpecHandlingCode_Out.
Zwischen der Stagevariablen und der Spalte wird eine Linie dargestellt und der Name TrimSetupDate wird im Feld **Ableitung** aufgelistet. Für jede Zeile, die verarbeitet wird, wird der Wert der Variablen TrimSetupDate in die Spalte SETUP_DATE geschrieben.
9. Klicken Sie auf **OK**, um das Fenster **Stage 'Transformer'** zu schließen und speichern Sie anschließend Ihren Job.

Ihr Stage-Editor sieht wie der folgende Stage-Editor aus:

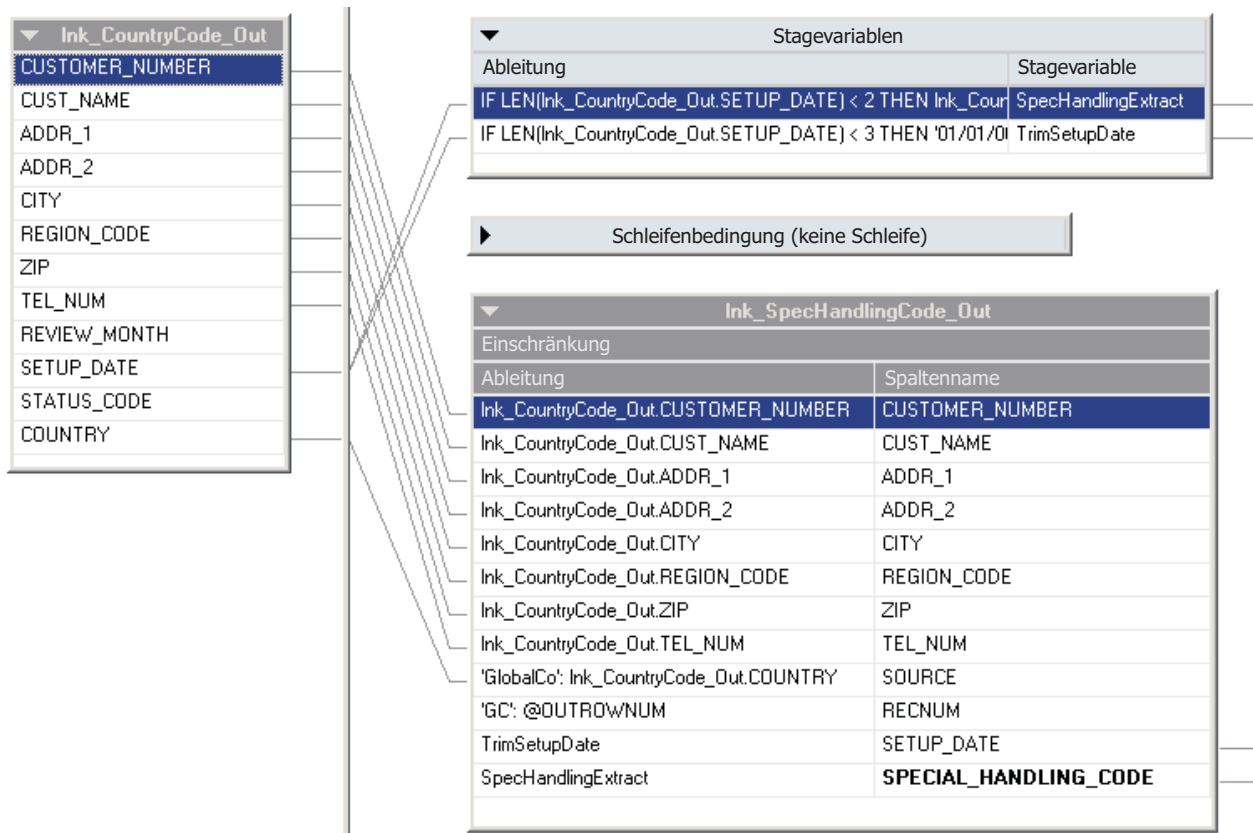


Abbildung 10. Sie ordnen die Spalten des Eingabelinks dem Ausgabelink zu. Die Stagevariablen, die Sie hinzufügen, stellen eine zusätzliche Umsetzungslogik zum Beheben von Inkonsistenzen und Fehlern in den Kundendaten von GlobalCo bereit.

Lookup-Operation konfigurieren

In dieser Übung konfigurieren Sie die Stages, die erforderlich sind, um den Code für Sonderbehandlung zu suchen und ihn in ein Ausgabedataset zu schreiben.

1. Öffnen Sie die Stage SpecHandling_Desc_FS.
 - a. Legen Sie die Eigenschaft **Lookup-Dateigruppe** so fest, dass sie auf den Parameter SpecHandlingCode_In_Source verweist.
 - b. Klicken Sie auf die Registerkarte **Spalten** und laden Sie den Tabellenparameter SpecialHandling.csv.
 - c. Klicken Sie auf **OK**, um den Stage-Editor zu schließen.
2. Öffnen Sie die Stage SpecHandling_Code_LKP.
 - a. Wählen Sie alle Spalten im Link-Container Ink_SpecHandlingCode_Out aus und ziehen Sie sie zum Link-Container Ink_CustData_Out.
 - b. Wählen Sie die Spalte DESCRIPTION im Link-Container Ink_SpecHandlingDesc_Ref aus und ziehen Sie sie zum Link-Container Ink_CustData_Out.
 - c. Geben Sie im Teilfenster rechts unten eine Beschreibung für die Spalte DESCRIPTION ein. Beispiel: Beschreibung für den Code für Sonderbehandlung.
 - d. Doppelklicken Sie im Link-Container Ink_SpecHandlingDesc_Ref auf die Leiste **Bedingung**.

- e. Wählen Sie **Fortfahren** im Feld **Lookup-Fehler** aus und klicken Sie dann auf **OK**.

Wenn Sie am Ende von Lerneinheit 3.3 die Ausgabedaten überprüft haben, konnten Sie feststellen, dass wenige Zeilen in den Beispieldaten einen Code für Sonderbehandlung enthalten. Daher soll der Job mit der Verarbeitung der Daten fortfahren, wenn die Zeilen, die keinen Code enthalten, zurückgewiesen werden.

Ihr Stage-Editor sieht wie der folgende Stage-Editor aus:

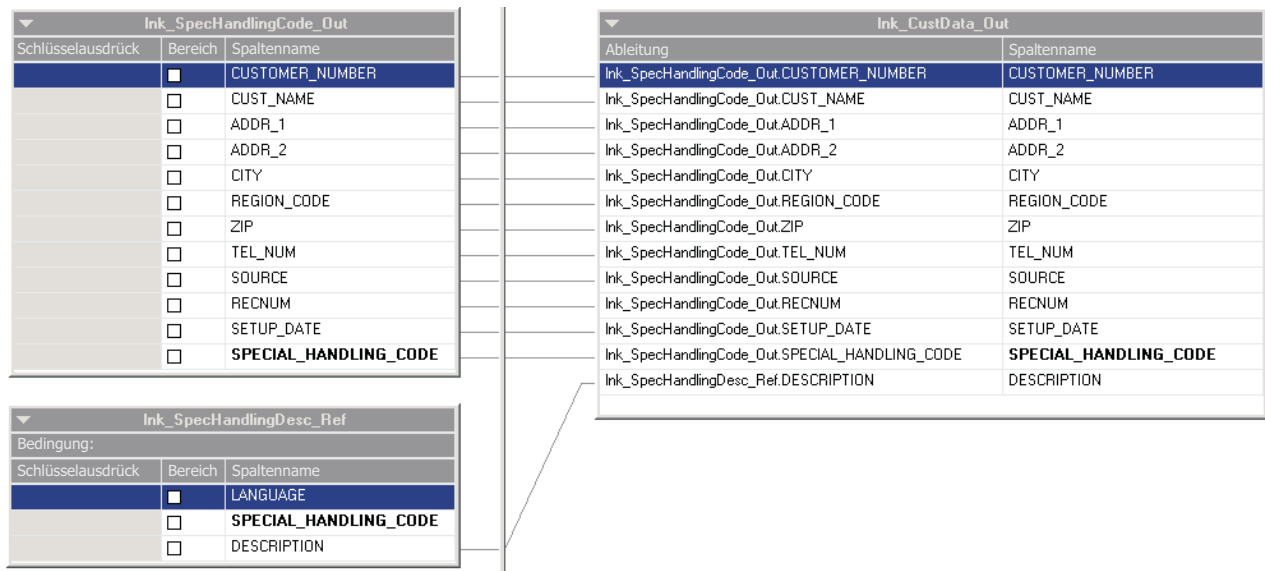


Abbildung 11. Sie ordnen die Spalten des Eingabelinks dem Ausgabeblick zu. Die Spalte **SPECIAL_HANDLING_CODE** wird in der Stage **Lookup** als Schlüssel zum Referenzieren von Kundendaten verwendet.

- f. Klicken Sie auf **OK**, um den Editor der Stage **SpecHandlingCode_LKP** zu schließen.
3. Öffnen Sie die Stage **GlobalCo_CustData_DS**.
- a. Klicken Sie im Feld **Datei** auf das Symbol **Durchsuchen** und klicken Sie dann auf **Jobparameter einfügen > Neu**.
- b. Geben Sie im Fenster **Jobeigenschaften** die folgenden Werte für die Jobparameter ein.

Parametername	Eingabe	Typ	Standardwert	Hilfetext
CustData_Output_dataset	Zieldatei für die Ausgabe der Kundendaten	Pfadname	Lernprogramm- ordner\ GlobalCo_Target _Output.ds	Zieldatei, die die bereinigten Kundendaten von GlobalCo enthält

Lernprogrammordner ist der Ordner, in dem Sie die Dateien für das Lernprogramm speichern wollen, beispielsweise C:\IBM\InformationServer\Server\Projects\Tutorial\Tutorial_files.

- c. Klicken Sie auf **OK**, um das Fenster **Jobeigenschaften** zu schließen.
- d. Geben Sie in der Eigenschaft **Datei** den Parameter **CustData_Output_dataset** an und drücken Sie dann die Eingabetaste.
- e. Klicken Sie auf **OK**, um den Editor der Stage **GlobalCo_CustData_DS** zu schließen.

4. Speichern und kompilieren Sie den Job `SpechHandling_Extract_TFM` und führen Sie ihn aus.
5. Wenn der Job erfolgreich ausgeführt wird, klicken Sie mit der rechten Maustaste auf die Stage `GlobalCo_CustData_DS` und klicken Sie auf **View Ink_CustData_Out data**.
6. Klicken Sie im Fenster **Jobparameter auflösen** auf **OK**.
In dem daraufhin geöffneten Fenster **Datenbrowser** werden die bereinigten Kundendaten angezeigt.

In der Spalte `SOURCE` ist dem zweistelligen Landescode der Name `GlobalCo` vorangestellt. Das Einschließen dieser Information ist wichtig, damit die Entwickler die `GlobalCo`-Kundendaten und die `WorldCo`-Kundendaten voneinander unterscheiden können, wenn die bereinigten Daten zusammengefasst werden.

Die Spalte `SPECIAL_HANDLING_CODE` enthält für jeden Kundendatensatz einen Code für Sonderbehandlung. Die entsprechende Beschreibung der einzelnen Codes ist in der Spalte `DESCRIPTION` aufgelistet. Diese Information ermöglicht es der Versandabteilung von `GlobalCo Worldwide` zu bestimmen, wann die einzelnen Kunden ihre Lieferung erhalten.

Weitere Schritte

In dieser Lerneinheit haben Sie eine strengere Datentypisierung angewendet, um an den Rechnungsstellungsdaten von `GlobalCo` weitere Umsetzungen und Bereinigungen vorzunehmen. Bislang haben Sie Ihre Ausgabedaten in Datasets geschrieben. Im folgenden Modul schreiben Sie die Ausgabedaten in eine relationale Datenbank, damit die Daten für andere Entwickler im Unternehmen `GlobalCo Worldwide` leicht zugänglich sind.

Sie haben die folgenden Tasks ausgeführt:

- Definieren von Stagevariablen in einer Stage **Transformer**
- Filtern von Spaltenmetadaten in einer Stage **Transformer**
- Verwenden von Systemvariablen zum Generieren von Ausgabespaltenwerten
- Konfigurieren von Stages **Lookup** zum Verbessern der Ausgabedaten
- Durch Einsichtnahme in das Jobprotokoll Aufschluss über die Komponenten gewonnen, aus denen sich der Job zusammensetzt

Kapitel 6. Modul 4: Daten in eine relationale Datenbank laden

In diesem Modul schreiben Sie die Daten `GlobalCo_billing` in eine relationale Datenbank. Diese Datenbank ist das Endziel für die GlobalCo-Rechnungsstellungsdaten, die Sie umgesetzt und bereinigt haben.

In den vorherigen Modulen, die Sie durchgearbeitet haben, verwendeten Sie CSV-Dateien und Dateien für die Zwischenspeicherung in internen Formaten wie Data-sets und Lookup-Dateigruppen. In diesem Modul verwenden Sie eine relationale Datenbank als Ziel.

Für diese Lerneinheiten verwenden Sie die Datenbank, die als Host für das Repository des Lernprogramms fungiert. Das Lernprogramm stellt Scripts zur Verfügung, die Ihr Datenbankadministrator ausführt, um die für diese Lerneinheiten erforderlichen Tabellen zu erstellen.

Da verschiedene Typen von relationalen Datenbanken als Host für das Repository des Lernprogramms verwendet werden können, wird in diesem Modul eine ODBC-Verbindung verwendet. Bei diesem Verbindungstyp können Sie unabhängig von dem von Ihnen verwendeten Datenbanktyp den Anweisungen folgen.

Lernziele

Nach dem Durcharbeiten der Lerneinheiten in diesem Modul werden Sie in der Lage sein, die folgenden Tasks auszuführen:

- Definieren eines Datenverbindungsobjekts zum Herstellen einer Verbindung zu einer Datenbank
- Importieren von Spaltenmetadaten aus einer Datenbank
- Schreiben von Daten in eine relationale Datenbank als Ziel

Erforderlicher Zeitaufwand

Für dieses Modul sind ungefähr 60 Minuten erforderlich.

Voraussetzungen

Stellen Sie sicher, dass Ihr Datenbankadministrator die Tasks in „Datenquellennamen für die Tabelle des Lernprogramms für parallele Jobs erstellen“ auf Seite 6 ausgeführt hat. Sie verwenden den von Ihrem Datenbankadministrator erstellten Datenquellennamen zum Herstellen einer ODBC-Verbindung.

Lerneinheit 4.1: Datenverbindungsobjekt erstellen

In dieser Lerneinheit erstellen Sie ein Datenverbindungsobjekt zum Speichern der Informationen, die für das Herstellen einer Verbindung zu der Datenbank erforderlich sind, in die Sie die bereinigten Kundenrechnungsstellungsdaten schreiben.

Übersicht


Sie verwenden dieses Datenverbindungsobjekt mit einer Stage **ODBC Connector** zum Definieren einer Verbindung zur Datenquelle von GlobalCo Worldwide. Sie

können Datenverbindungsobjekte auch verwenden, um die Details bereitzustellen, die zum Herstellen einer Verbindung zu einer Datenquelle und zum Importieren von Metadaten erforderlich sind.

Weitere Informationen zu Datenverbindungsobjekten: Wenn Sie die Details einer Datenverbindung ändern, während Sie einen Job entwerfen, spiegeln sich die Änderungen in dem Jobentwurf wider. Nachdem Sie jedoch den Job kompiliert haben, sind die Details der Datenverbindung in der ausführbaren Version des Jobs festgelegt. Bei nachfolgenden Änderungen am Jobentwurf wird ein Link zum Datenverbindungsobjekt hergestellt und die an diesem Objekt vorgenommenen Änderungen werden erfasst.

Sie können Datenverbindungsobjekte direkt im Repository von Designer-Client erstellen. Außerdem können Sie Datenverbindungsobjekte durch Speichern der Verbindungsdetails erstellen. Verwenden Sie diese Option, wenn Sie Metadaten mithilfe einer Connector-Stage importieren. In dieser Lerneinheit erstellen Sie ein Datenverbindungsobjekt direkt im Designer-Client-Repository.

Datenverbindungsobjekt erstellen

1. Klicken Sie im **Repository** des Designer-Clients auf den Ordner **Tutorial** und klicken Sie dann im Direktaufrufenmenü auf **Neu > Sonstige > Datenverbindung**.
 - a. Geben Sie im Fenster **Datenverbindung** einen Namen für das Datenverbindungsobjekt ein. Beispiel: `GlobalCo_billing_connect`
 - b. Geben Sie eine kurze Beschreibung ein. Beispiel: Verbindungsobjekt für die Rechnungsstellungsdaten von GlobalCo.
 - c. Geben Sie eine ausführliche Beschreibung ein, die die Details dieser Verbindung beschreibt, z. B. den Namen der Datenbank, zu der Sie eine Verbindung herstellen, und die Metadaten, die diese Verbindung darstellt.
2. Klicken Sie auf die Registerkarte **Parameter**.
 - a. Klicken Sie im Feld **Verbindung über Stagetyp herstellen** auf das Symbol **Durchsuchen** ().
 - b. Erweitern Sie im Fenster **Öffnen** den Ordner **Stagetypen > Parallel > Datenbank**.
 - c. Klicken Sie auf **ODBC-Connector** und klicken Sie dann auf **Öffnen**.
 - d. Geben Sie Werte für die einzelnen Parameter wie in der folgenden Tabelle dargestellt ein.

Parametername	Wert
Verbindungszeichenfolge	<i>Name des Datenquellennamens über den Sie eine Verbindung zur Datenbank herstellen</i>
Benutzername	<i>Benutzername für den Datenquellennamen über den Sie eine Verbindung zur Datenbank herstellen</i>
Kennwort	<i>Kennwort für den Benutzernamen, der dem Datenquellennamen entspricht</i>

- e. Klicken Sie auf **OK**, um die Datenverbindungsdetails zu übergeben.
3. Wählen Sie im Fenster **Datenverbindung - Speichern unter** den Ordner **Tutorial** aus und klicken Sie auf **Speichern**.

Sie haben ein Datenverbindungsobjekt erstellt und das Objekt im Repository gespeichert. In der nächsten Lerneinheit importieren Sie Spaltenmetadaten aus einer Datenbanktabelle. Sie ordnen die Spaltenmetadaten den Rechnungsstellungsdaten von GlobalCo aus Lerneinheit 3.4 zu.

Lerneinheit 4.2: Spaltenmetadaten aus einer Datenbanktabelle importieren

In dieser Lerneinheit importieren Sie Spaltenmetadaten aus einer Datenbanktabelle und speichern sie als Tabellendefinitionsobjekt im Repository.

Übersicht

In Lerneinheit 2.3 haben Sie Spaltenmetadaten aus einer CSV-Datei importiert. In dieser Lerneinheit importieren Sie Spaltenmetadaten über einen ODBC-Connector aus einer Datenbanktabelle. Sie verwenden die importierte Tabellendefinition in der nächsten Lerneinheit, um die bereinigten Rechnungsstellungsdaten der Kunden von GlobalCo in die Datenbank von GlobalCo Worldwide zu schreiben.

Wenn Sie Daten über einen Connector importieren, werden die Spaltendefinitionen als Tabellendefinition im Designer-Client-Repository gespeichert. Die Tabellendefinition kann dann von anderen Projekten und Komponenten in der Informationsintegrationssuite verwendet werden.

Vorgehensweise

1. Klicken Sie im Designer-Client-Menü auf **Importieren > Tabellendefinitionen > Assistenten für Import über Connector starten**.
2. Wählen Sie im Fenster für den Import von Connectormetadaten den Typ des Datenbankconnectors aus, aus dem Sie Daten importieren möchten, und klicken Sie auf **Weiter**.
 - a. Klicken Sie im Teilfenster **Verbindung** auf **Laden**.
 - b. Erweitern Sie im Fenster **Öffnen** den Ordner **Tutorial**, wählen Sie das Datenverbindungsobjekt aus, das Sie in Lerneinheit 4.1 erstellt haben, und klicken Sie dann auf **Öffnen**.

Die Felder **Datenquelle**, **Benutzername** und **Kennwort** werden mit den entsprechenden Daten gefüllt, die von Ihnen im Datenverbindungsobjekt gespeichert wurden.
 - c. Klicken Sie auf **Testverbindung**, um sicherzustellen, dass Sie eine Verbindung zur Datenbank herstellen können. Wenn die Verbindung fehlschlägt, geben Sie den Benutzernamen und das Kennwort ein und testen Sie die Verbindung erneut.
 - d. Klicken Sie auf **Weiter**, um fortzufahren.
3. Stellen Sie auf der Seite **Position der Datenquelle** sicher, dass der Name des Hosts, auf dem sich die Datenbank befindet, und der Name der Datenbank richtig angegeben sind und klicken Sie dann auf **Weiter**.
4. Wählen Sie auf der Seite **Filtern** das Schema für die Datenbank aus der Liste aus und klicken Sie dann auf **Weiter**. Wenn Sie sich über das zu verwendende Datenbankschema unsicher sind, wenden Sie sich an Ihren Datenbankadministrator.
5. Wählen Sie auf der Seite **Auswahl** die Tabelle für das Lernprogramm aus der Liste aus und klicken Sie dann auf **Weiter**. Die von Ihrem Administrator erstellte Tabelle könnte z. B. den Namen TUTORIAL haben.
 - a. Klicken Sie auf **Daten anzeigen**.

- b. Stellen Sie im Fenster zum Durchsuchen von Daten sicher, dass Sie die die Rechnungsstellungsdaten von GlobalCo importieren.
 - c. Klicken Sie auf **Schließen**, um das Fenster zum Durchsuchen von Daten zu schließen.
 - d. Klicken Sie auf der Seite **Auswahl** auf **Weiter**.
6. Klicken Sie im Teilfenster **Bestätigung** auf **Importieren**.
 7. Erweitern Sie im Fenster **Ordner auswählen** den Ordner Tutorial, wählen Sie den Ordner Tabellendefinitionen aus und klicken Sie dann auf **OK**.

Die Tabellendefinition wird importiert und im Ordner Tutorial\Table definitions des Repositorys aufgelistet. Die Tabellendefinition hat ein anderes Symbol als die zuvor von Ihnen verwendeten Tabellendefinitionen. Dieses Symbol gibt an, dass die Tabellendefinition über einen Connector importiert wurde und für andere Projekte und Suitekomponenten verfügbar ist.

Sie haben mithilfe eines Connectors eine Tabellendefinition aus einer Datenbank importiert. Die Tabellendefinition basiert auf dem Datenverbindungsobjekt, das Sie in Lerneinheit 4.1 erstellt haben.

In der nächsten Lerneinheit erstellen Sie einen Job, der die von Ihnen importierte Tabellendefinition in eine Datenbank schreibt. In dieser Datenbank werden die Rechnungsstellungsdaten von GlobalCo und die von WorldCo gespeichert, bevor Sie zusammengefasst werden.

Lerneinheit 4.3: Daten in eine Datenbank schreiben

In dieser Lerneinheit schreiben Sie die Rechnungsstellungsdaten von GlobalCo, die Sie in Modul 3 erstellt haben, über einen ODBC-Connector in eine vorhandene Tabelle in der Datenbank.

Übersicht

Im Rahmen der Konfiguration des Lernprogramms hat Ihr InfoSphere DataStage und QualityStage-Administrator ein Script ausgeführt, um eine Zieltabelle in einer Datenbank zu erstellen. In dieser Lerneinheit importieren Sie die Spaltendefinitionen aus dieser Tabelle, die im fertig gestellten Data-Warehouse die Dimensionstabelle GlobalCo_billing bildet. Ein anderer Entwickler fasst diese Dimensionstabelle mit der Dimensionstabelle WorldCo_billing zusammen, um die Rechnungsstellungsinformationen für GlobalCo Worldwide zu erstellen.

Weitere Informationen zu Connectors: Connector sind Stages, über die Sie eine Verbindung zu Datenquellen und Datenzielen herstellen. Abhängig vom Typ der Datenbank, aus der Sie lesen bzw. in die Sie schreiben sind verschiedene Connectorarten vorhanden. Erweitern Sie zum Anzeigen der Connector, die Sie verwenden können, in der Palette des Designer-Clients die Kategorie **Datenbank**. Einige Connector gibt es in verschiedenen Versionen. Beispielsweise enthalten IBM DB2-Datenbanken einen Standarddatenbankconnector und einen Enterprise-Connector.

Die Verwendung von Connector-Stages in Ihren Jobs bietet Ihnen im Vergleich zu anderen Stages die folgenden Vorteile:

- Erstellen von Jobparametern über die Connector-Stage, ohne die Jobparameter in den Jobeigenschaften definieren zu müssen
- Speichern von Verbindungsinformationen, die Sie in der Stage angeben, als Datenverbindungsobjekt

- Abstimmen der Datentypen zwischen Quellen und Zielen, um Fehler bei der Ausführung zu vermeiden
- Generieren detaillierter Fehlerinformationen, wenn bei einem Connector während der Ausführung des Jobs Probleme auftreten

Job erstellen

1. Erstellen Sie einen neuen parallelen Job mit dem Namen GlobalCo_Cust-Data_DB_Write.
2. Speichern Sie den Job im Ordner Tutorial\Jobs des Repositoriums.
3. Fügen Sie Ihrem Job eine Stage **Data Set** hinzu und nennen Sie die Stage GlobalCo_BillTo_Source.
4. Fügen Sie Ihrem Job einen Connector hinzu.
 - a. Erweitern Sie die Kategorie **Datenbank** in der Palette.
 - b. Klicken Sie auf **ODBC Connector** und fügen Sie den Connector rechts neben der Stage GlobalCo_BillTo_Source hinzu.
 - c. Benennen Sie die Stage in GlobalCo_BillTo_Target um.
5. Verlinken Sie die Stages und benennen Sie den Link in lnk_GlobalCo_BillTo_Out um.

Ihr Job sieht wie der folgende Job aus.

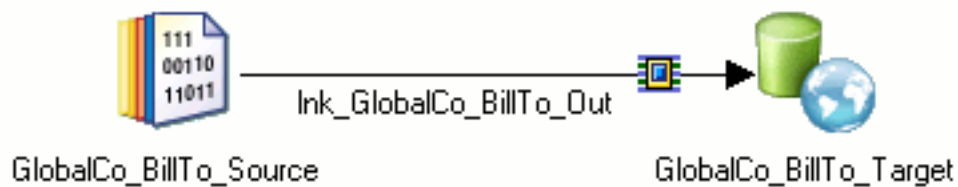


Abbildung 12. Ihr Job schreibt die bereinigten Rechnungsstellungsdaten der Kunden von GlobalCo in die gemeinsam genutzte Datenbank.

Ziel konfigurieren

In dieser Übung verwenden Sie das in Lerneinheit 3.4 erstellte Dataset mit Rechnungsstellungsdaten von GlobalCo als Quelle für Ihren Job.

In dieser Übung verwenden Sie die in Lerneinheit 4.2 importierte Tabellendefinition. Die Spaltendefinitionen sind mit der Tabellendefinition identisch, die Sie in Lerneinheit 3.4 durch Bearbeiten der Stage **Transformer** (Umsetzung) und der Stage **Lookup** im Job erstellt haben.

1. Öffnen Sie die Stage GlobalCo_BillTo_Source.
2. Klicken Sie auf die Eigenschaft **Datei** und erstellen Sie dann einen neuen Parameter, der die folgenden Werte verwendet. Sie verwenden das in Lerneinheit 3.4 erstellte Dataset GlobalCo_Target_Output.ds als Quelldaten.

Parametername	Eingabe	Typ	Standardwert	Hilfetext
CustData_Input_dataset	Quellen-datei für die Kundendaten-eingabe	Pfadname	Lernprogrammordner\GlobalCo_Target_Output.ds	Quellen-datei, die die bereinigten Kundendaten von GlobalCo enthält

3. Laden Sie die Spaltenmetadaten.
 - a. Klicken Sie auf die Registerkarte **Spalten** und klicken Sie dann auf **Laden**.
 - b. Erweitern Sie im Fenster **Tabellendefinitionen** den Ordner Tutorial\Table definitions. Wählen Sie die Datenbanktabelle aus, die Sie in Lerneinheit 4.2 erstellt haben, und klicken Sie dann auf **OK**.
 - c. Stellen Sie im Fenster **Spalten auswählen** sicher, dass alle Spalten ausgewählt sind und klicken Sie dann auf **OK**.
Die Registerkarte **Spalten** wird mit den von Ihnen importierten Spaltenmetadaten gefüllt.
4. Klicken Sie auf **OK**, um den Editor der Stage GlobalCo_BillTo_Source zu schließen, und speichern Sie dann Ihren Job.

ODBC-Connector konfigurieren

In dieser Übung konfigurieren Sie den Datenbankconnector, um Informationen bereitzustellen, die für das Schreiben in die Datenbanktabelle erforderlich sind. Sie definieren über SQL Builder die SQL-Anweisung, die zum Schreiben der Daten in die Datenbank verwendet wird, wenn Sie den Job ausführen.

1. Öffnen Sie die Stage GlobalCo_BillTo_Target.
2. Laden Sie die Verbindungseigenschaften.
 - a. Klicken Sie im Stage-Editor auf das Stagesymbol.
 - b. Klicken Sie im Teilfenster **Verbindung** der Registerkarte **Eigenschaften** auf **Laden**.
 - c. Erweitern Sie im Fenster **Öffnen** den Ordner **Tutorial**, wählen Sie das Datenverbindungsobjekt aus, das Sie in Lerneinheit 4.1 erstellt haben, und klicken Sie dann auf **Öffnen**. Ihr Objekt könnte beispielsweise den Namen GlobalCo_billing_connect haben.
Die Eigenschaften **Datenbank**, **Benutzername** und **Kennwort** werden mit den Werten aus dem Datenverbindungsobjekt gefüllt.
 - d. Klicken Sie im Abschnitt **Verbindung** der Registerkarte **Eigenschaften** auf **Test**, um die Verbindung zu testen.
Eine Nachricht wird angezeigt und informiert über die erfolgreiche Herstellung der Verbindung. Wenn die Verbindung fehlschlägt, überprüfen Sie die Verbindungsdetails und testen Sie die Verbindung dann erneut.
3. Fügen Sie die SQL-Anweisung ein und konfigurieren Sie sie.
 - a. Klicken Sie auf die Registerkarte **Eingabe**.
 - b. Erweitern Sie den Abschnitt **Verwendung** der Registerkarte **Eigenschaften** und setzen Sie den **Schreibzugriff** auf **Einfügen**. Diese Auswahl aktiviert das Feld Anweisung INSERT in der Kategorie **SQL**.

- c. Erweitern Sie die Kategorie **SQL**, klicken Sie auf die Zeile **Anweisung INSERT** und klicken Sie dann auf **Erzeugung > Neues SQL erzeugen**.
- d. Erweitern Sie im Fenster **SQL Builder** den Ordner Tutorial\Table definitions.
- e. Erweitern Sie das Projekt Tutorial und wählen Sie dann die Tabellendefinition TUTORIAL aus, die Sie in Lerneinheit 4.2 importiert haben.
- f. Ziehen Sie die Tabellendefinition TUTORIAL in das leere Teilfenster rechts neben dem Repository des Ordners.
- g. Klicken Sie in der Tabellendefinition TUTORIAL auf **Alles auswählen** und ziehen Sie dann die Spaltendefinitionen in das Teilfenster **Einfügespalten**.
Ihr Teilfenster **Einfügespalten** enthält die Spalten, die Sie in die Zieldatenbank schreiben wollen.

INSERT-Spalten		
	Einfügespalte	Einfügewert
1	CUSTOMER_NUMBER	ORCHESTRATE.CUSTOMER_NUMBER
2	CUST_NAME	ORCHESTRATE.CUST_NAME
3	ADDR_1	ORCHESTRATE.ADDR_1
4	ADDR_2	ORCHESTRATE.ADDR_2
5	CITY	ORCHESTRATE.CITY
6	REGION_CODE	ORCHESTRATE.REGION_CODE
7	ZIP	ORCHESTRATE.ZIP
8	TEL_NUM	ORCHESTRATE.TEL_NUM
9	SOURCE	ORCHESTRATE.SOURCE
10	RECNUM	ORCHESTRATE.RECNUM
11	SETUP_DATE	ORCHESTRATE.SETUP_DATE
12	SPECIAL_HANDLING_CODE	ORCHESTRATE.SPECIAL_HANDLING_CODE
*

INSERT SQL

Abbildung 13. Sie haben die Spalten angegeben, die in die Zieldatenbank geschrieben werden sollen.

- h. Klicken Sie auf die Registerkarte **SQL**, um die erstellte SQL-Anweisung anzuzeigen. Diese Anweisung weist die Syntax auf, mit der die von Ihnen importierten Spaltendefinitionen eingefügt werden.
 - i. Klicken Sie auf **OK**, um das Fenster **SQL Builder** zu schließen.
Die SQL-Anweisung wurde im Feld **Einfügeanweisung** eingefügt.
4. Klicken Sie auf **OK**, um die Stage GlobalCo_BillTo_Target zu schließen. Speichern und kompilieren Sie anschließend den Job und führen Sie ihn aus.

Sie haben die bereinigten Rechnungsstellungsdaten der Kunden von GlobalCo in die Datenbanktabelle des Lernprogramms geschrieben.

Weitere Schritte

In dieser Lerneinheit haben Sie mithilfe einer Connector-Stage eine Verbindung zur Datenbank von GlobalCo Worldwide hergestellt. Sie haben die Rechnungsstellungsdaten der Kunden von GlobalCo in diese Datenbank geschrieben, damit an-

dere Entwickler auf die Daten zugreifen und sie mit den Rechnungsstellungsdaten der Kunden von WorldCo kombinieren können.

Sie haben die folgenden Tasks ausgeführt:

- Konfigurieren einer Connector-Stage
- Bereitstellen von Datenbankverbindungsdetails über ein Datenverbindungsobjekt
- Verwenden von SQL Builder zum Definieren der SQL-Anweisung für das Schreiben von Daten in die Datenbank
- Schreiben der bereinigten Rechnungsstellungsdaten der Kunden von GlobalCo in die Datenbank

Sie haben das Lernprogramm für parallele Jobs für InfoSphere DataStage abgeschlossen.

Sie können jetzt eine der folgenden Tasks ausführen:

- Arbeiten Sie die Module von *InfoSphere DataStage Advanced Tutorial* durch.
- Lesen Sie das *IBM InfoSphere DataStage und QualityStage Designer-Client - Handbuch*, um mehr über InfoSphere DataStage zu erfahren.

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 Produkte 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 3. 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. 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 Produkte 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 E. 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 4. 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 F. 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 G. Feedback zu 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 5. 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

Tabelle 5. 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

A

Administrator-Client 3

B

Beispieljob
 Überblick 13
Bemerkungen 77

D

Data Set, Stage
 untersuchen 15
Daten
 kombinieren 41
 Lookup erstellen 41
 zurückgewiesene Daten 45
Daten kombinieren
 Lookup, Stage 41
Daten referenzieren
 Lookup File Set, Stage 41
Daten umsetzen
 Transformer, Stage 35
Daten zurückweisen
 erfassen 45
Datenbank
 Zieltabelle erstellen 5
Datenbankconnector
 konfigurieren 58
Datenquellenname
 Datei dsenv konfigurieren 7
 Datei odbc.ini konfigurieren 8
 Datei uvodbc.config konfigurieren 9
 DB2-Paket erstellen 10
 erstellen 6
 für die Datenbank erstellen 6
 ODBC-Konnektivität testen 10
 ODBC-Treiber konfigurieren 10
Datenverbindungsobjekte
 erstellen 55
Designer-Client
 Job öffnen 13
 starten 13
Director-Client
 starten 17
dsenv, Datei
 konfigurieren 7

J

Jobentwurf
 Stages hinzufügen 23
Jobparameter
 Ihren Jobs hinzufügen 29
 Parametersätze 31
Jobprotokoll
 Daten anzeigen 17
Jobs
 ausführen 17

Jobs (*Forts.*)
 erstellen 21
 kompilieren 15
 öffnen 13
Jobs entwerfen
 Lernprogramm 21

K

Kundenunterstützung
 kontaktieren 65, 71

L

Lernprogramm für parallele Jobs
 Beispieljob 13
 Dateien herunterladen 3
 Einrichtung 3
 Installation 3
 Job entwerfen 21
 Job erstellen 21
 Komponenten importieren 4
 Quellendateien kopieren 4
Lernprogrammdateien
 herunterladen 3
Lernprogramme
 parallele Jobs erstellen 1
Lernprogrammkomponenten
 importieren 4
Lookup, Stage
 Daten kombinieren 41
Lookup File Set, Stage
 Daten referenzieren 41

M

Marken
 Liste 77
Metadaten
 importieren 26

O

ODBC
 andere ODBC-Treiber konfigurieren 10
 Datei dsenv konfigurieren 7
 Datei odbc.ini konfigurieren 8
 Datei uvodbc.config konfigurieren 9
 DB2-Paket erstellen 10
 Konnektivität testen 10
odbc.ini, Datei
 konfigurieren 8
ODBC-Verbindung
 Datenquellennamen erstellen 6
ODBC-Verbindungen
 Datenquellennamen erstellen
 AIX, HP-UX, Solaris, Linux 6
 Windows 6

P

Parallele Jobs
 Lernprogramm 1
Parametersätze
 erstellen 31
Produktdokumentation
 Zugriff auf 67, 73
Produkteingabehilfen
 Eingabehilfen 63, 69
Projekt
 erstellen 3

Q

Quellendateien
 auf Computer kopieren 4

R

Repository-Objekte
 Tabellendefinitionen importieren 57

S

Schreiben, in eine Datenbank
 mit Datenbankconnector 58
Sequential File, Stage
 untersuchen 15
Software-Services
 kontaktieren 65, 71
Spaltenmetadaten
 importieren 26, 57
Stages
 Ihren Job hinzufügen 23
 Lookup File Set 23
 Stage-Eigenschaften 23
 Stages untersuchen 15

T

Tabellendefinitionen
 importieren 26
Transformer, Stage
 Daten umsetzen 35

U

Unterstützung
 Kundenunterstützung 65, 71
uvodbc.config, Datei
 konfigurieren 9

Z

Zieldatenbanktabelle
 erstellen 5



SC43-1104-00

