Getting Started

Informatica PowerCenter®
(Version 8.1)

Informatica PowerCenter Getting Started
Version 8.1
April 2006

Table of Contents

List of Figures

List of Tables

Preface

Welcome to PowerCenter, the Informatica software product that delivers an open, scalable data integration solution addressing the complete life cycle for all data integration projects including data warehouses, data migration, data synchronization, and information hubs. PowerCenter combines the latest technology enhancements for reliably managing data repositories and delivering information resources in a timely, usable, and efficient manner.

The PowerCenter repository coordinates and drives a variety of core functions, including extracting, transforming, loading, and managing data. The Integration Service can extract large volumes of data from multiple platforms, handle complex transformations on the data, and support high-speed loads. PowerCenter can simplify and accelerate the process of building a comprehensive data warehouse from disparate data sources.

About This Book

Getting Started is written for the IS developers and software engineers who are responsible for implementing a data warehouse. It provides a tutorial to help first-time users learn how to use PowerCenter. *Getting Started* assumes you have knowledge of your operating systems, relational database concepts, and the database engines, flat files, or mainframe systems in your environment. The guide also assumes you are familiar with the interface requirements for your supporting applications.

The material in this book is available for online use.

Document Conventions

This guide uses the following formatting conventions:

If you see...	It means...
italicized text	The word or set of words are especially emphasized.
boldfaced text	Emphasized subjects.
`italicized monospaced text`	This is the variable name for a value you enter as part of an operating system command. This is generic text that should be replaced with user-supplied values.
Note:	The following paragraph provides additional facts.
Tip:	The following paragraph provides suggested uses.
Warning:	The following paragraph notes situations where you can overwrite or corrupt data, unless you follow the specified procedure.
`monospaced text`	This is a code example.
`bold monospaced text`	This is an operating system command you enter from a prompt to run a task.

Other Informatica Resources

In addition to the product manuals, Informatica provides these other resources:

- Informatica Customer Portal
- Informatica web site
- Informatica Developer Network
- Informatica Knowledge Base
- Informatica Technical Support

Visiting Informatica Customer Portal

As an Informatica customer, you can access the Informatica Customer Portal site at http://my.informatica.com. The site contains product information, user group information, newsletters, access to the Informatica customer support case management system (ATLAS), the Informatica Knowledge Base, and access to the Informatica user community.

Visiting the Informatica Web Site

You can access the Informatica corporate web site at http://www.informatica.com. The site contains information about Informatica, its background, upcoming events, and sales offices. You will also find product and partner information. The services area of the site includes important information about technical support, training and education, and implementation services.

Visiting the Informatica Developer Network

You can access the Informatica Developer Network at http://devnet.informatica.com. The Informatica Developer Network is a web-based forum for third-party software developers. The site contains information about how to create, market, and support customer-oriented add-on solutions based on interoperability interfaces for Informatica products.

Visiting the Informatica Knowledge Base

As an Informatica customer, you can access the Informatica Knowledge Base at http://my.informatica.com. Use the Knowledge Base to search for documented solutions to known technical issues about Informatica products. You can also find answers to frequently asked questions, technical white papers, and technical tips.

Obtaining Technical Support

There are many ways to access Informatica Technical Support. You can contact a Technical Support Center by using the telephone numbers listed the following table, you can send email, or you can use the WebSupport Service.

Use the following email addresses to contact Informatica Technical Support:

- support@informatica.com for technical inquiries
- support_admin@informatica.com for general customer service requests

WebSupport requires a user name and password. You can request a user name and password at http://my.informatica.com.

North America / South America	Europe / Middle East / Africa	Asia / Australia
Informatica Corporation Headquarters 100 Cardinal Way Redwood City, California 94063 United States	**Informatica Software Ltd**. 6 Waltham Park Waltham Road, White Waltham Maidenhead, Berkshire SL6 3TN United Kingdom	**Informatica Business Solutions Pvt. Ltd.** Diamond District Tower B, 3rd Floor 150 Airport Road Bangalore 560 008 India
Toll Free 877 463 2435	**Toll Free** 00 800 4632 4357	**Toll Free** Australia: 00 11 800 4632 4357 Singapore: 001 800 4632 4357
Standard Rate United States: 650 385 5800	**Standard Rate** Belgium: +32 15 281 702 France: +33 1 41 38 92 26 Germany: +49 1805 702 702 Netherlands: +31 306 022 797 United Kingdom: +44 1628 511 445	**Standard Rate** India: +91 80 4112 5738

Chapter 1

Before You Begin

This chapter includes the following topics:

Overview

Getting Started includes multiple lessons that introduce you to PowerCenter and how to use it to load transformed data into file and relational targets. The lessons in this book are designed for PowerCenter beginners.

This tutorial walks you through the process of creating a data warehouse. The tutorial teaches you how to:

- Create users and groups.
- Add source definitions to the repository.
- Create targets and add their definitions to the repository.
- Map data between sources and targets.
- Instruct the Integration Service to write data to targets.
- Monitor the Integration Service as it writes data to targets.

In general, you can set the pace for completing the tutorial. However, you should complete an entire lesson in one session, since each lesson builds on a sequence of related tasks.

For additional information, case studies, and updates on using Informatica products, see the Informatica Knowledge Base at http://my.informatica.com.

Getting Started

Before you can begin the lessons, read "Product Overview" in the *Administrator Guide*. The product overview explains the different components that work together to extract, transform, and load data.

Also, the PowerCenter administrator must install and configure the PowerCenter Client and Services. Verify that the administrator has completed the following steps:

- **Install PowerCenter Client.** You use the PowerCenter Client to manage users, define sources and targets, build mappings and mapplets with the transformation logic, and create sessions and workflows to run the mapping logic.
- **Install and configure PowerCenter Services.** The Repository Service manages connections to the repository from client applications. It inserts, updates, and fetches objects from the repository database tables.

  The Integration Service extracts the source data, performs the data transformation, and loads the transformed data into the targets.
- **Create a repository.** The PowerCenter repository is at the center of the PowerCenter suite. When you create objects with the Informatica applications, you create a set of metadata tables within the repository database that PowerCenter accesses. The PowerCenter Client and Integration Service access the repository to save and retrieve metadata.

For an overview on installing and configuring all the PowerCenter components, see "Installation and Configuration Overview" in the *Installation and Configuration Guide*.

You also need information to connect to the source, target, and repository databases. Use the tables in "Connecting to Databases" on page 4 to enter the connectivity information. Contact the PowerCenter administrator for the necessary information.

Using the PowerCenter Client in the Tutorial

The PowerCenter Client consists of applications that you use to design mappings and mapplets, create sessions and workflows to load the data, and monitor workflow progress.

In this tutorial, you use the following applications and tools:

- **Repository Manager.** Use the Repository Manager to create and administer the metadata repository. In Lesson 1, you perform some tasks that you need to complete before you start the other lessons in the book. You use the Repository Manager to create a repository user and group. You create a folder to store the metadata you create in the lessons.
- **Designer.** Use the Designer to create mappings that contain transformation instructions for the Integration Service. Before you can create mappings, you must add source and target definitions to the repository. In this tutorial, you use the following tools in the Designer:
  - **Source Analyzer.** Import or create source definitions. In Lesson 2, you use the Designer to create source definitions.
  - **Target Designer.** Import or create target definitions. In Lesson 2, you use the Designer to create target definitions. You also create tables in the target database based on the target definitions.
  - **Mapping Designer.** Create mappings that the Integration Service uses to extract, transform, and load data. In Lessons 3 to 5, you use the Mapping Designer to create mappings that extract data from sources and load it to a target.
- **Workflow Manager.** Use the Workflow Manager to create and run workflows and tasks. A workflow is a set of instructions that describes how and when to run tasks related to extracting, transforming, and loading data. In Lessons 3 to 5, you use the Workflow Manager to create Session tasks to run the mappings you create, and workflows to run the sessions.
- **Workflow Monitor.** Use the Workflow Monitor to monitor scheduled and running workflows for each Integration Service. In Lessons 3 to 5, you run the workflows you create, and use the Workflow Monitor to monitor the workflow runs.

For more information about PowerCenter, see "Product Overview" in the *Administrator Guide*.

Connecting to Databases

To use the lessons in this book, you need to connect to the source, target, and repository databases. Use the tables in this section to record the connectivity information you need to connect to the databases. Contact the PowerCenter administrator for the connection information listed below.

Use Table 1-1 to enter the domain information you need to connect to the repository:

Table 1-1. Repository Domain Information

	Domain
Domain Name	
Gateway Host	
Gateway Port	

Use Table 1-2 to enter the information you need to connect to the repository as the PowerCenter administrator:

Table 1-2. Administrator Repository Login

	Repository
Repository Name	
Repository Administrator User Name	
Repository Administrator Password	

Note: Use the Administrator profile for the lessons "Creating Repository Users and Groups" on page 8 and "Creating a Folder" on page 14. For all other lessons, you use a user profile that you create to log in to the repository. If you do not have the repository administrator user name and password, ask a repository administrator to either provide this information or set up a repository user profile that you can use.

Use Table 1-3 to enter the information you need to connect to the repository in each PowerCenter Client tool. Use this information when you create the user profile in "Creating a User" on page 12:

Table 1-3. Repository Login

	Repository
Repository Name	
Repository User Name	
Repository Password	

You need to create an ODBC connection for the source and target databases, if not already created. For more information about database connections, see "Connecting to Databases from Windows" in the *Installation and Configuration Guide*.

Use Table 1-4 to enter the information you need to connect to the source and target databases from the Designer:

Table 1-4. ODBC Data Source Information

	Source Connection	Target Connection
ODBC Data Source Name		
Database User Name		
Database Password		

For more information about ODBC drivers, see "Connecting to Databases from Windows" in the *Installation and Configuration Guide*.

Use Table 1-5 to enter the information you need to create database connections in the Workflow Manager:

Table 1-5. Workflow Manager Connectivity Information

	Source Database Connection Object	Target Database Connection Object
Database Type		
User Name		
Password		
Connect String		
Code Page		
Database Name		
Server Name		
Domain Name		
Note: You may not need all properties in this table.		

Table 1-6 lists the native connect string syntax to use for different databases:

Table 1-6. Native Connect String Syntax for Database Platforms

Database	Native Connect String	Example
IBM DB2	*dbname*	mydatabase
Informix	*dbname@servername*	mydatabase@informix
Microsoft SQL Server	*servername@dbname*	sqlserver@mydatabase
Oracle	*dbname*.world (same as TNSNAMES entry)	oracle.world

Table 1-6. Native Connect String Syntax for Database Platforms

Database	Native Connect String	Example
Sybase ASE	*servername*@*dbname*	sambrown@mydatabase
Teradata	Teradata* ODBC_data_source_name or ODBC_data_source_name@db_name or ODBC_data_source_name@db_user_name	TeradataODBC TeradataODBC@mydatabase TeradataODBC@sambrown

Chapter 2

Tutorial Lesson 1

This chapter includes the following topics:

Creating Repository Users and Groups

You can create a repository user profile for each person working in the repository. Each profile has a separate user name and password. You can also create user groups and assign each user to one or more groups. Then, grant repository privileges to each group, so users in the group can perform repository tasks, such as using the Designer or creating workflows.

The repository user profile is not the same as the database user profile. Although a particular user might not have access to a database as a *database* user, that same person can have privileges to a repository in the database as a *repository* user.

PowerCenter provides the following levels of repository security:

- **Privileges.** Repository-wide security that controls which task or set of tasks a single user or group of users can access.
- **Permissions.** Security assigned to individual folders or global objects in the repository.

Privileges depend on a user's group membership. Every repository user belongs to at least one group. For example, the user who administers the repository belongs to the Administrators group. By default, users receive the privileges assigned to the group. While it is most common to assign privileges by group, the repository administrator, who has either the Super User *or* Admin Repository privilege, can also grant privileges to individual users.

A PowerCenter administrator can complete the following tasks:

- Create groups.
- Assign privileges to groups.
- Create users and assign them to groups.

In this lesson, you complete the following tasks:

1. Connect to the repository as the administrator user. If necessary, ask the PowerCenter administrator for the user name and password. Otherwise, ask the PowerCenter administrator to complete the lessons in this chapter for you.
2. Create the TUTORIAL group. To do this, you need to log in to the repository as the Administrator.
3. Assign privileges to the TUTORIAL group.
4. Create a new user.

Connecting to the Repository

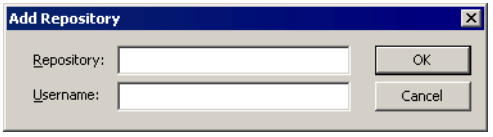
To complete the following tasks, you need to connect to the repository. If you are already connected to the repository, disconnect and connect again to log in as the Administrator.

Note: Use the Administrator user name and password you entered in Table 1-2 on page 4. Otherwise, ask the PowerCenter administrator to perform the tasks in this chapter for you.

To connect to the repository:

1. Launch the Repository Manager.
2. Click Repository > Add Repository.

   The Add Repository dialog box appears.

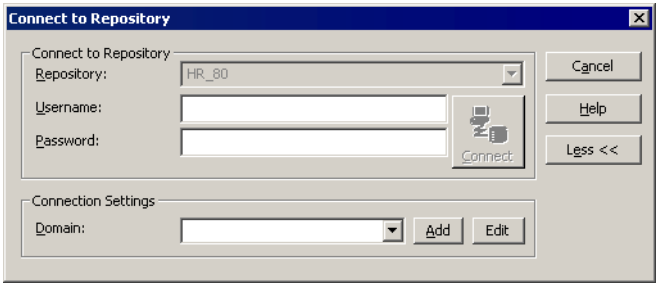


3. Enter the name of the repository and the user name for the Administrator user.

   Use the information you entered in Table 1-2 on page 4.
4. Click OK.

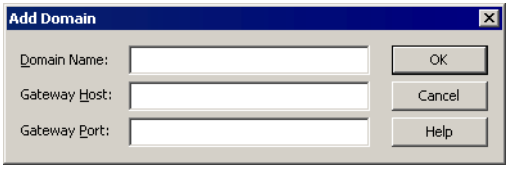
   The repository appears in the Navigator.
5. Click Repository > Connect, or double-click the repository to connect.

   The Connect to Repository dialog box appears.



6. In the connection settings, click Add to add domain connection information.

   The Add Domain dialog box appears.



7. Enter the domain name, gateway host, and gateway port number from Table 1-1 on page 4.
8. Click OK.

If a message indicates that the domain already exists, click Yes to replace the existing domain.

9. In the Connect to Repository dialog box, enter the password for the Administrator user.
10. Click Connect.

    You are now connected to the repository as the Administrator user.

Creating a Group

In the following steps, you create a new group.

To create the TUTORIAL group:

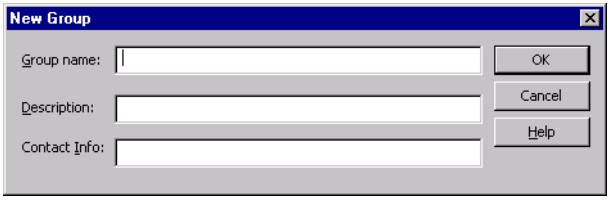
1. Select the repository in the Navigator, and click Security > Manage Users and Privileges.

   The Manage Users and Privileges dialog box contains tabs for users, groups, and privileges that list all existing users, groups, and privileges in the repository.
2. Click the Groups tab.

   The Groups tab includes the default groups, Administrators and Public. You cannot edit or remove these groups.

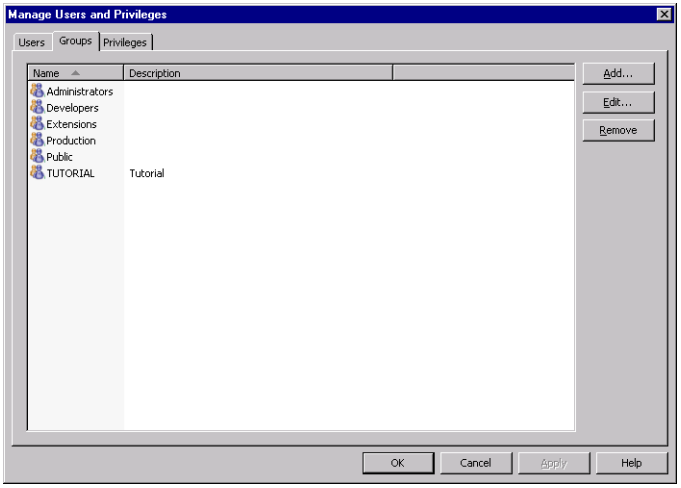
   Note: The Groups tab may display more groups if the administrator created additional groups.
3. Click Add.

   The New Group dialog box appears.



4. Type TUTORIAL for the name of the new group and Tutorial as the description. Click OK.

The TUTORIAL group appears on the Groups tab.



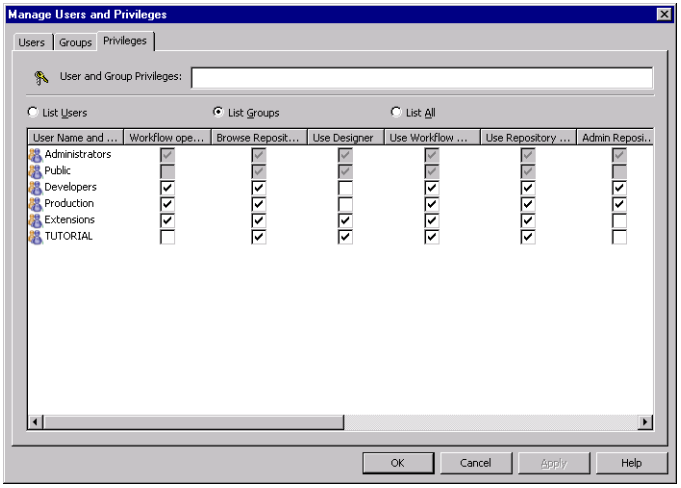
Assigning Privileges to a Group

The next step is to grant privileges to the TUTORIAL group.

To assign privileges to the TUTORIAL group:

1. In the Manage Users and Privileges dialog box, click the Privileges tab.

   The privileges currently assigned to users and groups appears.



   The TUTORIAL group has the following default privileges:

   - Browse Repository
   - Use Designer
   - Use Workflow Manager

- Use Repository Manager
- Manage Connection

Users in the TUTORIAL group now can create workflows in any folder for which they have read and write permission.

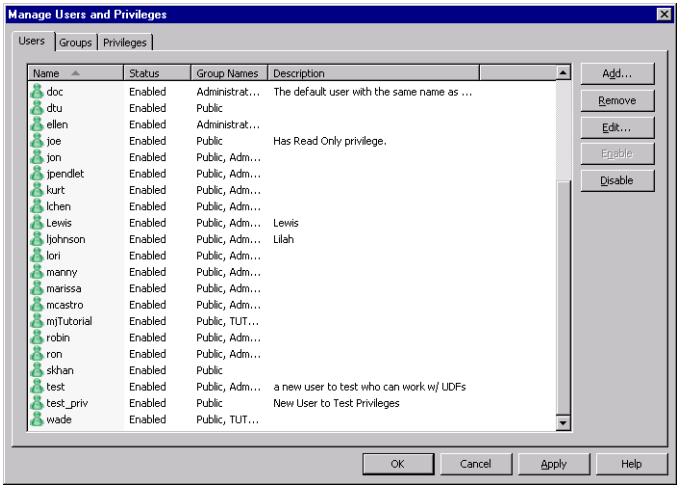
Creating a User

The final step is to create a new user and add that user to the TUTORIAL group. You use this profile throughout the rest of this tutorial.

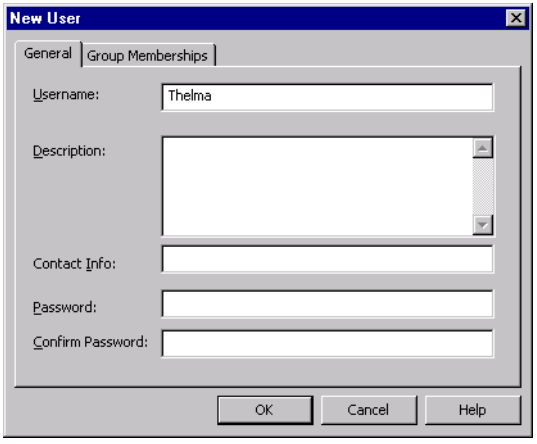
To create a new user:

1. In the Manage Users and Privileges dialog box, click the Users tab.

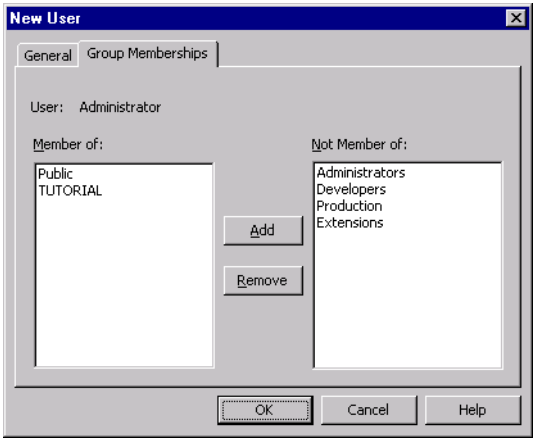
   The dialog box lists all the users in the repository.



2. Click Add.



3. In the New User dialog box, enter a name as the user name.
4. Enter contact information, such as a phone number, if needed.
5. In both the Password and Confirm Password fields, enter a password.
6. Click the Group Memberships tab.
7. Select TUTORIAL and click Add.



8. Click OK.

   You now have all the privileges associated with the TUTORIAL group.
9. Click OK to return to the Repository Manager.

Creating a Folder

In this section, you create a tutorial repository folder. You save all objects you create in these lessons in the tutorial folder.

Folders provide a way to organize and store all metadata in the repository, including mappings, schemas, and sessions. Folders are designed to be flexible to help you organize the repository logically. Each folder has a set of properties you can configure to define how users access the folder. For example, you can create a folder that allows all repository users to see objects within the folder, but not to edit them.

Folder Permissions

Permissions allow repository users to perform tasks within a folder. With folder permissions, you can control users' access to the folder and the tasks you permit them to perform.

Folder permissions work closely with repository privileges. Privileges grant access to specific tasks, while permissions grant access to specific folders with read, write, and execute access. However, any user with the Super User privilege can perform all tasks across all folders in the repository. Folders have the following types of permissions:

- **Read permission.** You can view the folder and objects in the folder.
- **Write permission.** You can create or edit objects in the folder.
- **Execute permission.** You can run or schedule workflows in the folder.

Creating Folders in This Tutorial

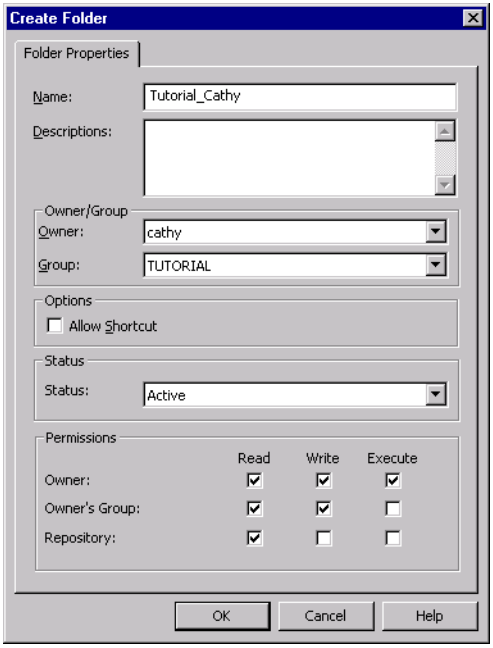
For this tutorial, you create a single folder, Tutorial_*yourname*. You add metadata to this folder about orders, customers, and products.

To create a new folder:

1. In the Repository Manager, click Folder > Create.
2. Enter Tutorial_*yourname* as the name of the folder.
3. Select a user name as the owner of the folder.

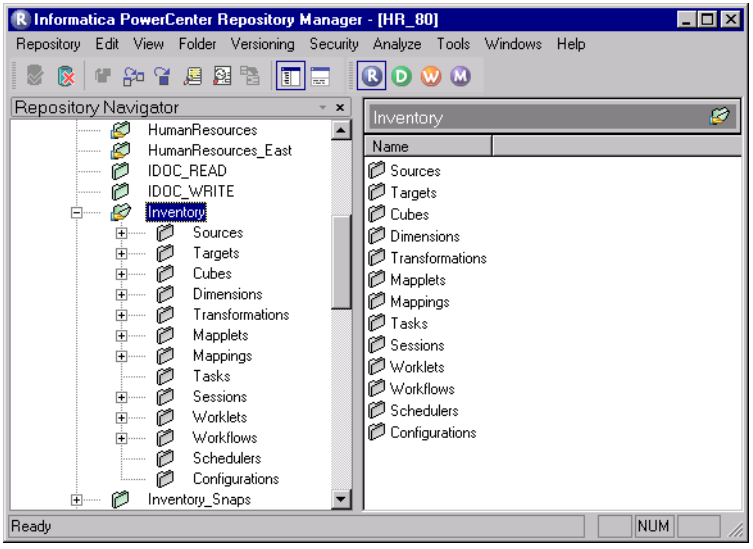
   Because the user name is a member of the Public group, that group appears by default in the Group field. Users who belong to more than one group must select a group from the list.

4. Select TUTORIAL as the owner group for the folder.



5. Click OK to create the folder.

   The new folder now appears as part of the repository.



6. Exit the Repository Manager.

Creating Source Tables

With most data warehouses, you already have existing source tables or flat files. Before you continue with the other lessons in this book, you need to create the source tables in the database. In this section, you run an SQL script in the Target Designer to create sample source tables. The SQL script creates sources with 7-bit ASCII table names and data.

When you run the SQL script, you create the following source tables:

- CUSTOMERS
- DEPARTMENT
- DISTRIBUTORS
- EMPLOYEES
- ITEMS
- ITEMS_IN_PROMOTIONS
- JOBS
- MANUFACTURERS
- ORDERS
- ORDER_ITEMS
- PROMOTIONS
- STORES

The Target Designer generates SQL based on the definitions in the workspace. Generally, you use the Target Designer to create target tables in the target database. In this lesson, you use this feature to generate the source tutorial tables from the tutorial SQL scripts that ship with the product.

When you run the SQL script, you also create a stored procedure that you will use to create a Stored Procedure transformation in Lesson 5. For more information about the stored procedure, see "Creating a Stored Procedure Transformation" on page 85.

To create the sample source tables:

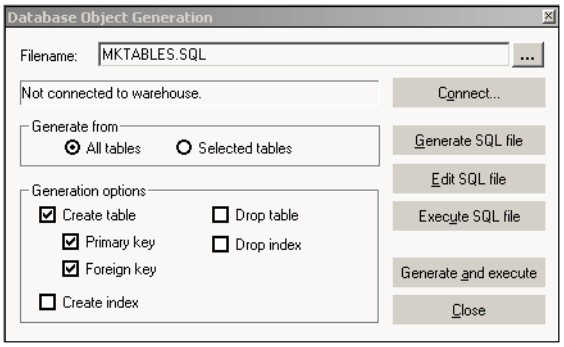
1. Launch the Designer, double-click the icon for the repository, and log into the repository.

   Use your user profile to open the connection.

2. Double-click the Tutorial_*yourname* folder.

3. Click Tools > Target Designer to switch to the Target Designer.

4. Click Targets > Generate/Execute SQL.



The Database Object Generation dialog box gives you several options for creating tables.

5. Click the Connect button to connect to the source database.
6. Select the ODBC data source you created to connect to the source database. Use the information you entered in Table 1-4 on page 5.
7. Enter the database user name and password and click the Connect button.

   You now have an open connection to the source database. When you are connected, the Disconnect button appears and the ODBC name of the source database appears in the dialog box.
8. Make sure the Output window is open at the bottom of the Designer.

   If it is not open, click View > Output.
9. Click the browse button to find the SQL file. The SQL file is installed in the following directory:

   ```
   C:\Program Files\Informatica PowerCenter\client\bin
   ```
10. Select the SQL file appropriate to the source database platform you are using. Click Open.

Platform	File
Informix	smpl_inf.sql
Microsoft SQL Server	smpl_ms.sql
Oracle	smpl_ora.sql
Sybase ASE	smpl_syb.sql
DB2	smpl_db2.sql
Teradata	smpl_tera.sql

Alternatively, you can enter the path and file name of the SQL file.

11. Click Execute SQL file.

    The database now executes the SQL script to create the sample source database objects and to insert values into the source tables. While the script is running, the Output window displays the progress. The Designer generates and executes SQL scripts in Unicode (UCS2) format.

12. When the script completes, click Disconnect, and then click Close.

What Comes Next

In the next lesson, you create source definitions in the repository based on the source tables you just created. You also create target definitions and create tables in the target database based on the target definitions.

Chapter 3

Tutorial Lesson 2

This chapter includes the following topics:

Creating Source Definitions

Now that you have added the source tables containing sample data, you are ready to create the source definitions in the repository. The repository contains a description of source tables, not the actual data contained in them. After you add these source definitions to the repository, you use them in a mapping.

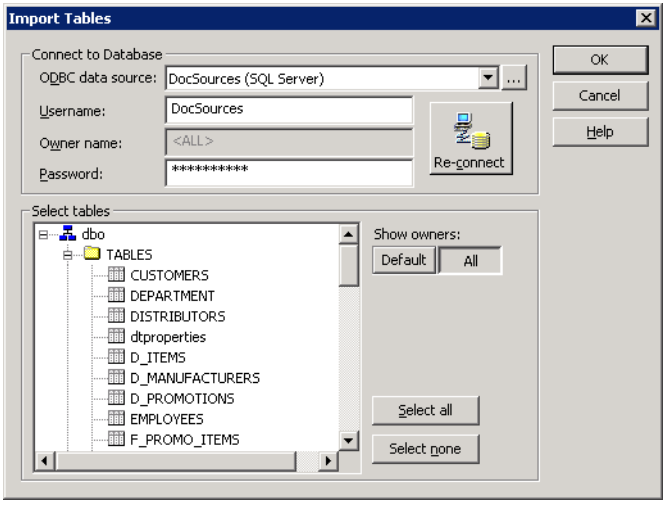
To import the sample source definitions:

1. In the Designer, click Tools > Source Analyzer to open the Source Analyzer.
2. Double-click the tutorial folder to view its contents.

   Every folder contains nodes for sources, targets, schemas, mappings, mapplets, cubes, dimensions and reusable transformations.
3. Click Sources > Import from Database.
4. Select the ODBC data source to access the database containing the source tables.
5. Enter the user name and password to connect to this database. Also, enter the name of the source table owner, if necessary.

   Use the database connection information you entered in Table 1-4 on page 5.

   In Oracle, the owner name is the same as the user name. Make sure that the owner name is in all caps (for example, JDOE).



6. Click Connect.
7. In the Select tables list, expand the database owner and the TABLES heading.

   If you click the All button, you can see all tables in the source database.

   You should now see a list of all the tables you created by running the SQL script in addition to any tables already in the database.

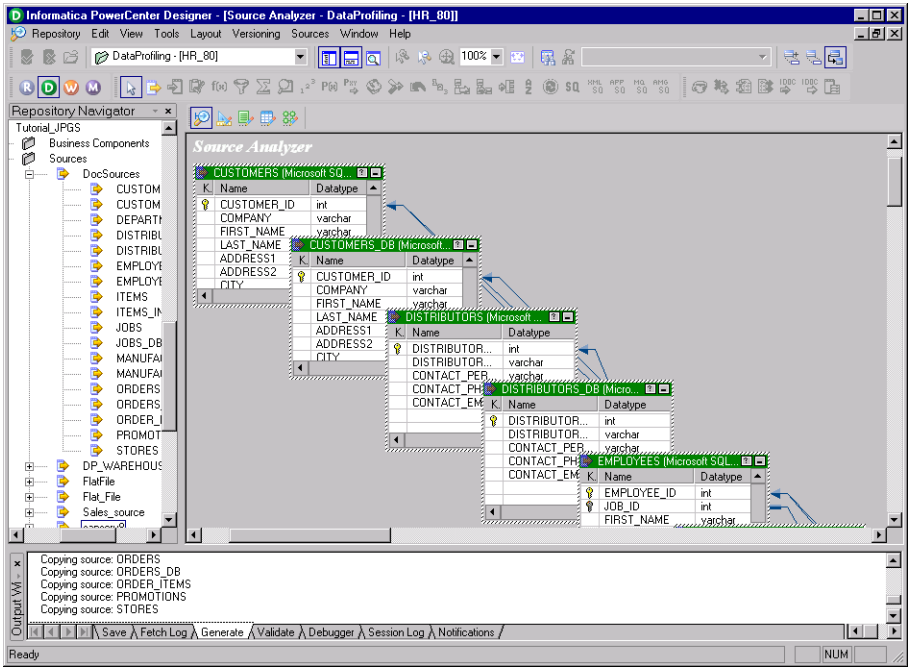
8. Select the following tables:
   - CUSTOMERS
   - DEPARTMENT
   - DISTRIBUTORS
   - EMPLOYEES
   - ITEMS
   - ITEMS_IN_PROMOTIONS
   - JOBS
   - MANUFACTURERS
   - ORDERS
   - ORDER_ITEMS
   - PROMOTIONS
   - STORES

   Hold down the Ctrl key to select multiple tables. Or, hold down the Shift key to select a block of tables. You may need to scroll down the list of tables to select all tables.

   Note: Database objects created in Informix databases have shorter names than those created in other types of databases. For example, the name of the table ITEMS_IN_PROMOTIONS is shortened to ITEMS_IN_PROMO.

9. Click OK to import the source definitions into the repository.

The Designer displays the newly imported sources in the workspace. You can click Layout > Scale to Fit to fit all the definitions in the workspace.



A new database definition (DBD) node appears under the Sources node in the tutorial folder. This new entry has the same name as the ODBC data source to access the sources you just imported. If you double-click the DBD node, the list of all the imported sources appears.

Viewing Source Definitions

You can view details for each source definition.

To view a source definition:

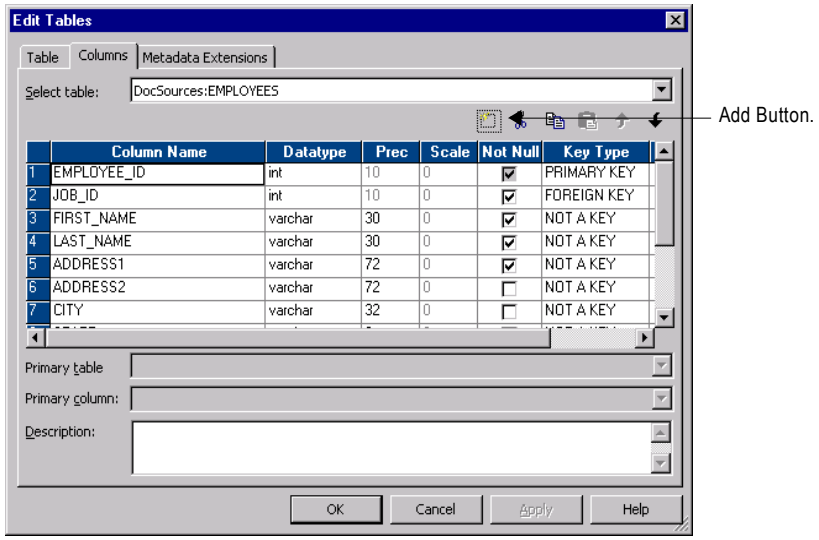
1. Double-click the title bar of the source definition for the EMPLOYEES table to open the EMPLOYEES source definition.

   The Edit Tables dialog box appears and displays all the properties of this source definition. The Table tab shows the name of the table, business name, owner name, and the database type. You can add a comment in the Description section.

2. Click the Columns tab.

   The Columns tab displays the column descriptions for the source table.

Note: The source definition must match the structure of the source table. Therefore, you must not modify source column definitions after you import them.



3. Click the Metadata Extensions tab.

   Metadata extensions allow you to extend the metadata stored in the repository by associating information with individual repository objects. For example, you can store contact information, such as name or email address, with the sources you create.

   In this lesson, you create user-defined metadata extensions that define the date you created the source definition and the name of the person who created the source definition.

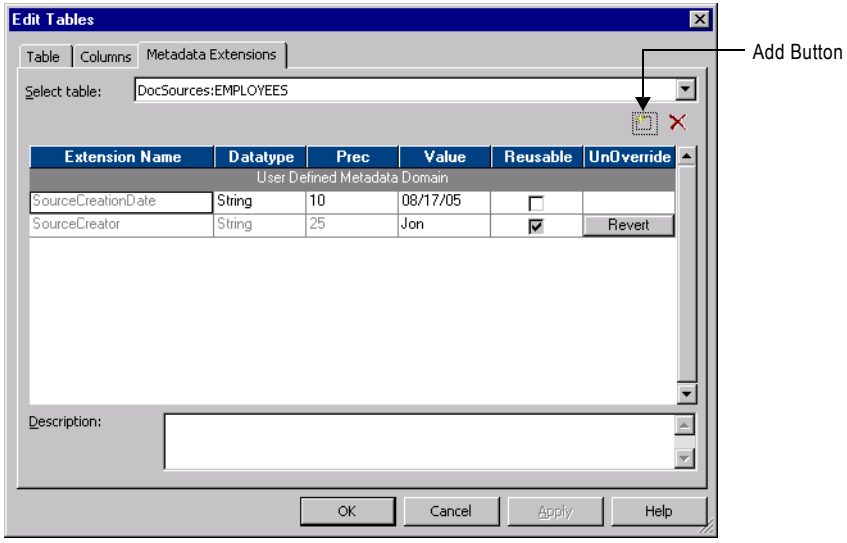
4. Click the Add button to add a new metadata extensions.

   The SourceCreator row is added by default.

5. Name the new row SourceCreationDate and enter today's date as the value.

6. Enter your first name as the value in the SourceCreator row.

   The Metadata Extensions tab should look similar to this:



7. Click Apply.
8. Click OK to close the dialog box.
9. Click Repository > Save to save the changes to the repository.

Creating Target Definitions and Target Tables

You can import target definitions from existing target tables, or you can create the definitions and then generate and run the SQL to create the target tables. In this lesson, you create a target definition in the Target Designer, and then create a target table based on the definition.

Creating Target Definitions

The next step is to create the metadata for the target tables in the repository. The actual tables that the target definitions describe do not exist yet.

Target definitions define the structure of tables in the target database, or the structure of file targets the Integration Service creates when you run a session. If you add a target definition to the repository that does not exist in the relational database, you need to create target tables in the target database. You do this by generating and executing the necessary SQL code within the Target Designer.

In the following steps, you copy the EMPLOYEES source definition into the Target Designer to create the target definition. Then, you modify the target definition by deleting and adding columns to create the definition you want.

To create the T_EMPLOYEES target definition:

1. In the Designer, switch to the Target Designer.
2. Drag the EMPLOYEES source definition from the Navigator to the Target Designer workspace.

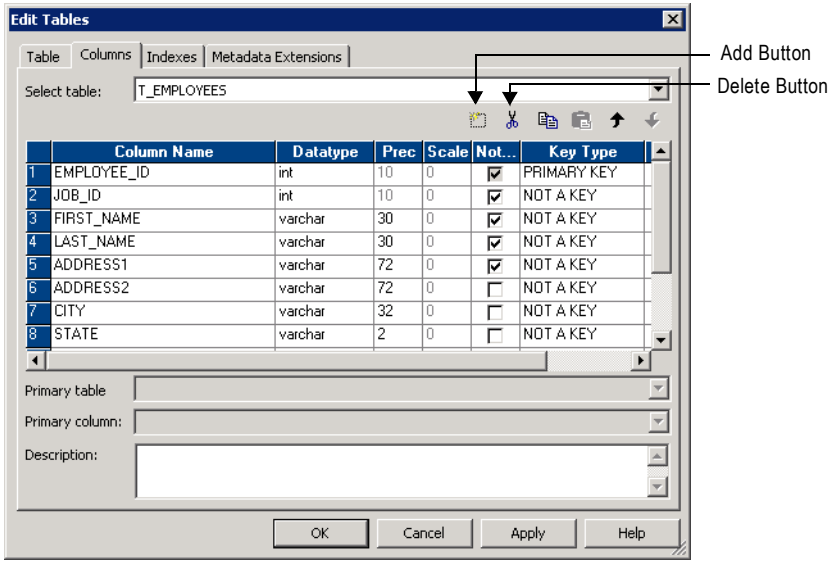
   The Designer creates a new target definition, EMPLOYEES, with the same column definitions as the EMPLOYEES source definition and the same database type.

   Next, modify the target column definitions.
3. Double-click the EMPLOYEES target definition to open it.

   The Edit Tables dialog box appears.
4. Click Rename and name the target definition T_EMPLOYEES.

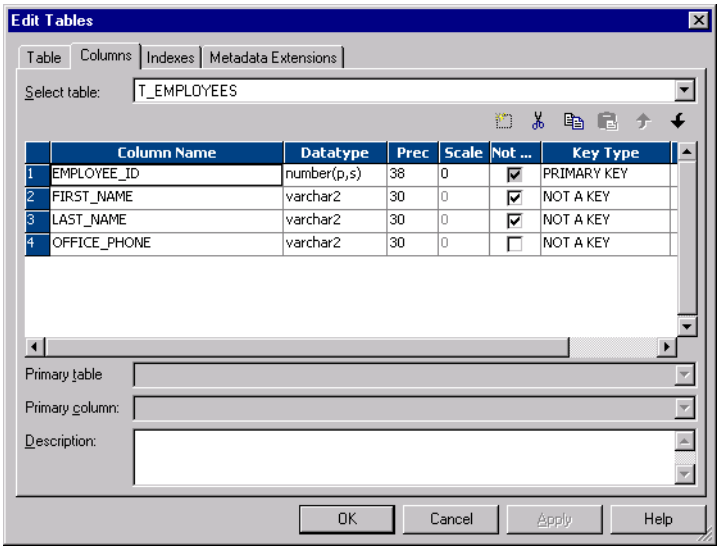
   Note: If you need to change the database type for the target definition, you can select the correct database type when you edit the target definition.
5. Click the Columns tab.

The target column definitions are the same as the EMPLOYEES source definition.



6. Select the JOB_ID column and click the delete button.
7. Delete the following columns:
   - ADDRESS1
   - ADDRESS2
   - CITY
   - STATE
   - POSTAL_CODE
   - HOME_PHONE
   - EMAIL

When you finish, the target definition should look similar to the following target definition:



Note that the EMPLOYEE_ID column is a primary key. The primary key cannot accept null values. The Designer selects Not Null and disables the Not Null option. You now have a column ready to receive data from the EMPLOYEE_ID column in the EMPLOYEES source table.

Note: If you want to add a business name for any column, scroll to the right and enter it.

8. Click OK to save the changes and close the dialog box.
9. Click Repository > Save.

Creating Target Tables

Use the Target Designer to run an existing SQL script to create target tables.

Note: When you use the Target Designer to generate SQL, you can choose to drop the table in the database before creating it. To do this, select the Drop Table option. If the target database already contains tables, make sure it does not contain a table with the same name as the table you plan to create. If the table exists in the database, you lose the existing table and data.

To create the target T_EMPLOYEES table:

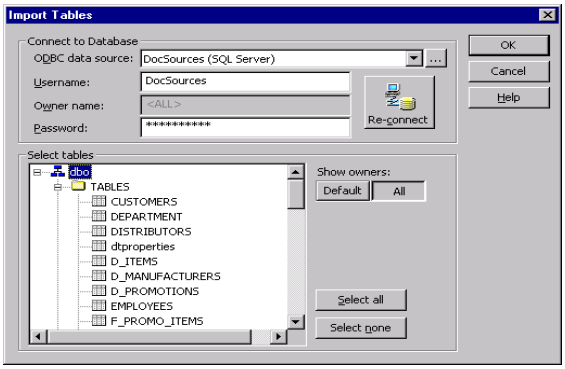
1. In the workspace, select the T_EMPLOYEES target definition.
2. Click Targets > Generate/Execute SQL.

   The Database Object Generation dialog box appears.
3. In the File Name field, enter the following text:

   `C:\[the installation directory]\MKT_EMP.SQL`

If you installed the client software in a different location, enter the appropriate drive letter and directory.

4. If you are connected to the source database from the previous lesson, click Disconnect, and then click Connect.
5. Select the ODBC data source to connect to the target database.
6. Enter the necessary user name and password, and then click Connect.



7. Select the Create Table, Drop Table, Foreign Key and Primary Key options.
8. Click the Generate and Execute button.

   The Designer runs the DDL code needed to create T_EMPLOYEES. If you want to review the actual code, click Edit SQL file to open the MKT_EMP.SQL file.
9. Click Close to exit.

What Comes Next

In the next lesson, you create a mapping to design the data flow between sources and targets. You specify how each column in the target table receives data from a column in the source table.

Chapter 4

Tutorial Lesson 3

This chapter includes the following topics:

Creating a Pass-Through Mapping

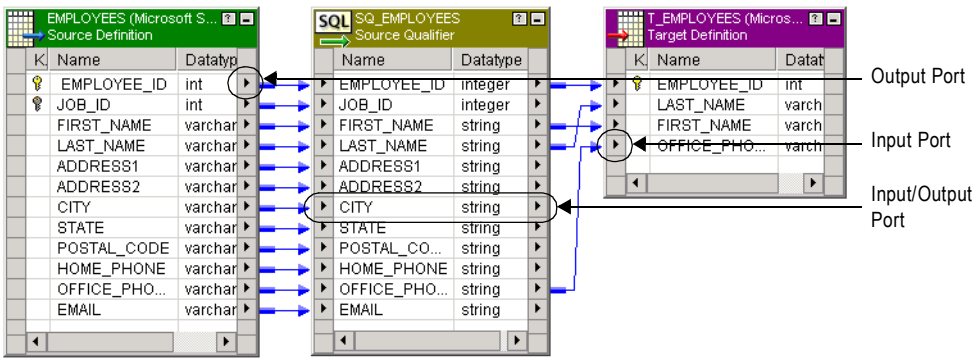
In the previous lesson, you added source and target definitions to the repository. You also generated and ran the SQL code to create target tables.

The next step is to create a mapping to depict the flow of data between sources and targets. For this step, you create a Pass-Through mapping. A Pass-Through mapping inserts all the source rows into the target.

To create and edit mappings, you use the Mapping Designer tool in the Designer. The mapping interface in the Designer is *component-based*. You add *transformations* to a mapping that depict how the Integration Service extracts and transforms data before it loads a target.

Figure 4-1 shows a mapping between a source and a target with a Source Qualifier transformation:

Figure 4-1. Pass-Through Mapping



The source qualifier represents the rows that the Integration Service reads from the source when it runs a session.

If you examine the mapping, you see that data flows from the source definition to the Source Qualifier transformation to the target definition through a series of input and output ports.

The source provides information, so it contains only output ports, one for each column. Each output port is connected to a corresponding input port in the Source Qualifier transformation. The Source Qualifier transformation contains both input and output ports. The target contains input ports.

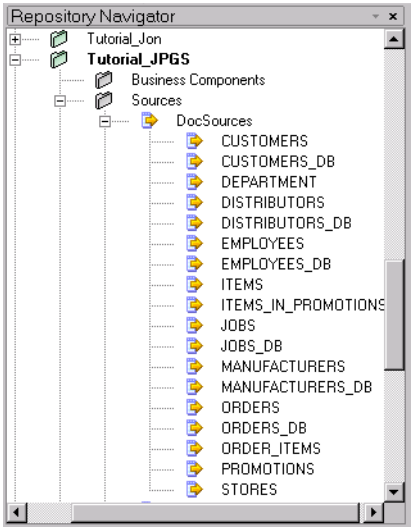
When you design mappings that contain different types of transformations, you can configure transformation ports as inputs, outputs, or both. You can rename ports and change the datatypes.

Creating a Mapping

In the following steps, you create a mapping and link columns in the source EMPLOYEES table to a Source Qualifier transformation.

To create a mapping:

1. Switch to the Mapping Designer.
2. In the Navigator, expand the Sources node in the tutorial folder, and then expand the DBD node containing the tutorial sources.



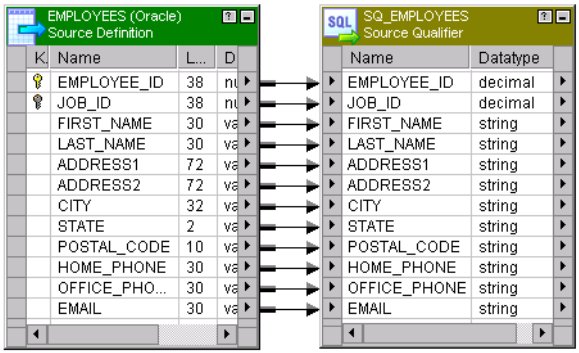
3. Drag the EMPLOYEES source definition into the Mapping Designer workspace.

   The Designer creates a new mapping and prompts you to provide a name.

4. In the Mapping Name dialog box, enter m_PhoneList as the name of the new mapping and click OK.

   The naming convention for mappings is m_*MappingName*.

The source definition appears in the workspace. The Designer creates a Source Qualifier transformation and connects it to the source definition.



5. Expand the Targets node in the Navigator to open the list of all target definitions.
6. Drag the T_EMPLOYEES target definition into the workspace.

   The target definition appears. The final step is to connect the Source Qualifier transformation to the target definition.

Connecting Transformations

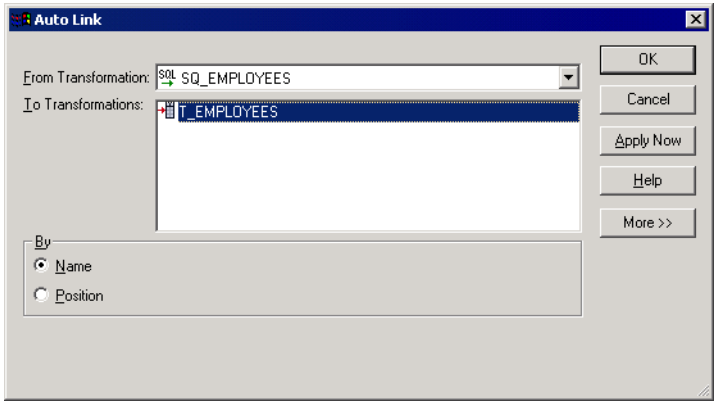
The port names in the target definition are the same as some of the port names in the Source Qualifier transformation. When you need to link ports between transformations that have the same name, the Designer can link them based on name.

In the following steps, you use the autolink option to connect the Source Qualifier transformation to the target definition.

To connect the Source Qualifier transformation to the target definition:

1. Click Layout > Autolink.

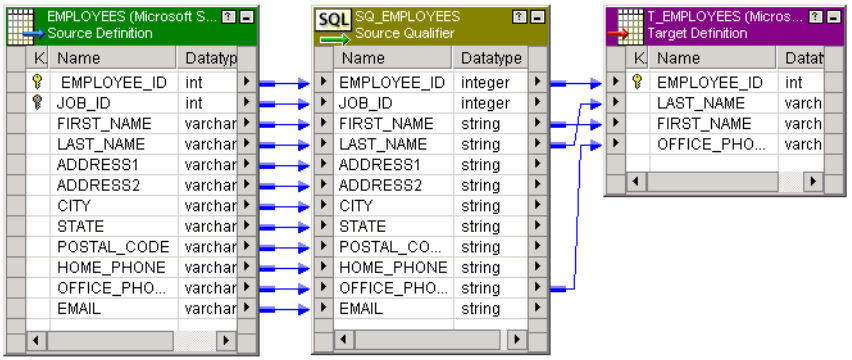
   The Auto Link dialog box appears.



2. Select T_EMPLOYEES in the To Transformations field. Verify that SQ_EMPLOYEES is in the From Transformation field.

3. Autolink by name and click OK.

   The Designer links ports from the Source Qualifier transformation to the target definition by name. A link appears between the ports in the Source Qualifier transformation and the target definition.



   Note: When you need to link ports with different names, you can drag from the port of one transformation to a port of another transformation or target. If you connect the wrong columns, select the link and press the Delete key.

4. Click Layout > Arrange.

5. In the Select Targets dialog box, select the T_EMPLOYEES target and click OK.

   The Designer rearranges the source, Source Qualifier transformation, and target from left to right, making it easy to see how one column maps to another.

6. Drag the lower edge of the source and Source Qualifier transformation windows until all columns appear.

7. Click Repository > Save to save the new mapping to the repository.

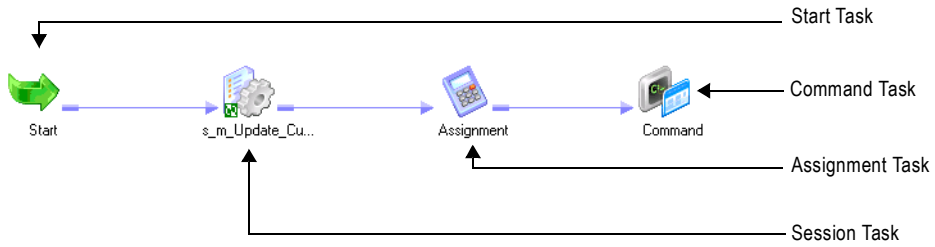
Creating Sessions and Workflows

A session is a set of instructions that tells the Integration Service how to move data from sources to targets. A session is a task, similar to other tasks available in the Workflow Manager. You create a session for each mapping that you want the Integration Service to run. The Integration Service uses the instructions configured in the session and mapping to move data from sources to targets.

A workflow is a set of instructions that tells the Integration Service how to execute tasks, such as sessions, email notifications, and shell commands. You create a workflow for sessions you want the Integration Service to run. You can include multiple sessions in a workflow to run sessions in parallel or sequentially. The Integration Service uses the instructions configured in the workflow to run sessions and other tasks.

Figure 4-2 shows a workflow with multiple branches and tasks:

Figure 4-2. Sample Workflow



You create and maintain tasks and workflows in the Workflow Manager.

In this lesson, you create a session and a workflow that runs the session. Before you create a session in the Workflow Manager, you need to configure database connections in the Workflow Manager.

Configuring Database Connections in the Workflow Manager

Before you can create a session, you need to provide the Integration Service with the information it needs to connect to the source and target databases. Configure database connections in the Workflow Manager. Database connections are saved in the repository.

For more information about working with database connections, see "Managing Connection Objects" in the *Workflow Administration Guide*.

To define a database connection:

1. Launch Workflow Manager.

   From the Windows Start menu, select Programs > Informatica PowerCenter > Informatica PowerCenter Client > Workflow Manager.

   From the Designer, click Tools > Workflow Manager.

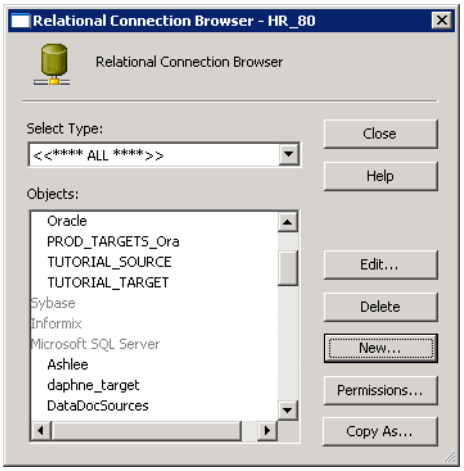
2. In the Workflow Manager, select the repository in the Navigator, and then click Repository > Connect.

3. Enter a user name and password to connect to the repository, and click Connect.

   Use the user profile and password you entered in Table 1-3 on page 4.

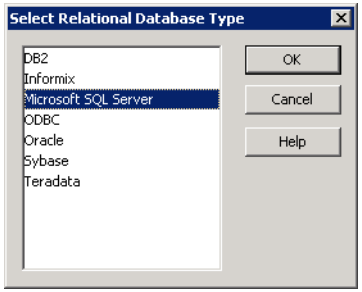
4. Click Connections > Relational.

   The Relational Connection Browser dialog box appears.



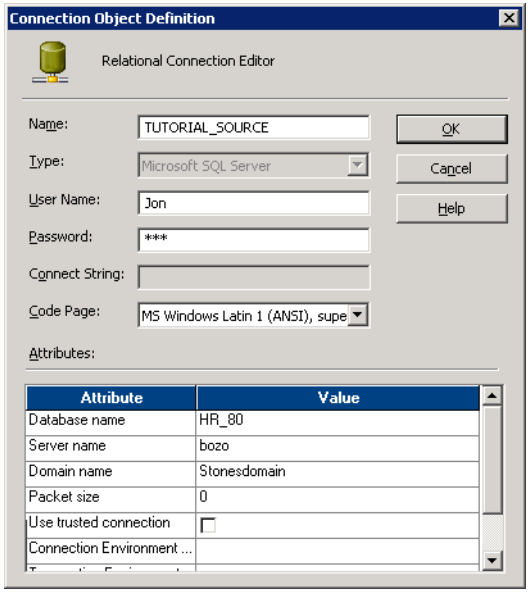
5. Click New in the Relational Connection Browser dialog box.

The Select Relational Database Type dialog box appears.



6. Select the appropriate database type and click OK.

   The Connection Object Definition dialog box appears with options appropriate to the selected database platform.



7. In the Name field, enter TUTORIAL_SOURCE as the name of the database connection.

   The Integration Service uses this name as a reference to this database connection.

8. Enter the database name. Enter the user name and password to connect to the database.

9. Select a code page for the database connection. The source code page must be a subset of the target code page.

   Use the database connection information you entered for the source database in Table 1-5 on page 5.

10. Enter additional information necessary to connect to this database, such as the native connect string, and click OK.

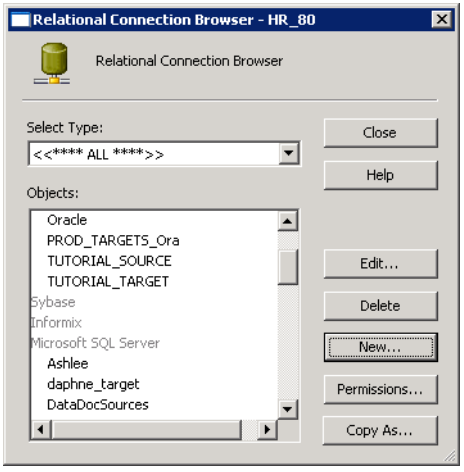
TUTORIAL_SOURCE now appears in the list of registered database connections in the Relational Connection Browser dialog box.

11. Repeat steps 5 to 10 to create another database connection called TUTORIAL_TARGET for the target database.

    The target code page must be a superset of the source code page.

    Use the database connection information you entered for the target database in Table 1-5 on page 5.

    When you finish, both TUTORIAL_SOURCE and TUTORIAL_TARGET appear in the list of registered database connections in the Relational Connection Browser dialog box.



12. Click Close.

13. Click Repository > Save to save the new database connections to the repository.

You have finished configuring the connections to the source and target databases. The next step is to create a session for the mapping m_PhoneList.

Creating a Reusable Session

You can create reusable or non-reusable sessions in the Workflow Manager. Create reusable sessions in the Task Developer. When you create a reusable session, you can use it in multiple workflows. Create non-reusable sessions in the Workflow Designer. When you create a non-reusable session, you can use it only in that workflow.

In the following steps, you create a reusable session that uses the mapping m_PhoneList. Then, you create a workflow that uses the reusable session.

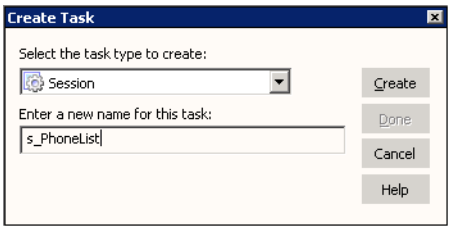
To create the session:

1. In the Workflow Manager Navigator, double-click the tutorial folder to open it.

2. Click Tools > Task Developer to open the Task Developer.

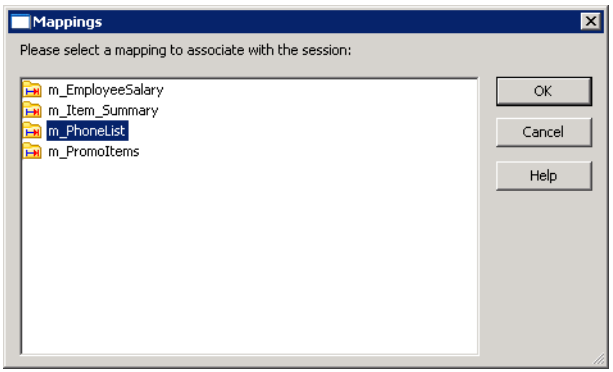
3. Click Tasks > Create.

   The Create Task dialog box appears.



4. Select Session as the task type to create.
5. Enter s_PhoneList as the session name and click Create.

   The Mappings dialog box appears.

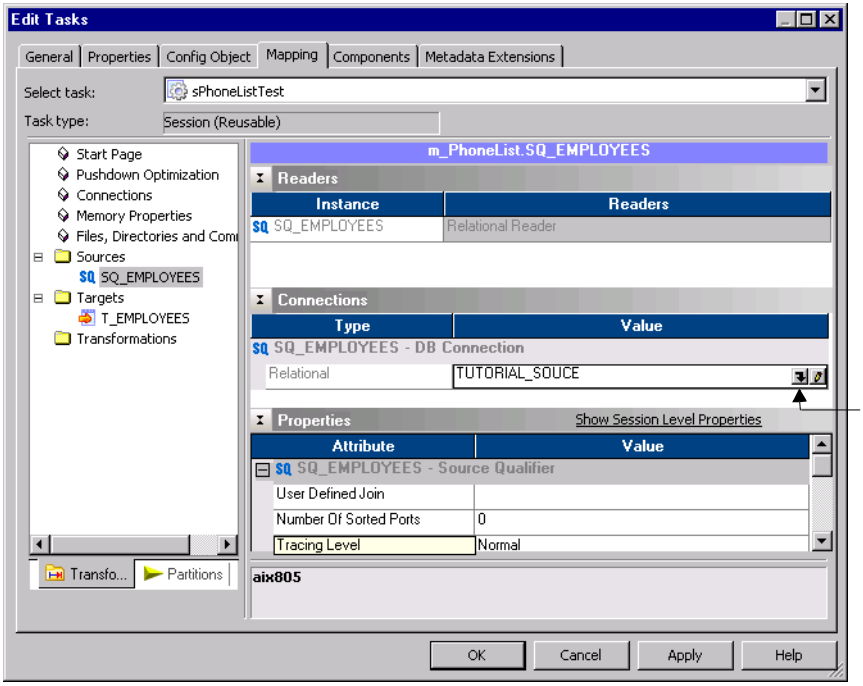


6. Select the mapping m_PhoneList, and click OK.

   The Workflow Manager creates a reusable Session task in the Task Developer workspace.
7. Click Done in the Create Task dialog box.
8. In the workspace, double-click s_PhoneList to open the session properties.

   The Edit Tasks dialog box appears. You use the Edit Tasks dialog box to configure and edit session properties, such as source and target database connections, performance properties, log options, and partitioning information. In this lesson, you use most default settings. You select the source and target database connections.
9. Click the Mapping tab.

10. Select Sources in the Transformations pane on the left.



11. In the Connections settings to the right, click the Browse Connections button in the Value column for the SQ_EMPLOYEES - DB Connection.

    The Relational Connection Browser appears.

12. Select TUTORIAL_SOURCE and click OK.

13. Select Targets in the Transformations pane.

14. In the Connections settings, click the Edit button in the Value column for the T_EMPLOYEES - DB Connection.

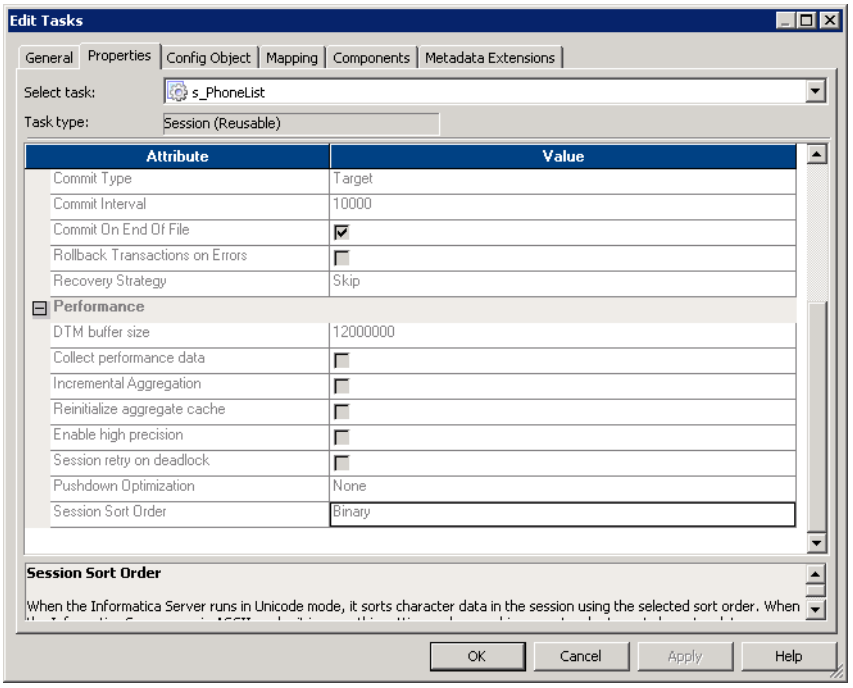
    The Relational Connection Browser appears.

15. Select TUTORIAL_TARGET and click OK.

16. Click the Properties tab.

17. Select a session sort order associated with the Integration Service code page.

For English data, use the Binary sort order. For more information about sort orders and code pages, see "Understanding Globalization" in the *Administrator Guide*.



These are the session properties you need to define for this session. For more information about the tabs and settings in the session properties, see "Session Properties Reference" in the *Workflow Administration Guide*.

18. Click OK to close the session properties with the changes you made.
19. Click Repository > Save to save the new session to the repository.

You have created a reusable session. The next step is to create a workflow that runs the session.

Creating a Workflow

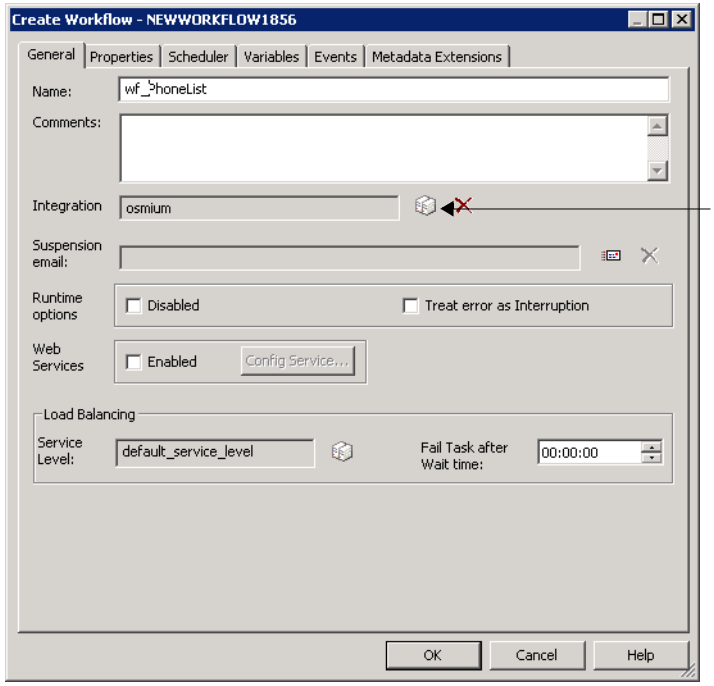
You create workflows in the Workflow Designer. When you create a workflow, you can include reusable tasks that you create in the Task Developer. You can also include non-reusable tasks that you create in the Workflow Designer.

In the following steps, you create a workflow that runs the session s_PhoneList.

To create a workflow:

1. Click Tools > Workflow Designer.
2. In the Navigator, expand the tutorial folder, and then expand the Sessions node.
3. Drag the session s_PhoneList to the Workflow Designer workspace.

The Create Workflow dialog box appears.



4. Enter wf_PhoneList as the name for the workflow.

   The naming convention for workflows is wf_*WorkflowName*.

5. Click the Browse Integration Services button to choose an Integration Service to run the workflow.

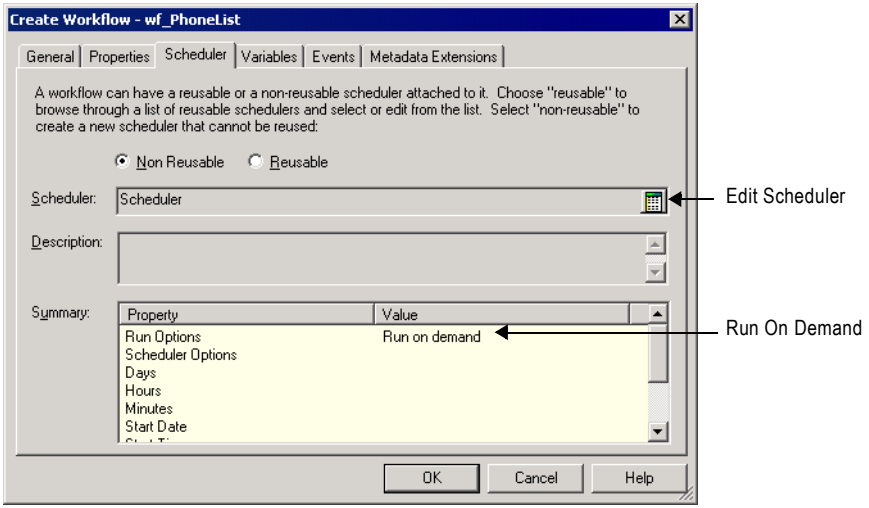
   The Integration Service Browser dialog box appears.

6. Select the appropriate Integration Service, and click OK.

7. Click the Properties tab.

8. Enter wf_PhoneList.log for the workflow log file name.

9. Click the Scheduler tab.



By default, the workflow is scheduled to run on demand. The Integration Service only runs the workflow when you manually start the workflow. You can configure workflows to run on a schedule. For example, you can schedule a workflow to run once a day or run on the last day of the month. Click the Edit Scheduler button to configure schedule options. For more information about scheduling workflows, see "Working with Workflows" in the *Workflow Administration Guide*.

10. Accept the default schedule for this workflow.

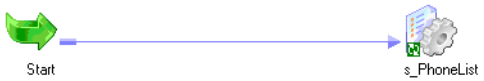
11. Click OK to close the Create Workflow dialog box.

    The Workflow Manager creates a new workflow in the workspace, including the reusable session you added. All workflows begin with the Start task, but you need to instruct the Integration Service which task to run next. To do this, you link tasks in the Workflow Manager.

    Note: You can click Workflows > Edit to edit the workflow properties at any time.

12. On the toolbar, click the link tasks button.

13. Drag from the Start task to the Session task.



14. Click Repository > Save to save the workflow in the repository.

    You can now run and monitor the workflow.

Running and Monitoring Workflows

When the Integration Service runs workflows, you can monitor workflow progress in the Workflow Monitor. You can view details about a workflow or task in either a Gantt Chart view or a Task view. You can start, stop, and abort workflows from the Workflow Monitor. The Workflow Monitor displays workflows that have run at least once.

In the following steps, you run a workflow and monitor it.

Opening the Workflow Monitor

You can configure the Workflow Manager to open the Workflow Monitor when you run a workflow in the Workflow Manager.

You can also open the Workflow Monitor from the Workflow Manager Navigator or from the Windows Start menu.

For more information about using the Workflow Monitor, see "Monitoring Workflows" in the *Workflow Administration Guide*.

To configure the Workflow Manager to open the Workflow Monitor:

1. In the Workflow Manager, click Tools > Options.
2. In the General tab, select Launch Workflow Monitor When Workflow Is Started.
3. Click OK.

Next, you run the workflow and open the Workflow Monitor.

Running the Workflow

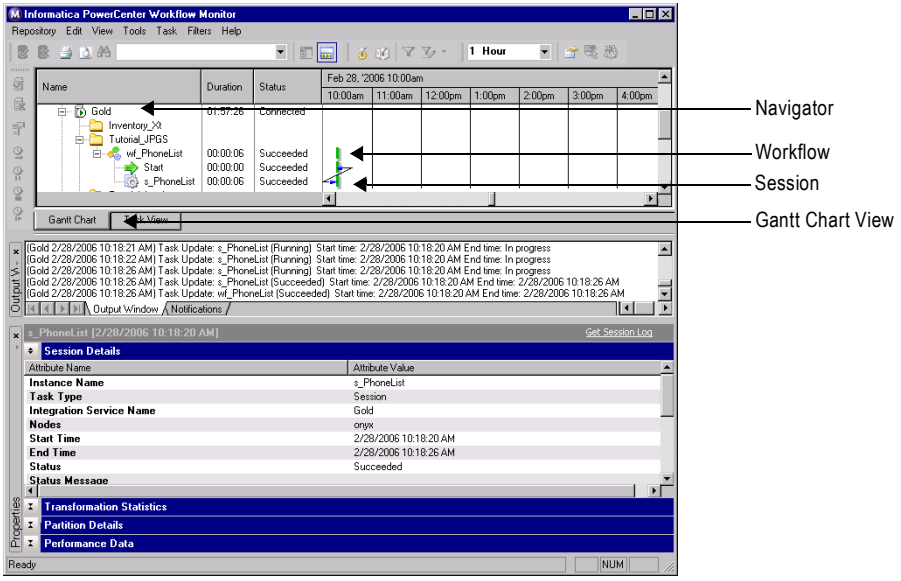
After you create a workflow containing a session, you can run it to move the data from the source to the target.

To run a workflow:

1. Verify the workflow is open in the Workflow Designer.
2. In the Workflow Manager, click Workflows > Start Workflow.

   Tip: You can also right-click the workflow in the Navigator and select Start Workflow.

The Workflow Monitor opens, connects to the repository, and opens the tutorial folder.



3. Click the Gantt Chart tab at the bottom of the Time window to verify the Workflow Monitor is in Gantt Chart view.

4. In the Navigator, expand the node for the workflow.

   All tasks in the workflow appear in the Navigator. For more information about Gantt Chart view, see "Monitoring Workflows" in the *Workflow Administration Guide*.

   The session returns the following results:

EMPLOYEE_ID	LAST_NAME	FIRST_NAME	OFFICE_PHONE
1921	Nelson	William	415-541-5145
1922	Page	Ian	415-541-5145
1923	Osborne	Lyle	415-541-5145
1928	DeSouza	Leo	415-541-5145
2001	S. MacDonald	Ira	415-541-5145
2002	Hill	Andy	415-541-5145
2003	Sawyer	Monisha	415-541-5145
2006	St. Jean	Bender	415-541-5145
2100	Johnson	Teddy	415-541-5145
2102	Steadman	Ono	415-541-5145
2103	Markowitz	John	415-541-5145

EMPLOYEE_ID	LAST_NAME	FIRST_NAME	OFFICE_PHONE
2109	Centre	Tom	415-541-5145

(12 rows affected)

What Comes Next

In Lesson 4, you create a mapping using aggregate values to learn how to modify data before it reaches the target. Then, you create a reusable session with the mapping, and you run the session in a workflow.

Chapter 5

Tutorial Lesson 4

This chapter includes the following topics:

Using Transformations

In this lesson, you create a mapping that contains a source, multiple transformations, and a target.

A transformation is a part of a mapping that generates or modifies data. Every mapping includes a Source Qualifier transformation, representing all data read from a source and temporarily stored by the Integration Service. In addition, you can add transformations that calculate a sum, look up a value, or generate a unique ID before the source data reaches the target.

Figure 5-1 shows the Transformation toolbar:

Figure 5-1. Transformation Toolbar

Table 5-1 lists the transformations displayed in the Transformations toolbar in the Designer:

Table 5-1. Transformation Descriptions

Transformation	Description
Aggregator	Performs aggregate calculations.
Application Multi-Group Source Qualifier	Represents the rows that the Integration Service reads from an application, such as an ERP source, when it runs a workflow.
Custom	Calls a procedure in a shared library or DLL.
Expression	Calculates a value.
External Procedure	Calls a procedure in a shared library or in the COM layer of Windows.
Filter	Filters data.
Input	Defines mapplet input rows. Available in the Mapplet Designer.
Java	Executes user logic coded in Java.
Joiner	Joins data from different databases or flat file systems.
Lookup	Looks up values.
MQ Series Source Qualifier	Represents the rows that the Integration Service reads from an MQSeries message queue source when it runs a workflow.
Normalizer	Source qualifier for COBOL sources. Can also use in the pipeline to normalize data from relational or flat file sources.
Output	Defines mapplet output rows. Available in the Mapplet Designer.
Rank	Limits records to a top or bottom range.
Router	Routes data into multiple transformations based on group conditions.
Sequence Generator	Generates primary keys.

Table 5-1. Transformation Descriptions

Transformation	Description
Sorter	Sorts data based on a sort key.
Source Qualifier	Represents the rows that the Integration Service reads from a relational or flat file source when it runs a workflow.
Stored Procedure	Calls a stored procedure.
Transaction Control	Defines commit and rollback transactions.
Union	Merges data from multiple databases or flat file systems.
Update Strategy	Determines whether to insert, delete, update, or reject records.
XML Generator	Reads data from one or more input ports and outputs XML through a single output port.
XML Parser	Reads XML from one input port and outputs data to one or more output ports.
XML Source Qualifier	Represents the rows that the Integration Service reads from an XML source when it runs a workflow.

For more information about using transformations, see "Working with Transformations" in the *Transformation Guide*. For more information about each transformation, see the corresponding chapter in the *Transformation Guide*.

In this lesson, you complete the following steps:

1. Create a new target definition to use in a mapping, and create a target table based on the new target definition.
2. Create a mapping using the new target definition. Add the following transformations to the mapping:
   - **Lookup transformation.** Finds the name of a manufacturer.
   - **Aggregator transformation.** Calculates the maximum, minimum, and average price of items from each manufacturer.
   - **Expression transformation.** Calculates the average profit of items, based on the average price.
3. Learn some tips for using the Designer.
4. Create a session and workflow to run the mapping, and monitor the workflow in the Workflow Monitor.

Creating a New Target Definition and Target

Before you create the mapping in this lesson, you need to design a new target definition that holds summary data about products from various manufacturers. This table includes the maximum and minimum price for products from a given manufacturer, an average price, and an average profit.

After you create the target definition, you create the table in the target database.

Creating a Target Definition

To create the target definition in this lesson, you copy the MANUFACTURERS source definition into the Target Designer. Then, you modify the target definition by adding columns to create the definition you want.

Note: You can also manually create a target definition, import the definition for an existing target from a database, or create a relational target from a transformation in the Designer. For more information about creating target definitions, see "Working with Targets" in the *Designer Guide*.

To create the new target definition:

1. Open the Designer, connect to the repository, and open the tutorial folder.
2. Click Tools > Target Designer.
3. Drag the MANUFACTURERS source definition from the Navigator to the Target Designer workspace.

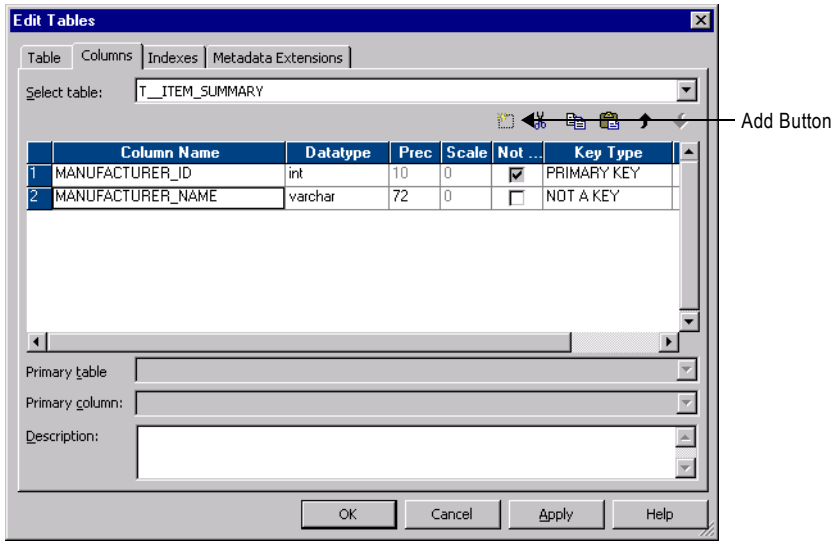
   The Designer creates a new target definition, MANUFACTURERS, with the same column definitions as the MANUFACTURERS source definition and the same database type.

   Next, you add new target column definitions.
4. Double-click the MANUFACTURERS target definition to open it.

   The Edit Tables dialog box appears.
5. Click Rename and name the target definition T_ITEM_SUMMARY.
6. Optionally, change the database type for the target definition. You can select the correct database type when you edit the target definition.
7. Click the Columns tab.

   The target column definitions are the same as the MANUFACTURERS source definition.

8. For the MANUFACTURER_NAME column, change precision to 72, and clear the Not Null column.

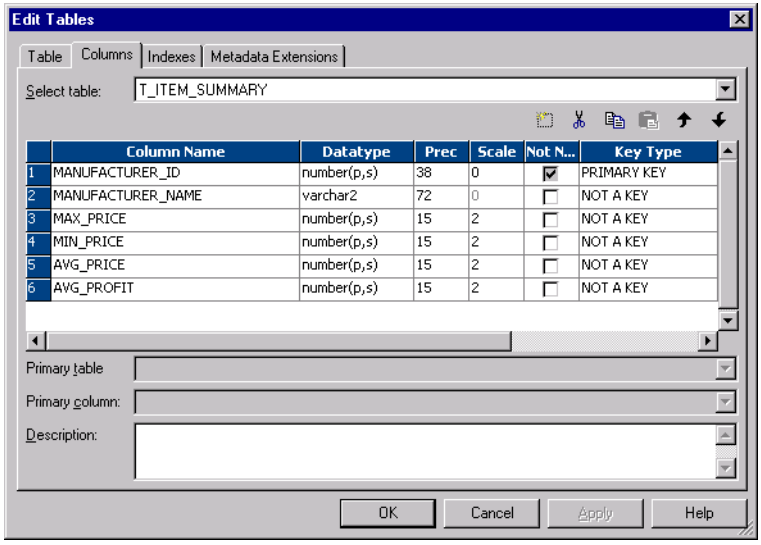


9. Add the following columns with the Money datatype, and select Not Null:

   - MAX_PRICE
   - MIN_PRICE
   - AVG_PRICE
   - AVG_PROFIT

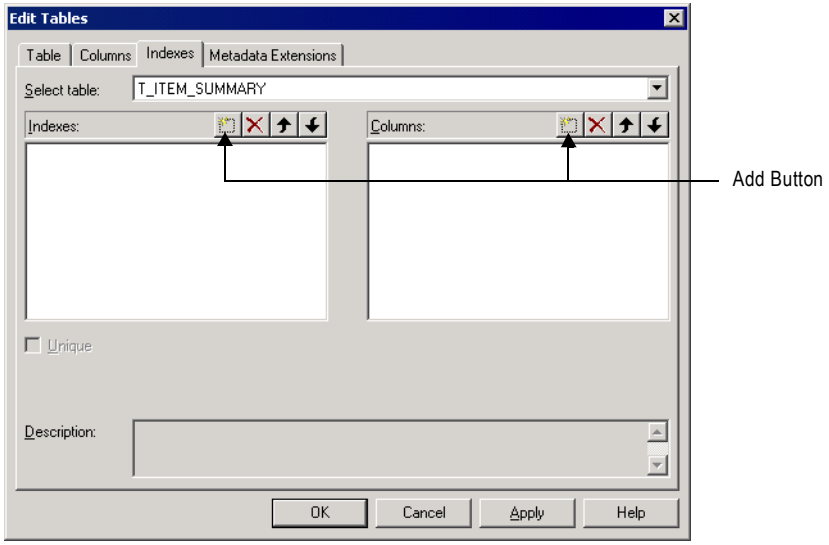
   Use the default precision and scale with the Money datatype. If the Money datatype does not exist in the database, use Number (p,s) or Decimal. Change the precision to 15 and the scale to 2.

10. Click Apply.



11. Click the Indexes tab to add an index to the target table.

If the target database is Oracle, skip to the final step. You cannot add an index to a column that already has the PRIMARY KEY constraint added to it.



12. In the Indexes section, click the Add button.

13. Enter IDX_MANUFACTURER_ID as the name of the new index, and then press Enter.

14. Select the Unique index option.

15. In the Columns section, click Add.

The Add Column To Index dialog box appears. It lists the columns you added to the target definition.

16. Select MANUFACTURER_ID and click OK.
17. Click OK to save the changes to the target definition, and then click Repository > Save.

Creating a Target Table

In the following steps, you use the Designer to generate and execute the SQL script to create a target table based on the target definition you created.

To create the table in the database:

1. Select the table T_ITEM_SUMMARY, and then click Targets > Generate/Execute SQL.
2. In the Database Object Generation dialog box, connect to the target database.
3. Click Generate from Selected tables, and select the Create Table, Primary Key, and Create Index options.

   Leave the other options unchanged.
4. Click Generate and Execute.

   The Designer notifies you that the file MKT_EMP.SQL already exists.
5. Click OK to override the contents of the file and create the target table.

   The Designer runs the SQL script to create the T_ITEM_SUMMARY table.
6. Click Close.

Creating a Mapping with Aggregate Values

In the next step, you create a mapping with the following mapping logic:

- Finds the most expensive and least expensive item in the inventory for each manufacturer. Use an Aggregator transformation to perform these calculations.
- Calculates the average price and profitability of all items from a given manufacturer. Use an Aggregator and an Expression transformation to perform these calculations.

You need to configure the mapping to perform both simple and aggregate calculations. For example, use the MIN and MAX functions to find the most and least expensive items from each manufacturer.

Creating a Mapping with T_ITEM_SUMMARY

First, create a mapping with the target definition you just created.

To create the new mapping:

1. Switch from the Target Designer to the Mapping Designer.
2. Click Mappings > Create.
3. When prompted to close the current mapping, click Yes.
4. In the Mapping Name dialog box, enter m_ItemSummary as the name of the mapping.
5. From the list of sources in the tutorial folder, drag the ITEMS source definition into the mapping.
6. From the list of targets in the tutorial folder, drag the T_ITEM_SUMMARY target definition into the mapping.

Creating an Aggregator Transformation

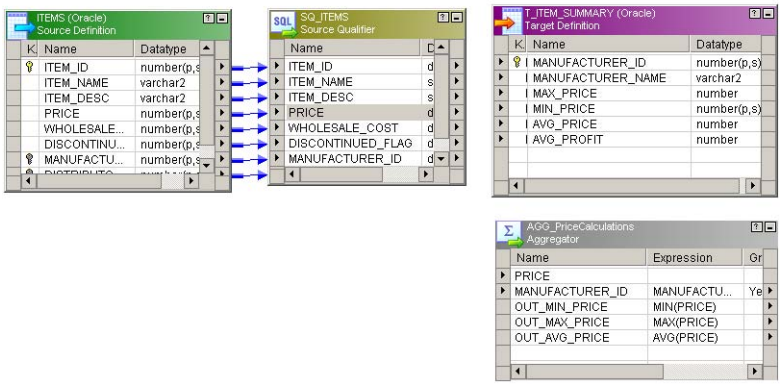
Next, add an Aggregator transformation to calculate the average, maximum, and minimum prices of items from each manufacturer.

To add the Aggregator transformation:

1. Click Transformation > Create to create an Aggregator transformation.
2. Click Aggregator and name the transformation AGG_PriceCalculations. Click Create, and then click Done.

   The naming convention for Aggregator transformations is AGG_*TransformationName*.

The Mapping Designer adds an Aggregator transformation to the mapping.



3. Click Layout > Link Columns.

   When you drag ports from one transformation to another, the Designer copies the port description and links the original port to its copy.

   If you click Layout > Copy Columns, every port you drag is copied, but not linked.

4. From the Source Qualifier transformation, drag the PRICE column into the Aggregator transformation.

   A copy of the PRICE port now appears in the new Aggregator transformation. The new port has the same name and datatype as the port in the Source Qualifier transformation.

   The Aggregator transformation receives data from the PRICE port in the Source Qualifier transformation. You need this information to calculate the maximum, minimum, and average product price for each manufacturer.

5. Drag the MANUFACTURER_ID port into the Aggregator transformation.

   You need another input port, MANUFACTURER_ID, to provide the information for the equivalent of a GROUP BY statement. By adding this second input port, you can define the groups (in this case, manufacturers) for the aggregate calculation. This organizes the data by manufacturer.

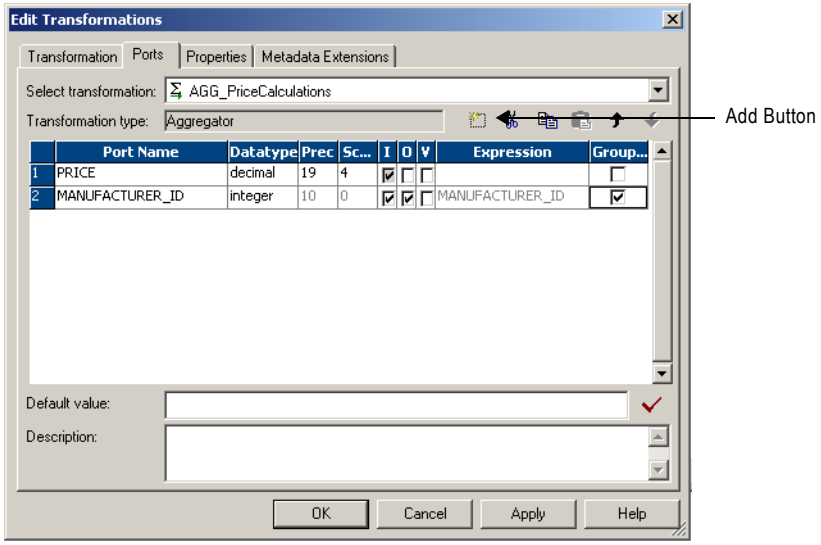
6. Double-click the Aggregator transformation, and then click the Ports tab.

7. Clear the Output (O) column for PRICE.

   You want to use this port as an input (I) only, not as an output (O). Later, you use data from PRICE to calculate the average, maximum, and minimum prices.

8. Select the Group By option for the MANUFACTURER_ID column.

9. Click the Add button three times to add three new ports.



When you select the Group By option for MANUFACTURER_ID, the Integration Service groups all incoming rows by manufacturer ID when it runs the session.

10. Configure the following ports:

Name	Datatype	Precision	Scale	I	O	V
OUT_MIN_PRICE	Decimal	19	2	No	Yes	No
OUT_MAX_PRICE	Decimal	19	2	No	Yes	No
OUT_AVG_PRICE	Decimal	19	2	No	Yes	No

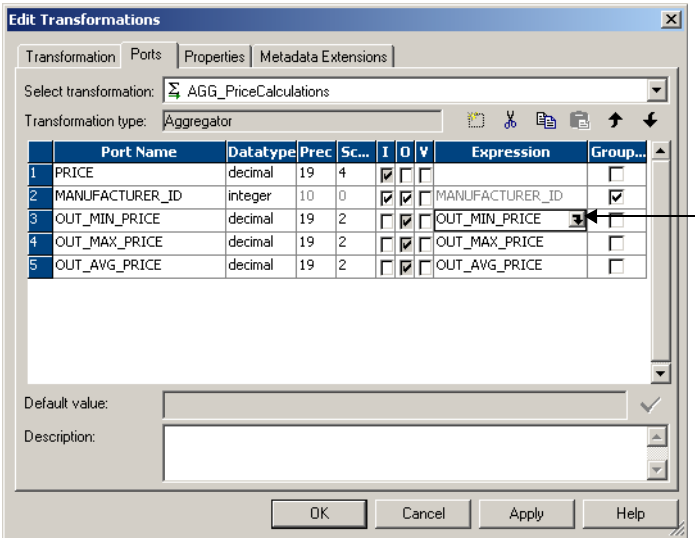
Tip: You can select each port and click the Up and Down buttons to position the output ports after the input ports in the list.

11. Click Apply to save the changes.

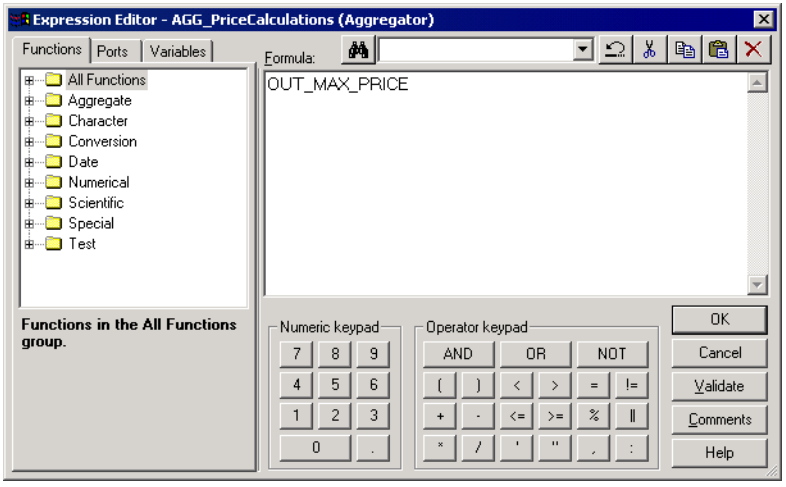
Now, you need to enter the expressions for all three output ports, using the functions MAX, MIN, and AVG to perform aggregate calculations.

To enter the first aggregate calculation:

1. Click the open button in the Expression column of the OUT_MAX_PRICE port to open the Expression Editor.



2. Delete the text OUT_MAX_PRICE.



The Formula section of the Expression Editor displays the expression as you develop it. Use other sections of this dialog box to select the input ports to provide values for an expression, enter literals and operators, and select functions to use in the expression.

For more information about using the Expression Editor, see "Working with Transformations" in the *Transformation Guide*.

3. Double-click the Aggregate heading in the Functions section of the dialog box.

   A list of all aggregate functions now appears.

4. Double-click the MAX function on the list.

   The MAX function appears in the window where you enter the expression. To perform the calculation, you need to add a reference to an input port that provides data for the expression.

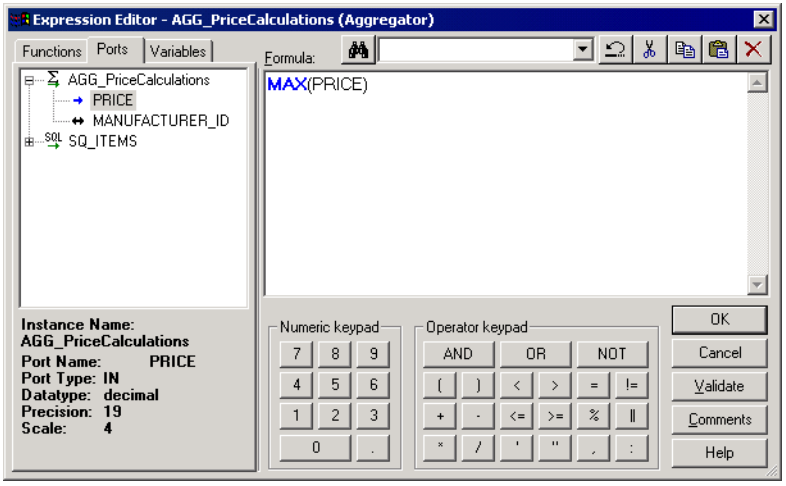
5. Move the cursor between the parentheses next to MAX.

6. Click the Ports tab.

   This section of the Expression Editor displays all the ports from all transformations appearing in the mapping.

7. Double-click the PRICE port appearing beneath AGG_PriceCalculations.

   A reference to this port now appears within the expression. The final step is to validate the expression.



8. Click Validate.

   The Designer displays a message that the expression parsed successfully. The syntax you entered has no errors.

9. Click OK to close the message box from the parser, and then click OK again to close the Expression Editor.

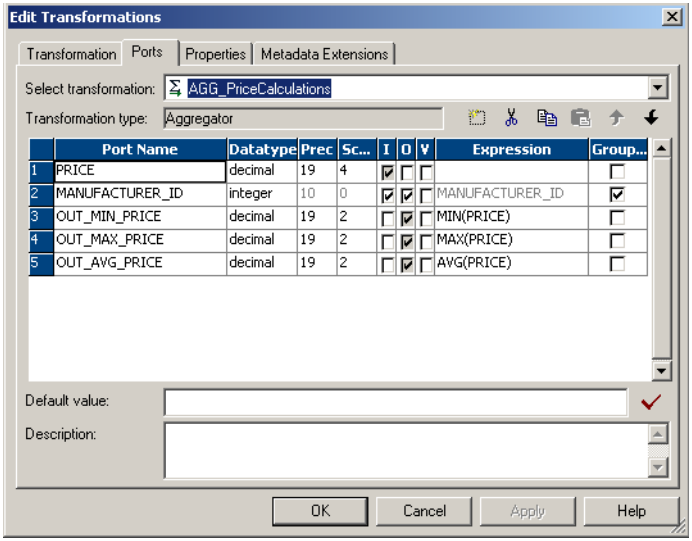
To enter the remaining aggregate calculations:

1. Enter and validate the following expressions for the other two output ports:

Port	Expression
OUT_MIN_PRICE	MIN(PRICE)
OUT_AVG_PRICE	AVG(PRICE)

   Both MIN and AVG appear in the list of Aggregate functions, along with MAX.

2. Click OK to close the Edit Transformations dialog box.



3. Click Repository > Save, and view the messages in the Output window.

   When you save changes to the repository, the Designer validates the mapping. You may notice an error message indicating that you have not connected the targets. You connect the targets later in this lesson.

Creating an Expression Transformation

Now that you have calculated the highest, lowest, and average prices for items, the next step is to calculate the average profitability of items from each manufacturer. While such calculations are normally more complex, you simply multiply the average price by 0.2 (20%).

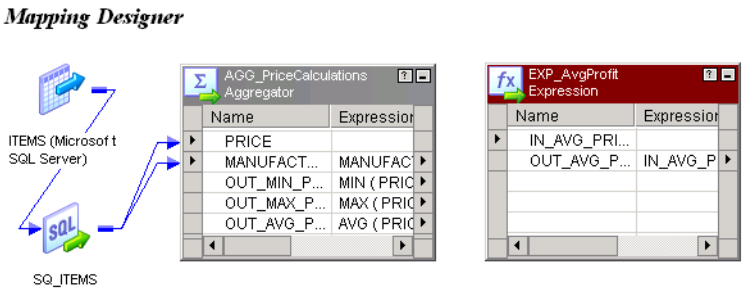
To add this information to the target, you need to create an Expression transformation that takes the average price of items from a manufacturer, performs the calculation, and then passes the result along to the target. As you develop transformations, you connect transformations using the output of one transformation as an input for others.

To add an Expression transformation:

1. Click Transformation > Create.
2. Select Expression and name the transformation EXP_AvgProfit. Click Create, and then click Done.

   The naming convention for Expression transformations is EXP_*TransformationName*.

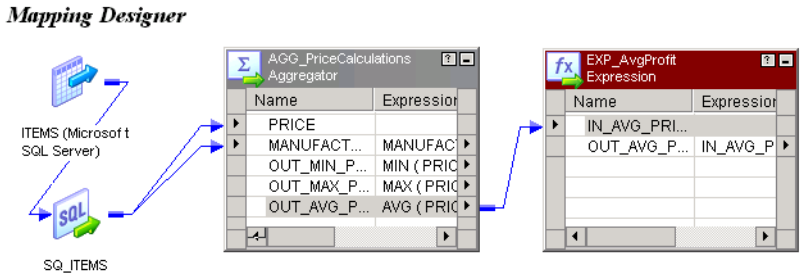
The Mapping Designer adds an Expression transformation to the mapping.



3. Open the Expression transformation.
4. Add a new input port, IN_AVG_PRICE, using the Decimal datatype with precision of 19 and scale of 2.
5. Add a new output port, OUT_AVG_PROFIT, using the Decimal datatype with precision of 19 and scale of 2.

   Note: Verify OUT_AVG_PROFIT is an output port, not an input/output port. You cannot enter expressions in input/output ports.
6. Enter the following expression for OUT_AVG_PROFIT:

   ```
   IN_AVG_PRICE * 0.2
   ```
7. Validate the expression.
8. Close the Expression Editor, and then close the EXP_AvgProfit transformation.
9. Connect OUT_AVG_PRICE from the Aggregator to the new input port.



10. Click Repository > Save.

Creating a Lookup Transformation

The source table in this mapping includes information about the manufacturer ID. However, you want the manufacturer name in the target table to make the summary data easier to read. In the following steps, you use a Lookup transformation to find each manufacturer name in the MANUFACTURERS table based on the manufacturer ID in the source table.

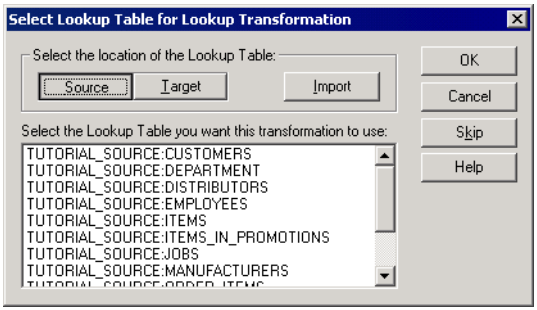
To add the Lookup transformation:

1. Create a Lookup transformation and name it LKP_Manufacturers.

   The naming convention for Lookup transformations is LKP_*TransformationName*.

   A dialog box prompts you to identify the source or target database to provide data for the lookup. When you run a session, the Integration Service must access the lookup table.

2. Click Source.



3. Select the MANUFACTURERS table from the list and click OK.

4. Click Done to close the Create Transformation dialog box.

   The Designer now adds the transformation.

   Use source and target definitions in the repository to identify a lookup source for the Lookup transformation. Alternatively, using the Import button, you can import a lookup source.

5. Open the Lookup transformation.

6. Add a new input port, IN_MANUFACTURER_ID, using the same datatype as MANUFACTURER_ID.

   In a later step, you connect the MANUFACTURER_ID port from the Aggregator transformation to this input port. IN_MANUFACTURER_ID receives MANUFACTURER_ID values from the Aggregator transformation. When the Lookup transformation receives a new value through this input port, it looks up the matching value from MANUFACTURERS.

   Note: By default, the Lookup transformation queries and stores the contents of the lookup table before the rest of the transformation runs, so it performs the join through a local copy of the table that it has cached. For more information about caching the lookup table, see “Lookup Caches” in the *Transformation Guide*.

7. Click the Condition tab, and click the Add button.

   An entry for the first condition in the lookup appears. Each row represents one condition in the WHERE clause that the Integration Service generates when querying records.

8. Verify the following settings for the condition:

Lookup Table Column	Operator	Transformation Port
MANUFACTURER_ID	=	IN_MANUFACTURER_ID

Note: If the datatypes (including precision and scale) of these two columns do not match, the Designer displays a message and marks the mapping invalid.

9. View the Properties tab.

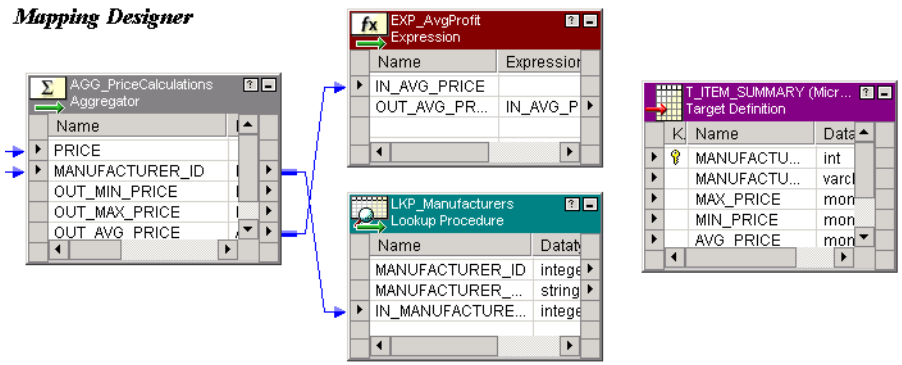
   Do not change settings in this section of the dialog box. For more information about the Lookup properties, see "Lookup Transformation" in the *Transformation Guide*.

10. Click OK.

    You now have a Lookup transformation that reads values from the MANUFACTURERS table and performs lookups using values passed through the IN_MANUFACTURER_ID input port. The final step is to connect this Lookup transformation to the rest of the mapping.

11. Click Layout > Link Columns.

12. Connect the MANUFACTURER_ID output port from the Aggregator transformation to the IN_MANUFACTURER_ID input port in the Lookup transformation.



13. Click Repository > Save.

Connecting the Target

You have set up all the transformations needed to modify data before writing to the target. So far, you have performed the following tasks:

- Created a target definition and target table.
- Created a mapping.
- Added transformations.

The final step is to connect to the target.

To connect to the target:

1. Drag the following output ports to the corresponding input ports in the target:

Transformation	Output Port	Target Input Port
Lookup	MANUFACTURER_ID	MANUFACTURER_ID
Lookup	MANUFACTURER_NAME	MANUFACTURER_NAME
Aggregator	OUT_MIN_PRICE	MIN_PRICE
Aggregator	OUT_MAX_PRICE	MAX_PRICE
Aggregator	OUT_AVG_PRICE	AVG_PRICE
Expression	OUT_AVG_PROFIT	AVG_PROFIT

2. Click Repository > Save.

   Verify mapping validation in the Output window.

Designer Tips

The steps below include tips for using the Designer. You learn how to complete the following tasks:

- Use the Overview window to navigate the workspace.
- Arrange the transformations in the workspace.

Using the Overview Window

When you create a mapping with many transformations, you may not be able to see the entire mapping in the workspace. In the following steps, you use the Overview window to navigate around the workspace containing the mapping you just created.

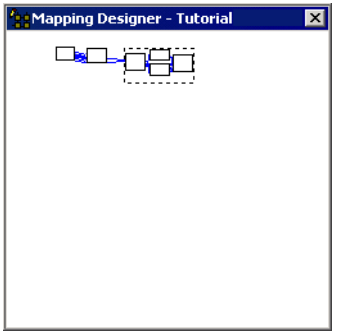
To use the Overview window:

1. Click View > Overview Window.

   You can also use the Toggle Overview Window icon.

   A window appears, displaying a smaller version of the mapping.

2. Drag the viewing rectangle (the dotted square) within this window.

   As you move the viewing rectangle, the perspective on the mapping changes.

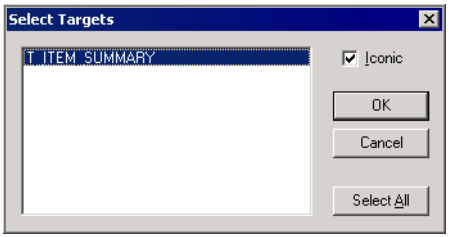
Arranging Transformations

The Designer can arrange the transformations in a mapping. When you use this option to arrange the mapping, you can arrange the transformations in normal view, or as icons.

To arrange a mapping:

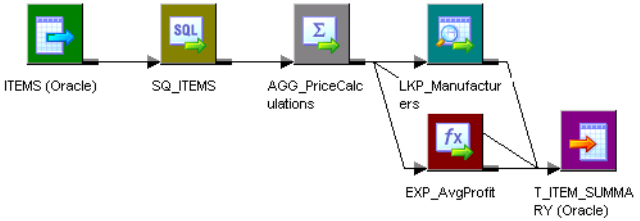
1. Click Layout > Arrange.

   The Select Targets dialog box appears showing all target definitions in the mapping.



2. Select Iconic to arrange the transformations as icons in the workspace.
3. Select T_ITEM_SUMMARY and click OK.

   The Designer arranges all transformations in the pipeline connected to the T_ITEM_SUMMARY target definition.



Creating a Session and Workflow

You have two mappings:

- **m_PhoneList.** A pass-through mapping that reads employee names and phone numbers.
- **m_ItemSummary.** A more complex mapping that performs simple and aggregate calculations and lookups.

You have a reusable session based on m_PhoneList. Next, you create a session for m_ItemSummary in the Workflow Manager. You create a workflow that runs both sessions.

Creating the Session

Open the Workflow Manager and connect to the repository if it is not open already.

To create the session:

1. Open the Task Developer and click Tasks > Create.
2. Create a Session task and name it s_ItemSummary. Click Create.

   In the Mappings dialog box, select the mapping m_ItemSummary and click OK.
3. Click Done.
4. Open the session properties for s_ItemSummary.
5. Click the Connections setting on the Mapping tab. Select the source database connection TUTORIAL_SOURCE for SQ_ITEMS.

   Use the database connection you created in “Configuring Database Connections in the Workflow Manager” on page 39.
6. Click the Connections setting on the Mapping tab. Select the target database connection TUTORIAL_TARGET for T_ITEM_SUMMARY.

   Use the database connection you created in “Configuring Database Connections in the Workflow Manager” on page 39.
7. Close the session properties and click Repository > Save.

Now that you have two sessions, you can create a workflow and include both sessions in the workflow. When you run the workflow, the Integration Service runs all sessions in the workflow, either simultaneously or in sequence, depending on how you arrange the sessions in the workflow.

Creating the Workflow

You can group sessions in a workflow to improve performance, or to ensure that targets load in a set order. In the following steps, you create a workflow that runs the sessions s_PhoneList and s_ItemSummary concurrently.

To create a workflow:

1. Click Tools > Workflow Designer.
2. Click Workflows > Create to create a new workflow.

   When you have a workflow already open, the Workflow Manager prompts you to close the current workflow. Click Yes to close any current workflow.

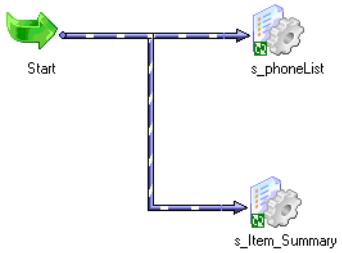
   The workflow properties appear.
3. Name the workflow wf_ItemSummary_PhoneList.
4. Click the Browse Integration Service button to select an Integration Service to run the workflow.

   The Integration Service Browser dialog box appears.
5. Select the Integration Service you use, and click OK.
6. Click the Properties tab and select Write Workflow Log to File.

   The default name of the workflow log file is wf_ItemSummary_PhoneList.log.
7. Click the Scheduler tab.

   By default, the workflow is scheduled to run on demand. Keep this default.
8. Click OK to close the Create Workflow dialog box.

   The Workflow Manager creates a new workflow in the workspace including the Start task.
9. From the Navigator, drag the s_ItemSummary session to the workspace. Then, drag the s_PhoneList session to the workspace.
10. Click the link tasks button on the toolbar.
11. Drag from the Start task to the s_ItemSummary Session task.
12. Drag from the Start task to the s_PhoneList Session task.



By default, when you link both sessions directly to the Start task, the Integration Service runs both sessions at the same time when you run the workflow. If you want the Integration Service to run the sessions one after the other, connect the Start task to one session, and connect that session to the other session.

13. Click Repository > Save to save the workflow in the repository.

    You can now run and monitor the workflow.

Running the Workflow

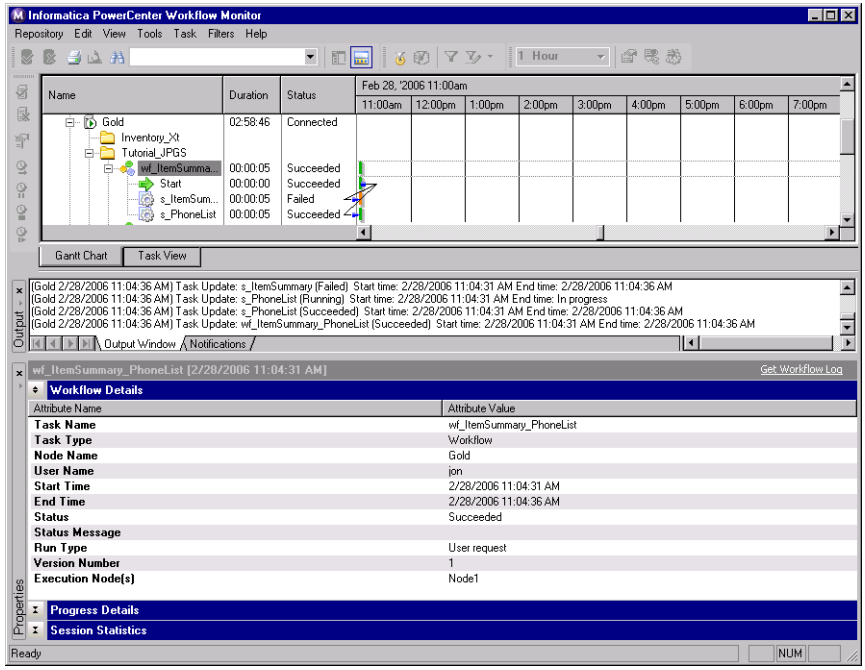
After you create the workflow containing the sessions, you can run it and use the Workflow Monitor to monitor the workflow progress.

To run a workflow:

1. Right-click the Start task in the workspace and select Start Workflow from Task.

   Tip: You can also right-click the workflow in the Navigator and select Start Workflow.

The Workflow Monitor opens and connects to the repository and opens the tutorial folder.



Tip: If the Workflow Monitor does not show the current workflow tasks, right-click the tutorial folder and select Get Previous Runs.

2. Click the Gantt Chart tab at the bottom of the Time window to verify the Workflow Monitor is in Gantt Chart view.

Note: You can also click the Task View tab at the bottom of the Time window to view the Workflow Monitor in Task view. You can switch back and forth between views at any time.

3. In the Navigator, expand the node for the workflow.

All tasks in the workflow appear in the Navigator.

The results from running the s_ItemSummary session are as follows:

MANUFACTURER_ID	MANUFACTURER_ NAME	MAX_ PRICE	MIN_ PRICE	AVG_ PRICE	AVG_ PROFIT
100	Nike	365.00	169.95	261.24	52.25
101	OBrien	188.00	44.95	134.32	26.86
102	Mistral	390.00	70.00	200.00	40.00
103	Spinnaker	70.00	29.00	52.98	10.60
104	Head	179.00	52.00	98.67	19.73

MANUFACTURER_ID	MANUFACTURER_ NAME	MAX_ PRICE	MIN_ PRICE	AVG_ PRICE	AVG_ PROFIT
105	Jesper	325.00	34.95	133.65	26.73
106	Acme	195.00	56.95	143.65	28.73
107	Medallion	235.00	19.95	98.65	19.73
108	Sportstar	280.00	18.00	149.00	29.80
109	WindJammer	430.00	395.00	412.50	82.50
110	Monsoon	280.00	280.00	280.00	56.00

(11 rows affected)

Viewing the Logs

You can view workflow and session logs using the Log Viewer or using the log files.

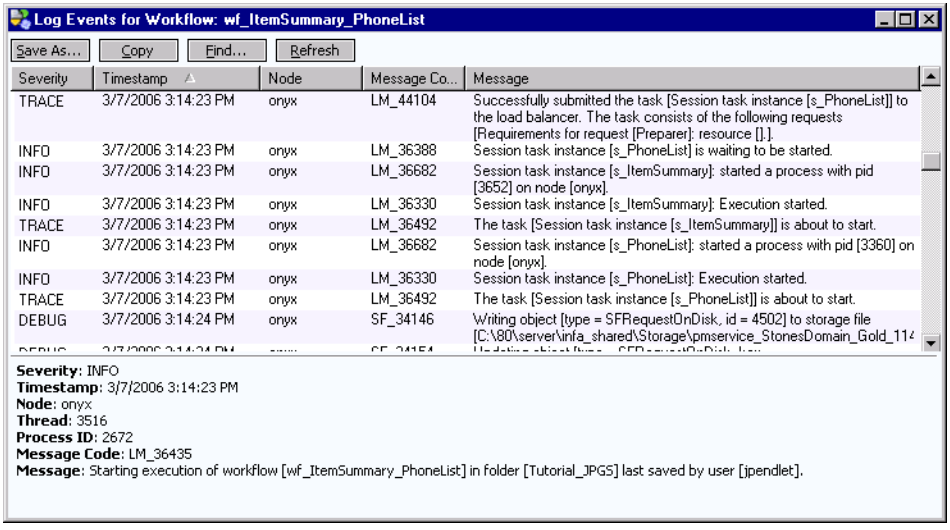
To view logs using the Log Viewer, in the Workflow Monitor right-click the workflow and select Get Workflow Log. The Log Events window for the selected workflow appears. You can also right-click a session and select Get Session Log.

The Log Viewer provides detailed information on each event performed during the workflow run. You can select a row in the log and the full text of the message appears in the section at the bottom of the window. You can use the Search button to search the log for specific text within a message. Sort the log file by column by clicking on the column heading.

You can click Save As to save the log as an XML document. Click Find to search for keywords in the log.

Figure 5-2 shows a sample Log Viewer window:

Figure 5-2. Sample Log Viewer



Log Files

When you created the workflow, the Designer assigned default workflow and session log names and locations on the Properties tab. During the session run, the Integration Service writes the log files to the locations specified in the session properties.

What Comes Next

In Lesson 5, you create a mapping with Filter, Stored Procedure, and Sequence Generator transformations that outputs data to a fact table and its dimension tables. Then, you create a reusable session with the mapping and run the session in a workflow. You define a link condition so the Integration Service runs the session only if the current date is before a date you specify.

Chapter 6

Tutorial Lesson 5

This chapter includes the following topics:

Creating a Mapping with Fact and Dimension Tables

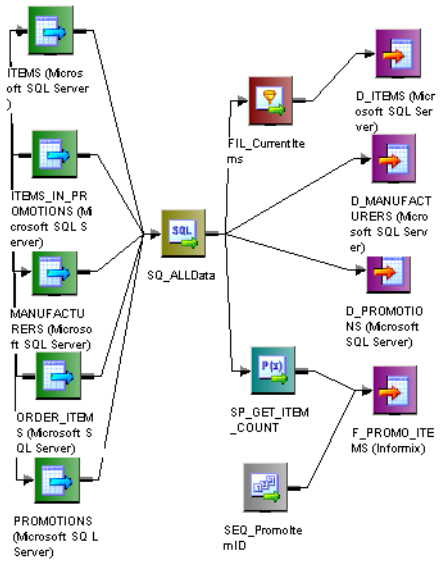
In previous lessons, you used the Source Qualifier, Expression, Aggregator, and Lookup transformations in mappings. In this lesson, you learn how to use the following transformations:

- **Stored Procedure.** Call a stored procedure and capture its return values.
- **Filter.** Filter data that you do not need, such as discontinued items in the ITEMS table.
- **Sequence Generator.** Generate unique IDs before inserting rows into the target.

You create a mapping that outputs data to a fact table and its dimension tables.

Figure 6-1 displays the mapping you create in this lesson:

Figure 6-1. Mapping with Fact and Dimension Tables



Creating Targets

Before you create the mapping, create the following target tables:

- F_PROMO_ITEMS (a fact table of promotional items)
- D_ITEMS, D_PROMOTIONS, and D_MANUFACTURERS (the dimensional tables)

For more information about fact and dimension tables, see "Creating Cubes and Dimensions" in the *Designer Guide*.

To design the new targets:

1. Open the Designer, connect to the repository, and open the tutorial folder.
2. Switch to the Target Designer.

To clear the workspace, right-click the workspace and select Clear All.

3. Click Targets > Create.
4. In the Create Target Table dialog box, enter F_PROMO_ITEMS as the name of the new target table, select the database type, and click Create.
5. Repeat step 4 to create the other tables needed for this schema: D_ITEMS, D_PROMOTIONS, and D_MANUFACTURERS. When you have created all these tables, click Done.
6. Open each new target, and add the following columns to the appropriate table:

D_ITEMS

Column	Datatype	Precision	Not Null	Key
ITEM_ID	Integer	NA	Not Null	Primary Key
ITEM_NAME	Varchar	72		
PRICE	Money	default		

D_PROMOTIONS

Column	Datatype	Precision	Not Null	Key
PROMOTION_ID	Integer	NA	Not Null	Primary Key
PROMOTION_NAME	Varchar	72		
DESCRIPTION	Varchar	default		
START_DATE	Datetime	default		
END_DATE	Datetime	default		

D_MANUFACTURERS

Column	Datatype	Precision	Not Null	Key
MANUFACTURER_ID	Integer	NA	Not Null	Primary Key
MANUFACTURER_NAME	Varchar	72		

F_PROMO_ITEMS

Column	Datatype	Precision	Not Null	Key
PROMO_ITEM_ID	Integer	NA	Not Null	Primary Key
FK_ITEM_ID	Integer	NA		Foreign Key
FK_PROMOTION_ID	Integer	NA		Foreign Key
FK_MANUFACTURER_ID	Integer	NA		Foreign Key

Column	Datatype	Precision	Not Null	Key
NUMBER_ORDERED	Integer	NA		
DISCOUNT	Money	default		
COMMENTS	Varchar	default		

The datatypes may vary, depending on the database you choose.

Note: For F_PROMO_ITEMS, you include foreign key columns that correspond to the primary keys in each of the dimension tables.

7. Click Repository > Save.

The next step is to generate and execute the SQL script to create each of these new target tables.

To create the tables:

1. Select all the table definitions.
2. Click Targets > Generate/Execute SQL.
3. In the Database Object Generation dialog box, connect to the target database.
4. Select Generate from Selected Tables, and select the options for creating the tables and generating keys.
5. Click Generate and Execute.
6. Click Close.

Creating the Mapping

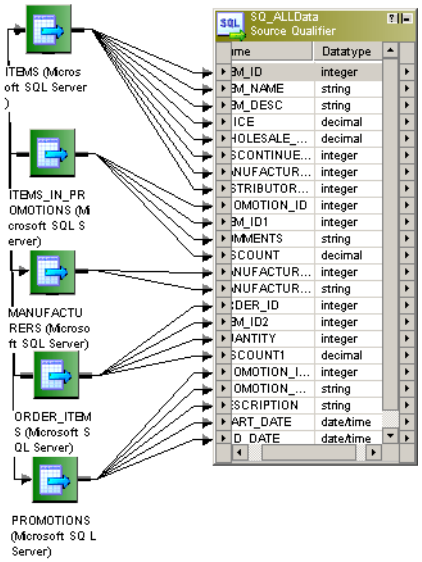
Create a mapping to filter out discontinued items, call a stored procedure to find how many of each item customers have ordered, and generate a unique ID for each row in the fact table.

To create the new mapping:

1. In the Designer, switch to the Mapping Designer and create a new mapping.
2. Name the mapping m_PromoItems.
3. From the list of targets, select the tables you just created and drag them into the mapping.
4. From the list of sources, add the following sources to the mapping:
   - PROMOTIONS
   - ITEMS_IN_PROMOTIONS
   - ITEMS
   - MANUFACTURERS
   - ORDER_ITEMS

5. Delete all Source Qualifier transformations that the Designer creates when you add these source definitions.

6. Add a Source Qualifier transformation named SQ_AllData to the mapping, and connect all the sources to it.



When you create a single Source Qualifier transformation, the Integration Service increases performance with a single read on the source database instead of multiple reads.

7. Click View > Navigator to close the Navigator window to allow extra space in the workspace.

8. Click Repository > Save.

Creating a Filter Transformation

The Filter transformation filters rows from a source. If you connect a Filter transformation to a Source Qualifier transformation, you can filter rows passed through the Source Qualifier transformation using any condition you want to apply. In this exercise, you remove discontinued items from the mapping.

The mapping contains a Filter transformation that limits rows queried from the ITEMS table to those items that have not been discontinued.

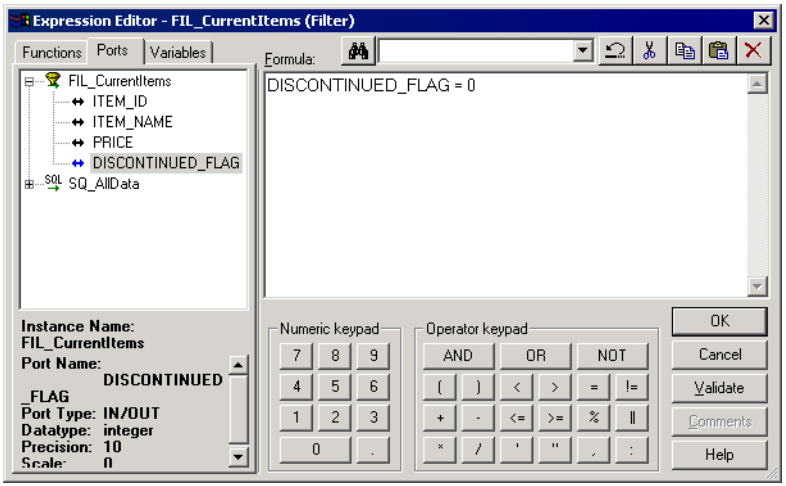
To add the Filter transformation:

1. Create a Filter transformation and name it FIL_CurrentItems.
2. Drag the following ports from the Source Qualifier transformation into the Filter transformation:
   - ITEM_ID
   - ITEM_NAME
   - PRICE
   - DISCONTINUED_FLAG
3. Open the Filter transformation.
4. Click the Properties tab to specify the filter condition.
5. Click the Open button in the Filter Condition field.

   The Expression Editor dialog box appears.
6. Select the word TRUE in the Formula field and press Delete.
7. Click the Ports tab.
8. Enter DISCONTINUED_FLAG = 0.

The following example shows the complete condition:

```
DISCONTINUED_FLAG = 0
```



9. Click Validate, and then click OK.

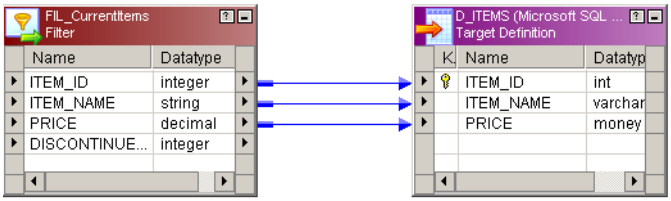
The new filter condition now appears in the Value field.

10. Click OK to return to the workspace.

Now, you need to connect the Filter transformation to the D_ITEMS target table. Currently sold items are written to this target.

To connect the Filter transformation:

1. Connect the ports ITEM_ID, ITEM_NAME, and PRICE to the corresponding columns in D_ITEMS.



2. Click Repository > Save.

Creating a Sequence Generator Transformation

A Sequence Generator transformation generates unique values, such as primary keys, for a target in a mapping. You can also use it to cycle through a closed set of values. Many relational databases include sequences, which are special database objects that generate values. The Sequence Generator transformation functions like a sequence object in a database. However, in PowerCenter, you do not need to write SQL code to create and use the sequence in a mapping.

The Sequence Generator transformation has the following properties:

- The starting number (normally 1).
- The current value stored in the repository.
- The number that the Sequence Generator transformation adds to its current value for every request for a new ID.
- The maximum value in the sequence.
- A flag indicating whether the Sequence Generator transformation counter resets to the minimum value once it has reached its maximum value.

The Sequence Generator transformation has two output ports, NEXTVAL and CURRVAL, which correspond to the two pseudo-columns in a sequence. When you query a value from the NEXTVAL port, the transformation generates a new value.

In the new mapping, you add a Sequence Generator transformation to generate IDs for the fact table F_PROMO_ITEMS. Every time the Integration Service inserts a new row into the target table, it generates a unique ID for PROMO_ITEM_ID.

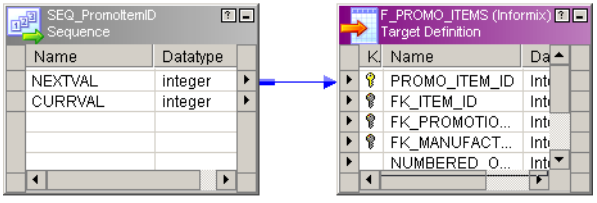
To add the Sequence Generator transformation:

1. Create a Sequence Generator transformation and name it SEQ_PromoItemID.
2. Open the Sequence Generator transformation.
3. Click the Ports tab.

   The two output ports, NEXTVAL and CURRVAL, appear in the list.

   Note: You cannot add any new ports to this transformation or reconfigure NEXTVAL and CURRVAL.
4. Click the Properties tab.

   The properties for the Sequence Generator transformation appear. You do not have to change any of these settings.
5. Click OK.
6. Connect the NEXTVAL column from the Sequence Generator transformation to the PROMO_ITEM_ID column in the target table F_PROMO_ITEMS.



7. Click Repository > Save.

Creating a Stored Procedure Transformation

When you installed the sample database objects to create the source tables, you also created a stored procedure, SP_GET_ITEM_COUNT. This procedure takes one argument, an ITEM_ID value, and returns the number of times that item has been ordered.

Table 6-1 describes the syntax for the stored procedure:

Table 6-1. Stored Procedure Syntax

Database	Syntax

Oracle	`CREATE FUNCTION SP_GET_ITEM_COUNT`
	`(ARG_ITEM_ID IN NUMBER) RETURN NUMBER`
	`IS SP_RESULT NUMBER;`
	`BEGIN`
	`SELECT COUNT(*)`
	`INTO SP_RESULT`
	`FROM ORDER_ITEMS`
	`WHERE ITEM_ID = ARG_ITEM_ID;`
	`RETURN (SP_RESULT);`
	`END;`

Microsoft SQL Server	`CREATE PROCEDURE SP_GET_ITEM_COUNT (@ITEM_ID INT)`
	`AS SELECT COUNT(*) FROM ORDER_ITEMS`
	`WHERE ITEM_ID = @ITEM_ID`

Sybase ASE	`CREATE PROCEDURE SP_GET_ITEM_COUNT (@ITEM_ID INT)` `AS SELECT COUNT(*) FROM ORDER_ITEMS WHERE ITEM_ID = @ITEM_ID`

Informix	`CREATE PROCEDURE SP_GET_ITEM_COUNT (ITEM_ID_INPUT INT)`
	`RETURNING INT;`
	`DEFINE CNT INT;`
	`SELECT COUNT(*) INTO CNT FROM ORDER_ITEMS WHERE ITEM_ID =`
	`ITEM_ID_INPUT;`
	`RETURN CNT;`

DB2	`CREATE PROCEDURE SP_GET_ITEM_COUNT (IN ARG_ITEM_ID INT,`
	`OUT SP_RESULT INT,`
	`OUT SQLCODE_OUT INT )`
	`LANGUAGE SQL`
	`P1: BEGIN`
	`-- Declare variables`
	`DECLARE SQLCODE INT DEFAULT 0;`
	`-- Declare handler`
	`DECLARE EXIT HANDLER FOR SQLEXCEPTION`
	`SET SQLCODE_OUT = SQLCODE;`
	`SELECT COUNT(*) INTO SP_RESULT`
	`FROM ORDER_ITEMS`
	`WHERE ITEM_ID=ARG_ITEM_ID;`
	`SET SQLCODE_OUT = SQLCODE;`
	`END P1`

Teradata	`CREATE PROCEDURE SP_GET_ITEM_COUNT`
	`(IN ARG_ITEM_ID integer, OUT SP_RESULT integer)`
	`BEGIN`
	`SELECT COUNT(*)`
	`INTO: SP_RESULT`
	`FROM ORDER_ITEMS`
	`WHERE ITEM_ID =: ARG_ITEM_ID;`
	`END;`

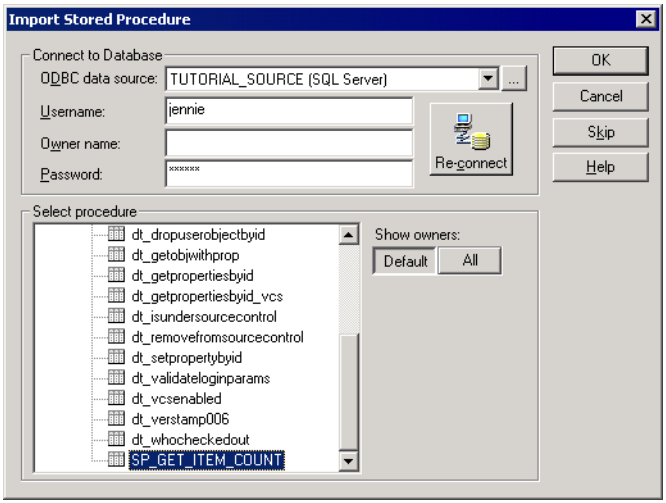
In the mapping, add a Stored Procedure transformation to call this procedure. The Stored Procedure transformation returns the number of orders containing an item to an output port.

To create the Stored Procedure transformation:

1. Create a Stored Procedure transformation and name it SP_GET_ITEM_COUNT.

   The Import Stored Procedure dialog box appears.

2. Select the ODBC connection for the source database. Enter a user name, owner name, and password. Click Connect.



3. Select the stored procedure named SP_GET_ITEM_COUNT from the list and click OK.

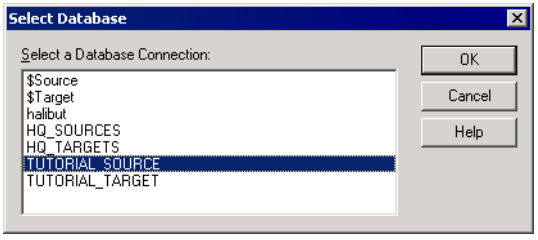
4. In the Create Transformation dialog box, click Done.

   The Stored Procedure transformation appears in the mapping.

5. Open the Stored Procedure transformation, and click the Properties tab.

6. Click the Open button in the Connection Information section.

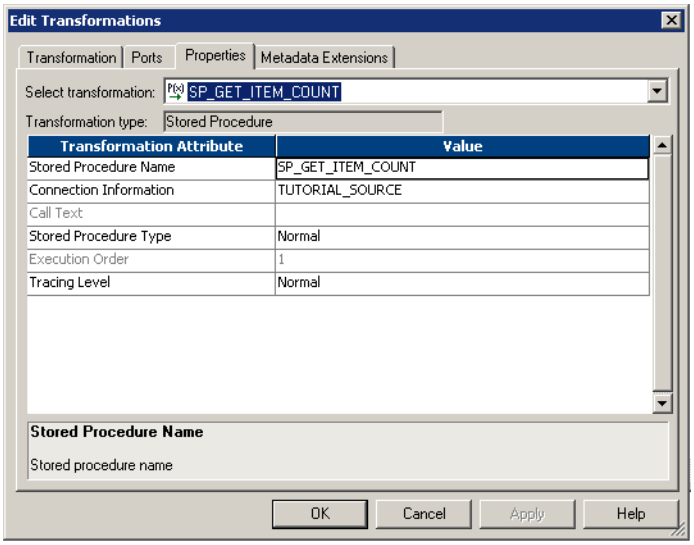
   The Select Database dialog box appears.



7. Select the source database and click OK.

   You can call stored procedures in both source and target databases.

Note: You can also select the built-in database connection variable, $Source. When you use $Source or $Target, the Integration Service determines which source database connection to use when it runs the session. If it cannot determine which connection to use, it fails the session. For more information about using $Source and $Target in a Stored Procedure transformation, see "Stored Procedure Transformation" in the *Transformation Guide*.



8. Click OK.
9. Connect the ITEM_ID column from the Source Qualifier transformation to the ITEM_ID column in the Stored Procedure transformation.
10. Connect the RETURN_VALUE column from the Stored Procedure transformation to the NUMBER_ORDERED column in the target table F_PROMO_ITEMS.
11. Click Repository > Save.

Completing the Mapping

The final step is to map data to the remaining columns in targets.

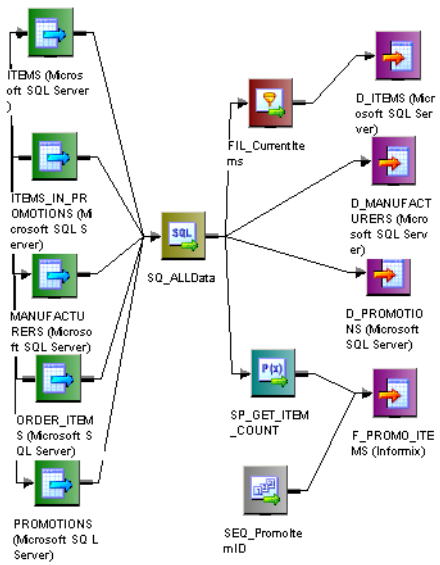
To complete the mapping:

1. Connect the following columns from the Source Qualifier transformation to the targets:

Source Qualifier	Target Table	Column
PROMOTION_ID	D_PROMOTIONS	PROMOTION_ID
PROMOTION_NAME	D_PROMOTIONS	PROMOTION_NAME
DESCRIPTION	D_PROMOTIONS	DESCRIPTION
START_DATE	D_PROMOTIONS	START_DATE
END_DATE	D_PROMOTIONS	END_DATE
MANUFACTURER_ID	D_MANUFACTURERS	MANUFACTURER_ID
MANUFACTURER_NAME	D_MANUFACTURERS	MANUFACTURER_NAME

2. Click Repository > Save.

The mapping is now complete. You can create and run a new workflow with this mapping.



Creating a Workflow

In this part of the lesson, you complete the following steps:

1. Create a workflow.
2. Add a non-reusable session to the workflow.
3. Define a link condition before the Session task.

Creating the Workflow

Open the Workflow Manager and connect to the repository.

To create a workflow:

1. Click Tools > Workflow Designer.
2. Click Workflows > Create to create a new workflow.

   When you have a workflow open, the Workflow Manager prompts you to close to the current workflow. Click Yes to close any current workflow.

   The workflow properties appear.
3. Name the workflow wf_PromoItems.
4. Click the Browse Integration Service button to select the Integration Service to run the workflow.

   The Integration Service Browser dialog box appears.
5. Select the Integration Service you use and click OK.
6. Click the Scheduler tab.

   By default, the workflow is scheduled to run on demand. Keep this default.
7. Click OK to close the Create Workflow dialog box.

   The Workflow Manager creates a new workflow in the workspace including the Start task.

Next, you create a non-reusable session in the workflow.

To create the non-reusable session:

1. Click Tasks > Create.

   The Create Task dialog box appears. The Workflow Designer provides more task types than the Task Developer. These tasks include the Email and Decision tasks.
2. Create a Session task and name it s_PromoItems. Click Create.

   In the Mappings dialog box, select the mapping m_PromoItems and click OK.
3. Click Done.

4. Open the session properties for s_PromoItems.
5. Click the Mapping tab.
6. Select the source database connection for the sources connected to the SQ_AllData Source Qualifier transformation.
7. Select the target database for each target definition.
8. Click OK to save the changes.
9. Click the Link Tasks button on the toolbar.
10. Drag from the Start task to s_PromoItems.
11. Click Repository > Save to save the workflow in the repository.

You can now create a link condition in the workflow.

Creating a Link Condition

After you create links between tasks, you can specify conditions for each link to determine the order of execution in the workflow. If you do not specify conditions for each link, the Integration Service executes the next task in the workflow by default.

If the link condition evaluates to True, the Integration Service runs the next task in the workflow. The Integration Service does not run the next task in the workflow if the link condition evaluates to False. You can also use pre-defined or user-defined workflow variables in the link condition.

You can add comments using -- or // comment indicators with the Expression Editor. Use comments to describe the expression.

You can view results of link evaluation during workflow runs in the workflow log.

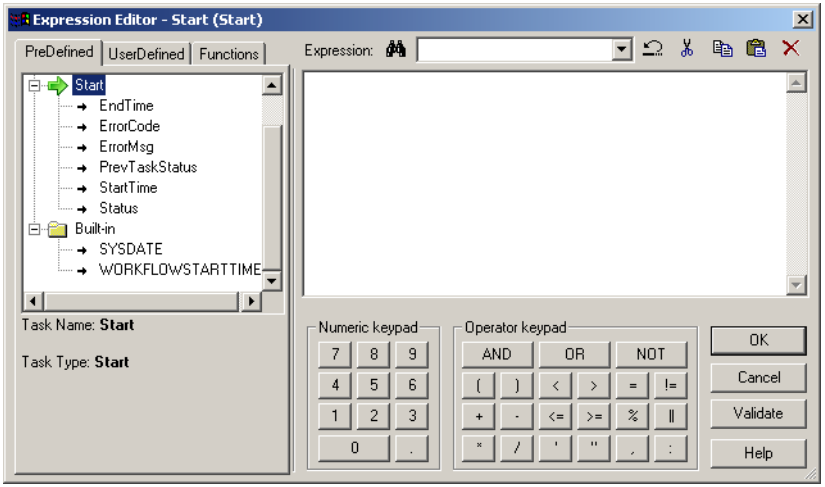
For more information about working with links or workflow variables, see "Working with Workflows" in the *Workflow Administration Guide*.

In the following steps, you create a link condition before the Session task and use the built-in workflow variable WORKFLOWSTARTTIME. You define the link condition so the Integration Service runs the session if the workflow start time is before the date you specify.

To define a link condition:

1. Double-click the link from the Start task to the Session task.

   The Expression Editor appears.



2. Expand the Built-in node on the PreDefined tab.

   The Workflow Manager displays the two built-in workflow variables, SYSDATE and WORKFLOWSTARTTIME.

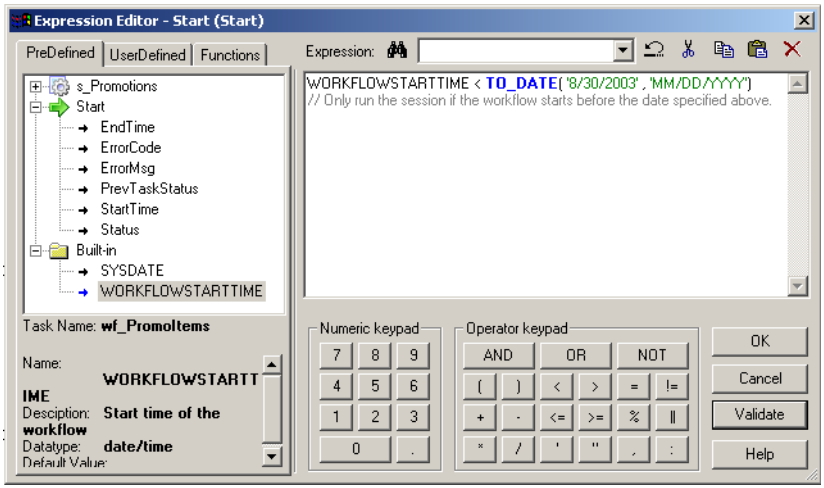
3. Enter the following expression in the expression window. Be sure to enter a date later than today's date:

   ```
   WORKFLOWSTARTTIME < TO_DATE('8/30/2003', 'MM/DD/YYYY')
   ```

   Tip: You can double-click the built-in workflow variable on the PreDefined tab and double-click the TO_DATE function on the Functions tab to enter the expression. For more information about using functions in the Expression Editor, see "The Transformation Language" in the *Transformation Language Reference*.

4. Press Enter to create a new line in the Expression. Add a comment by typing the following text:

```
// Only run the session if the workflow starts before the date specified
above.
```

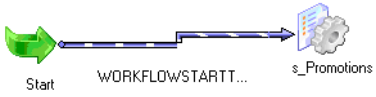


5. Validate the expression using the Validate button.

   The Workflow Manager displays a message in the Output window.

6. Click OK.

   After you specify the link condition in the Expression Editor, the Workflow Manager validates the link condition and displays it next to the link in the workflow.



7. Click Repository > Save to save the workflow in the repository.

Next, you run and monitor the workflow.

Running the Workflow

After you create the workflow, you can run it and use the Workflow Monitor to monitor the workflow progress.

To run the workflow:

1. Right-click the workflow in the workspace and select Start Workflow.

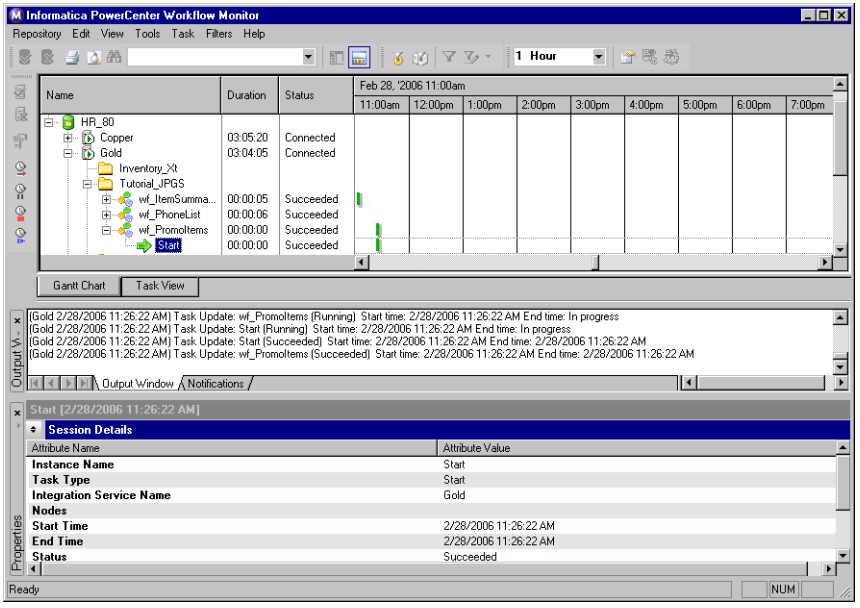
   Tip: You can also right-click the workflow in the Navigator and select Start Workflow.

   The Workflow Monitor opens and connects to the repository and opens the tutorial folder.

2. Click the Gantt Chart tab at the bottom of the Time window to verify the Workflow Monitor is in Gantt Chart view.

3. In the Navigator, expand the node for the workflow.

   All tasks in the workflow appear in the Navigator.



The results from running the s_PromoItems session are as follows:

```
F_PROMO_ITEMS 40 rows inserted

D_ITEMS 13 rows inserted

D_MANUFACTURERS 11 rows inserted

D_PROMOTIONS 3 rows inserted
```

Chapter 7

Tutorial Lesson 6

This chapter includes the following topics:

Using XML Files

XML is a common means of exchanging data on the web. Use XML files as a source of data and as a target for transformed data.

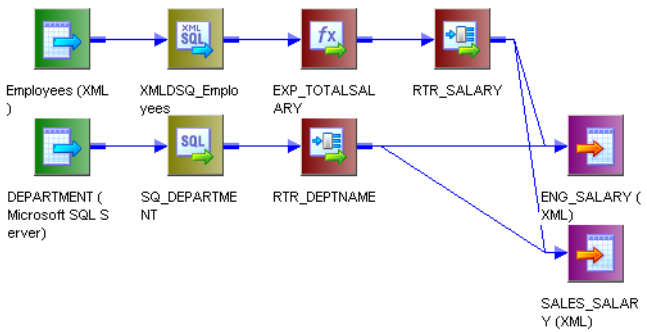
In this lesson, you have an XML schema file that contains data on the salary of employees in different departments, and you have relational data that contains information about the different departments. You want to find out the total salary for employees in two departments, and you want to write the data to a separate XML target for each department.

In the XML schema file, employees can have three types of wages, which appear in the XML schema file as three occurrences of salary. You pivot the occurrences of employee salaries into three columns: BASESALARY, COMMISSION, and BONUS. Then, you calculate the total salary in an Expression transformation.

You use a Router transformation to test for the department ID. You use another Router transformation to get the department name from the relational source. You send the salary data for the employees in the Engineering department to one XML target and the salary data for the employees in the Sales department to another XML target.

Figure 7-1 shows the mapping you create in this lesson:

Figure 7-1. Mapping with XML Sources and Targets



Creating the XML Source

You use the XML Wizard to import an XML source definition, and then you use the XML Editor to edit the definition.

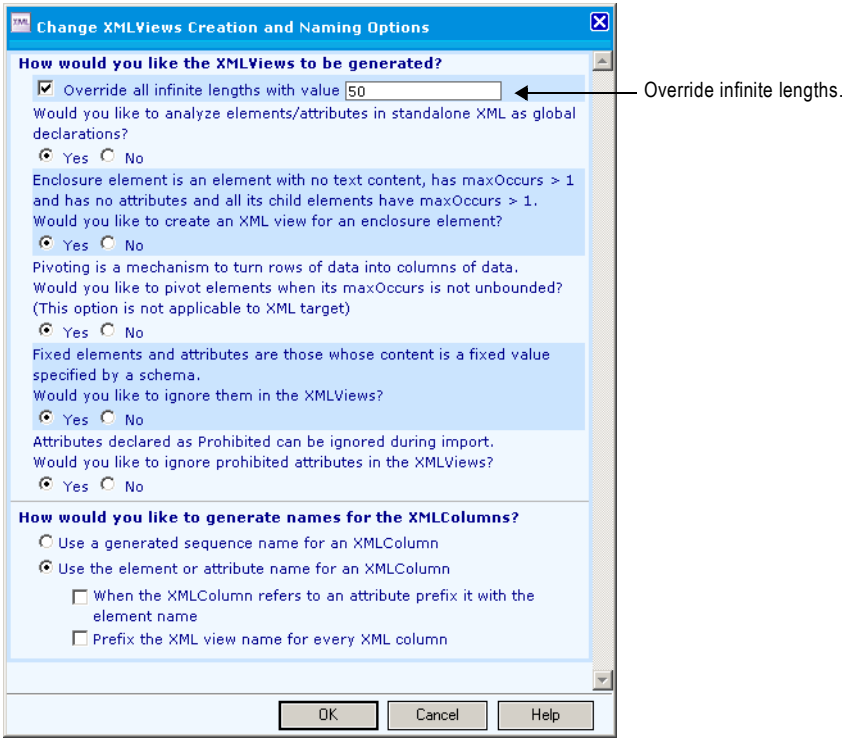
Importing the XML Source

Use the Employees.xsd file to create the XML source definition.

To import the XML source definition:

1. Open the Designer if it is not already open, connect to the repository, and open the tutorial folder.
2. Open the Source Analyzer.
3. Click Sources > Import XML Definition.
4. Click Advanced Options.

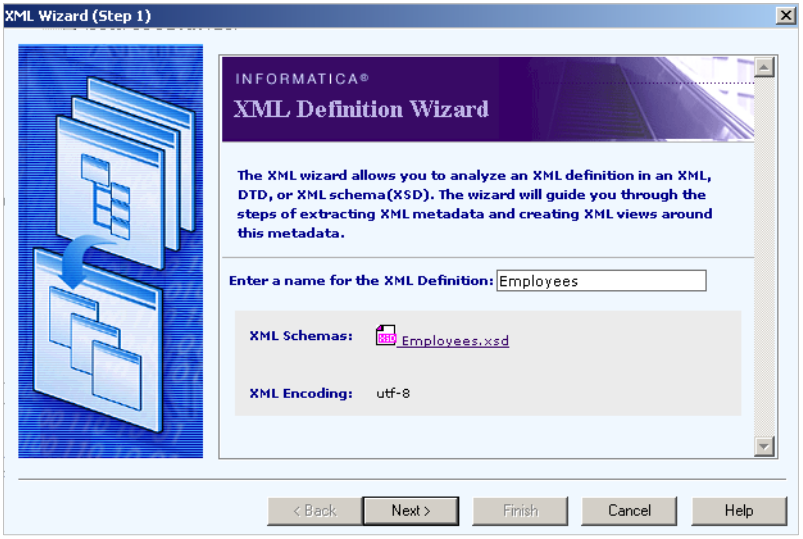
   The Change XML Views Creation and Naming Options dialog box opens.



5. Select Override All Infinite Lengths and enter 50.
6. Configure all other options as shown and click OK to save the changes.

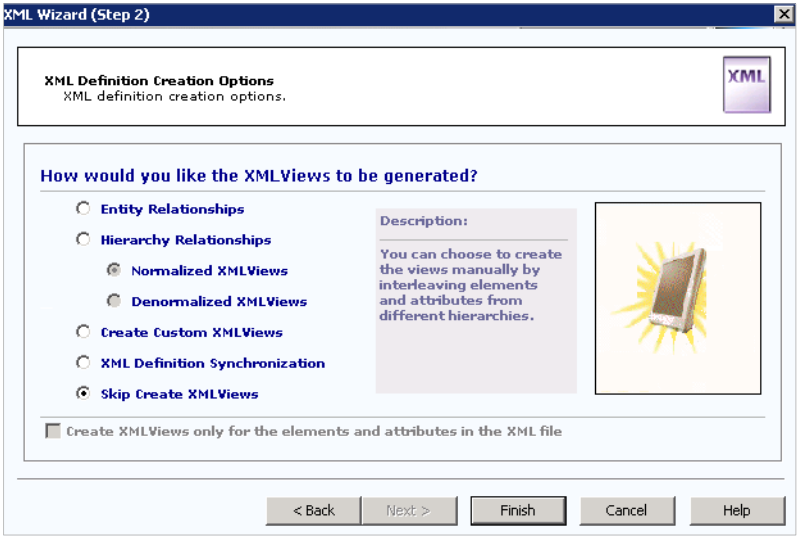
7. In the Import XML Definition dialog box, navigate to the client\bin directory under the PowerCenter installation directory and select the Employees.xsd file. Click Open.

   The XML Definition Wizard opens.



8. Verify that the name for the XML definition is Employees, and click Next.

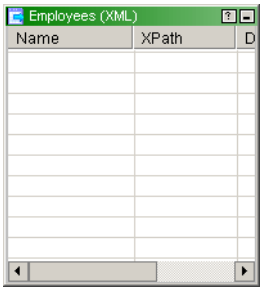
9. Select Skip Create XML Views.



   Because you only need to work with a few elements and attributes in the Employees.xsd file, you skip creating a definition using the XML Wizard. Instead, you create a custom view in the XML Editor. With a custom view in the XML Editor, you can exclude the elements and attributes that you do not need in the mapping.

10. Click Finish to create the XML definition.

    The XML Wizard creates an XML definition with no columns or groups.

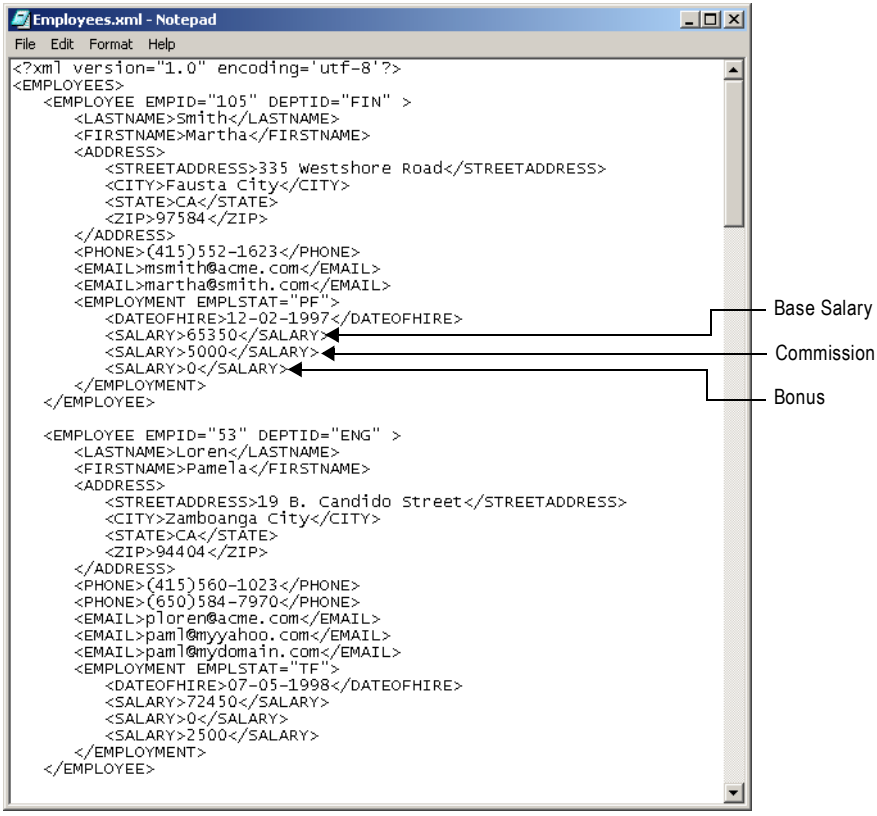


When you skip creating XML views, the Designer imports metadata into the repository, but it does not create the XML view. In the next step, you use the XML Editor to add groups and columns to the XML view.

Editing the XML Definition

The Designer represents an XML hierarchy in an XML definition as a set of views. Each view represents a subset of the XML hierarchy. A view consists of columns and rows. Columns represent elements and attributes, and rows represent occurrences of elements. You use the XML Editor to edit the XML views.

In this lesson, you use the XML Editor to pivot the three occurrences of SALARY into three columns in an XML group. You do this because the multiple-occurring element SALARY represents three types of salary: a base salary, a commission, and a bonus that appear in the XML file as three instances of the SALARY element.

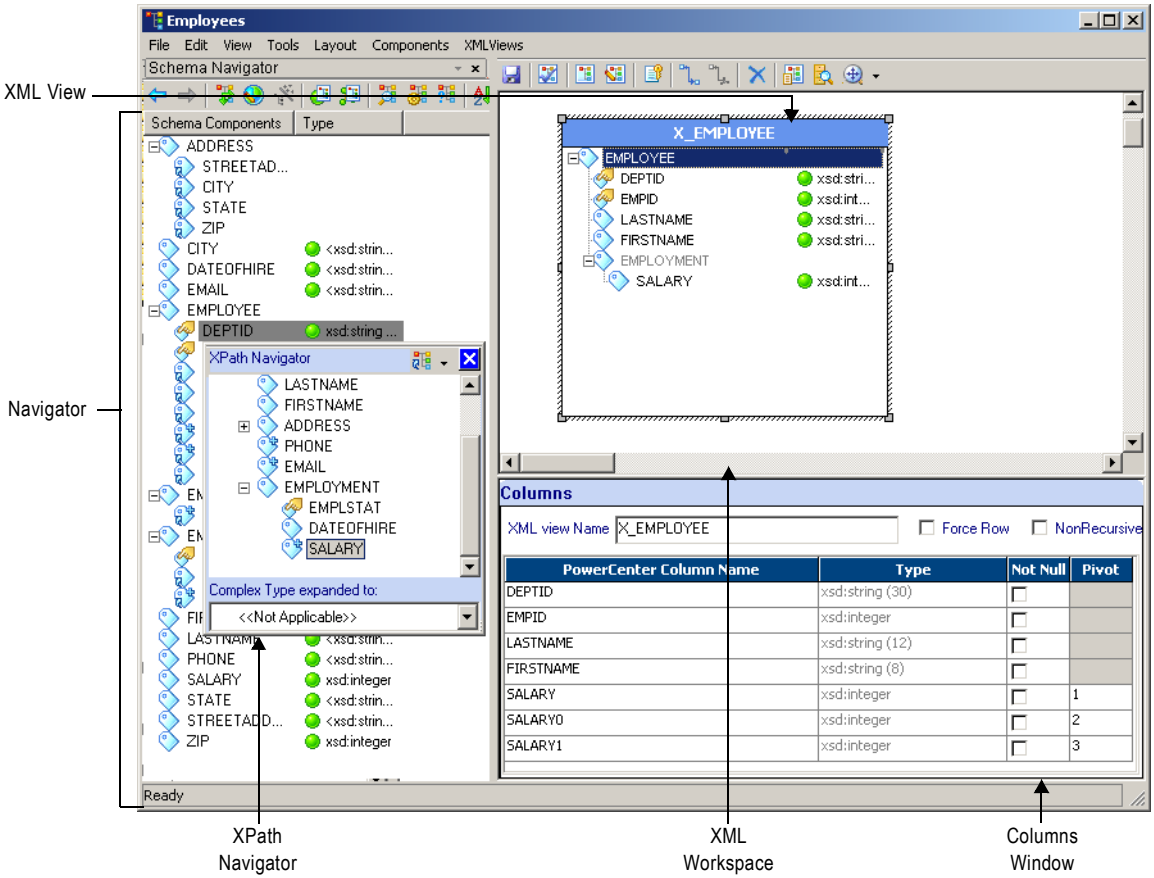


To work with these three instances separately, you pivot them to create three separate columns in the XML definition.

You create a custom XML view with columns from several groups. You then pivot the occurrence of SALARY to create the columns, BASESALARY, COMMISSION, and BONUS.

Figure 7-2 shows the XML Editor:

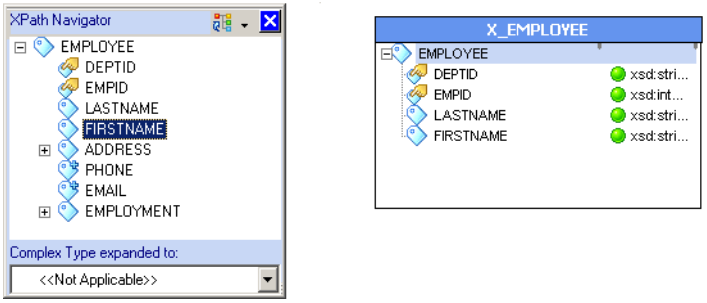
Figure 7-2. XML Editor



To edit the XML definition:

1. Double-click the XML definition or right-click the XML definition and select Edit XML Definition to open the XML Editor.
2. Click XMLViews > Create XML View to create a new XML view.
3. From the EMPLOYEE group, select DEPTID and right-click it.
4. Choose Show XPath Navigator.
5. From the XPath Navigator, select the following elements and attributes and drag them into the new view:
   - DEPTID
   - EMPID
   - LASTNAME

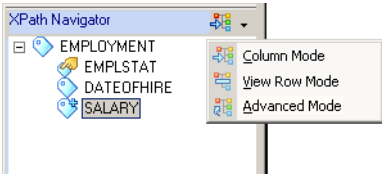
- FIRSTNAME



The XML Editor names the view X_EMPLOYEE.

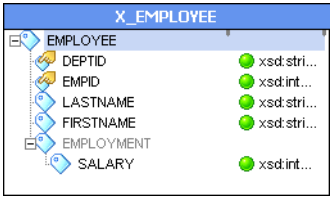
Note: The XML Wizard may transpose the order of the DEPTID and EMPID attributes when it imports them. If this occurs, you can add the columns in the order they appear in the Schema Navigator or XPath Navigator. Transposing the order of attributes does not affect data consistency.

6. Expand the EMPLOYMENT group so that the SALARY column appears.
7. Click the Mode icon on the XPath Navigator, and choose Advanced Mode.



8. Select the SALARY column, and drag it into the XML view.

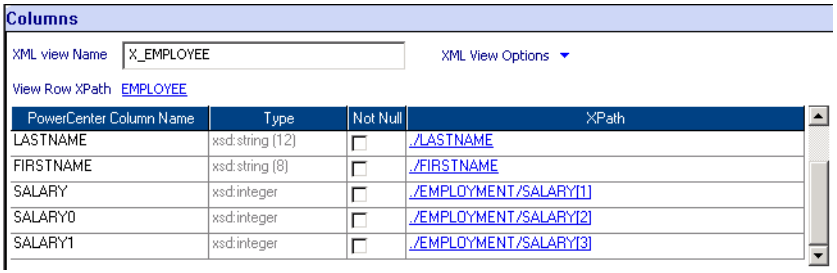
   The resulting view includes the elements and attributes shown in the following view.



9. Drag the SALARY column into the new XML view two more times to create three pivoted columns.

   Note: Although the new columns appear in the column window, the view shows one instance of SALARY.

The wizard adds three new columns in the column view and names them SALARY, SALARY0, and SALARY1.



10. Rename the new columns.

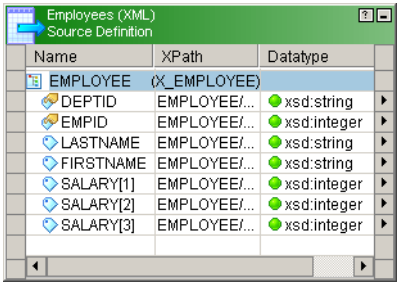
Use information on the following table to modify the name and pivot properties:

Column Name	New Column Name	Not Null	Pivot Occurrence
SALARY	BASESALARY	Yes	1
SALARY0	COMMISSION		2
SALARY1	BONUS		3

11. Click File > Apply Changes to save the changes to the view.

12. Click File > Exit to close the XML Editor.

The following source definition appears in the Source Analyzer.



Note: The pivoted SALARY columns do not display the names you entered in the Columns window. However, when you drag the ports to another transformation, the edited column names appear in the transformation.

13. Click Repository > Save to save the changes to the XML definition.

Creating the Target

In this lesson, you import an XML schema and create a custom view based on the schema. The custom XML target definition you create meets the following criteria:

- Each department has a separate target, and the structure for each target is the same.
- Each target contains salary and department information for employees in the Sales or Engineering department.

Because the structure for the target data is the same for the Engineering and Sales groups, use two instances of the target in the mapping. In the following steps, you import the Sales_Salary schema file and create a custom view based on the schema.

To import and edit the XML target:

1. In the Designer, switch to the Target Designer.

   If the workspace contains targets from other lessons, right-click the workspace and choose Clear All.

2. Click Targets > Import XML Definition.
3. Navigate to the Tutorial directory in the PowerCenter installation directory, and select the Sales_Salary.xsd file. Click Open.

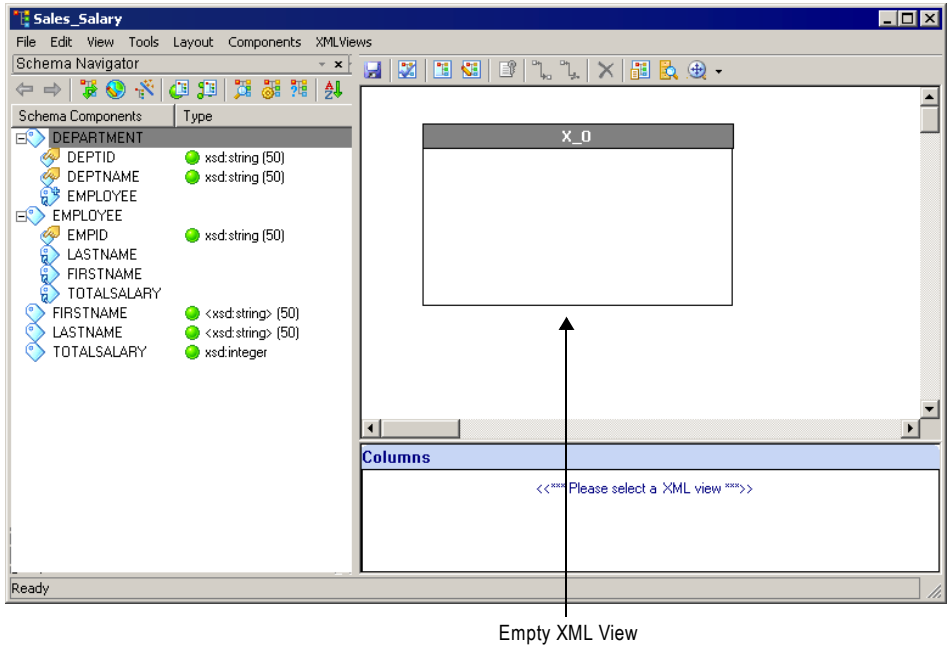
   The XML Definition Wizard appears.

4. Name the XML definition SALES_SALARY and click Next.
5. Select Skip Create XML Views, and click Finish.

   The XML Wizard creates the SALES_SALARY target with no columns or groups.

6. Double-click the XML definition to open the XML Editor.
7. Click XMLViews > Create XML View.

The XML Editor creates an empty view.



8. Right-click DEPARTMENT group in the Schema Navigator and select Show XPath Navigator.

9. From the XPath Navigator, drag DEPTNAME and DEPTID into the empty XML view.

   The XML Editor names the view X_DEPARTMENT.

   Note: The XML Editor may transpose the order of the attributes DEPTNAME and DEPTID. If this occurs, add the columns in the order they appear in the Schema Navigator. Transposing the order of attributes does not affect data consistency.

10. In the X_DEPARTMENT view, right-click the DEPTID column, and choose Set as Primary Key.

11. Click XMLViews > Create XML View.

    The XML Editor creates an empty view.

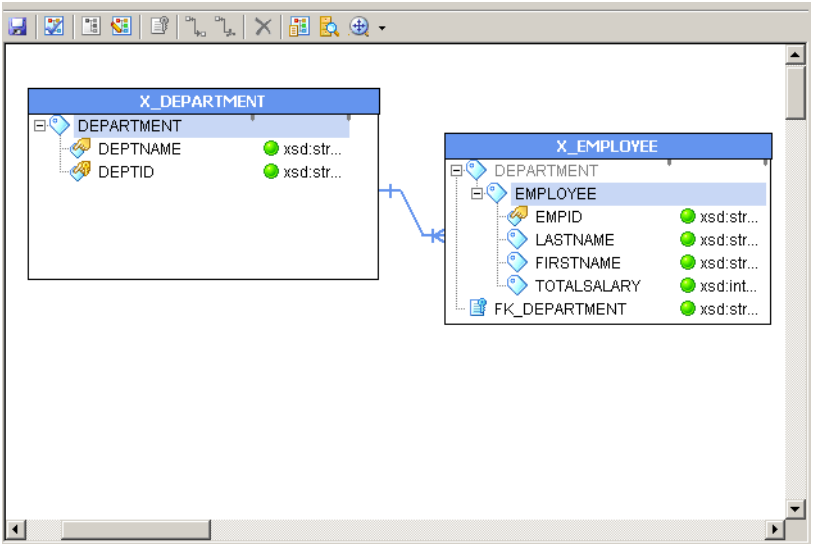
12. From the EMPLOYEE group in the Schema Navigator, open the XPath Navigator.

13. From the XPath Navigator, drag EMPID, FIRSTNAME, LASTNAME, and TOTALSALARY into the empty XML view.

    The XML Editor names the view X_EMPLOYEE.

14. Right-click the X_EMPLOYEE view, and choose Create Relationship. Drag the pointer from the X_EMPLOYEE view to the X_DEPARTMENT view to create a link.

15. The XML Editor creates a DEPARTMENT foreign key in the X_EMPLOYEE view that corresponds to the DEPTID primary key.



16. Click File > Apply Changes and close the XML Editor.

    The XML definition now contains the groups DEPARTMENT and EMPLOYEE.

17. Click Repository > Save to save the XML target definition.

Creating a Mapping with XML Sources and Targets

In the following steps, you create a mapping to transform the employee data. You add the following objects to the mapping:

- The Employees XML source definition you created
- The DEPARTMENT relational source definition you created in "Creating Source Definitions" on page 22
- Two instances of the SALES_SALARY target definition you created
- An Expression transformation to calculate the total salary for each employee
- Two Router transformations to route salary and department

You pass the data from the Employees source through the Expression and Router transformations before sending it to two target instances. You also pass data from the relational table through another Router transformation to add the department names to the targets. You need data for the sales and engineering departments.

To create the mapping:

1. In the Designer, switch to the Mapping Designer and create a new mapping.
2. Name the mapping m_EmployeeSalary.
3. Drag the Employees XML source definition into the mapping.
4. Drag the DEPARTMENT relational source definition into the mapping.

   By default, the Designer creates a source qualifier for each source.
5. Drag the SALES_SALARY target into the mapping two times.
6. Rename the second instance of SALES_SALARY as ENG_SALARY.
7. Click Repository > Save.

   Because you have not completed the mapping, the Designer displays a warning that the mapping m_EmployeeSalary is invalid.

Next, you add an Expression transformation and two Router transformations. Then, you connect the source definitions to the Expression transformation. You connect the pipeline to the Router transformations and then to the two target instances.

Creating an Expression Transformation

In the following steps, you use an Expression transformation to calculate the total salary for each employee. You use BASESALARY, COMMISSION, and BONUS as input columns to the Expression transformation and create a TotalSalary column as output.

To calculate the total salary:

1. Create an Expression transformation and name it EXP_TotalSalary.

   The new transformation appears.

2. Click Done.

3. Drag all the ports from the XML Source Qualifier transformation to the EXP_TotalSalary Expression transformation.

   The input/output ports in the XML Source Qualifier transformation are linked to the input/output ports in the Expression transformation.

4. Open the Expression transformation.

5. On the Ports tab, add an output port named TotalSalary. Use decimal datatype with precision of 10 and scale of 2.

6. Enter the following expression for TotalSalary:

   ```
   BASESALARY + COMMISSION + BONUS
   ```

7. Validate the expression and click OK.

8. Click OK to close the transformation.

9. Click Repository > Save.

Creating Router Transformations

A Router transformation tests data for one or more conditions and gives you the option to route rows of data that do not meet any of the conditions to a default output group.

In the following steps, you add two Router transformations to the mapping, one for each department. In each Router transformation you create two groups. One group returns True for rows where the DeptID column contains 'SLS'. The other group returns True where the DeptID column contains 'ENG'. All rows that do not meet either condition go into the default group.

To route the employee salary data:

1. Create a Router transformation and name it RTR_Salary. Then click Done.

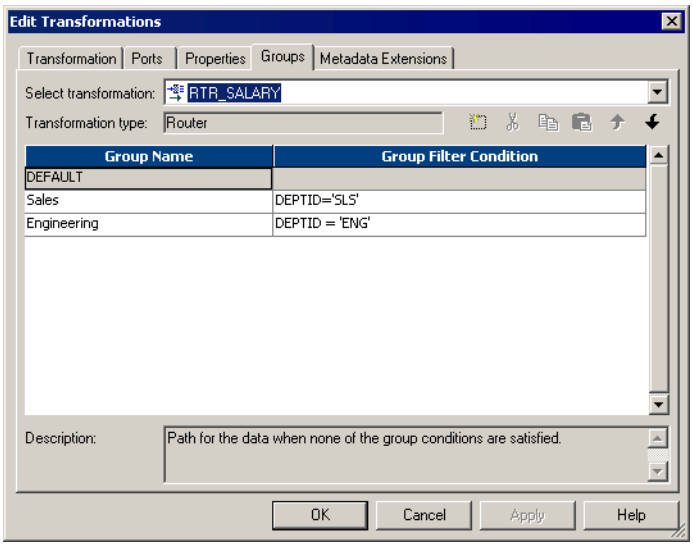
2. In the Expression transformation, select the following columns and drag them to RTR_Salary:

   - EmpID
   - DeptID
   - LastName
   - FirstName
   - TotalSalary

   The Designer creates an input group and adds the columns you drag from the Expression transformation.

3. Open the RTR_Salary Router transformation.

   The Edit Transformations dialog box appears.



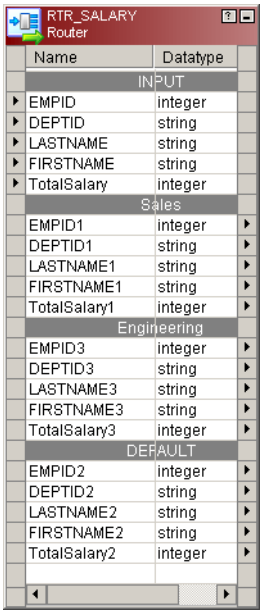
4. On the Groups tab, add two new groups. Change the group names and set the filter conditions. Use the following table as a guide:

Group Name	Filter Condition
Sales	DEPTID = ‘SLS’
Engineering	DEPTID = ‘ENG’

   The Designer adds a default group to the list of groups. All rows that do not meet the condition you specify in the group filter condition are routed to the default group. If you do not connect the default group, the Integration Service drops the rows.

5. Click OK to close the transformation.

6. In the workspace, expand the RTR_Salary Router transformation to see all groups and ports.



7. Click Repository > Save.

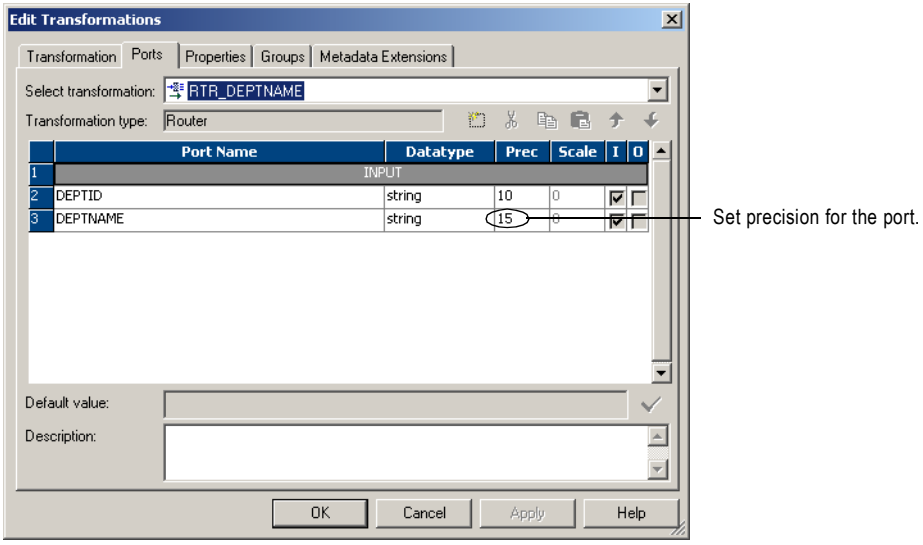
Next, you create another Router transformation to filter the Sales and Engineering department data from the DEPARTMENT relational source.

To route the department data:

1. Create a Router transformation and name it RTR_DeptName. Then click Done.
2. Drag the DeptID and DeptName ports from the DEPARTMENT Source Qualifier transformation to the RTR_DeptName Router transformation.
3. Open RTR_DeptName.
4. On the Groups tab, add two new groups. Change the group names and set the filter conditions using the following table as a guide:

Group Name	Filter Condition
Sales	DEPTID = 'SLS'
Engineering	DEPTID = 'ENG'

5. On the Ports tab, change the precision for DEPTNAME to 15.



6. Click OK to close the transformation.

7. In the workspace, expand the RTR_DeptName Router transformation to see all groups and columns.

8. Click Repository > Save.

Completing the Mapping

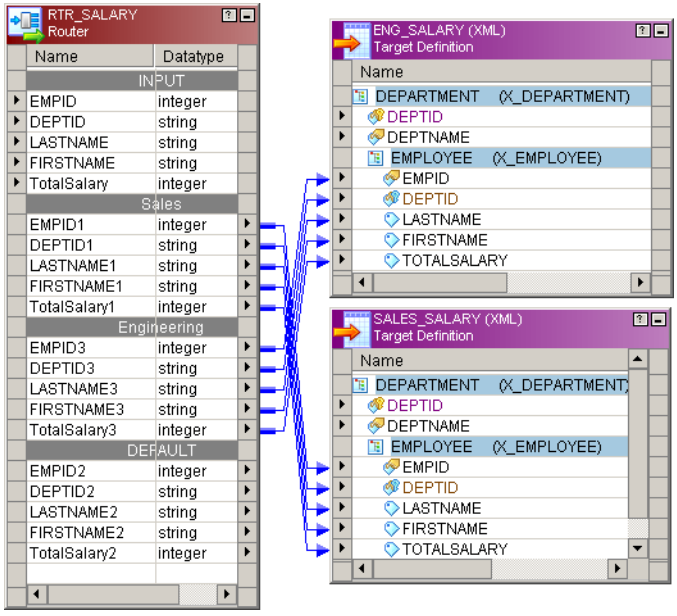
Connect the Router transformations to the targets to complete the mapping.

To complete the mapping:

1. Connect the following ports from RTR_Salary groups to the ports in the XML target definitions:

Router Group	Router Port	Target	Target Group	Target Port
Sales	EMPID1	SALES_SALARY	EMPLOYEE	EMPID
	DEPTID1			DEPTID (FK)
	LASTNAME1			LASTNAME
	FIRSTNAME1			FIRSTNAME
	TotalSalary1			TOTALSALARY
Engineering	EMPID3	ENG_SALARY	EMPLOYEE	EMPID
	DEPTID3			DEPTID (FK)
	LASTNAME3			LASTNAME
	FIRSTNAME3			FIRSTNAME
	TotalSalary3			TOTALSALARY

The mapping should now look like this:

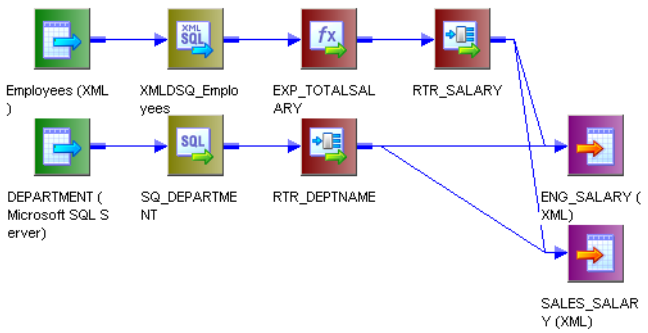


2. Connect the following ports from RTR_DeptName groups to the ports in the XML target definitions:

Router Group	Router Port	Target	Target Group	Target Port
Sales	DEPTID1	SALES_SALARY	DEPARTMENT	DEPTID
	DEPTNAME1			DEPTNAME
Engineering	DEPTID3	ENG_SALARY	DEPARTMENT	DEPTID
	DEPTNAME3			DEPTNAME

3. Click Repository > Save.

The mapping is now complete. When you save the mapping, the Designer displays a message that the mapping m_EmployeeSalary is valid.



Creating a Workflow

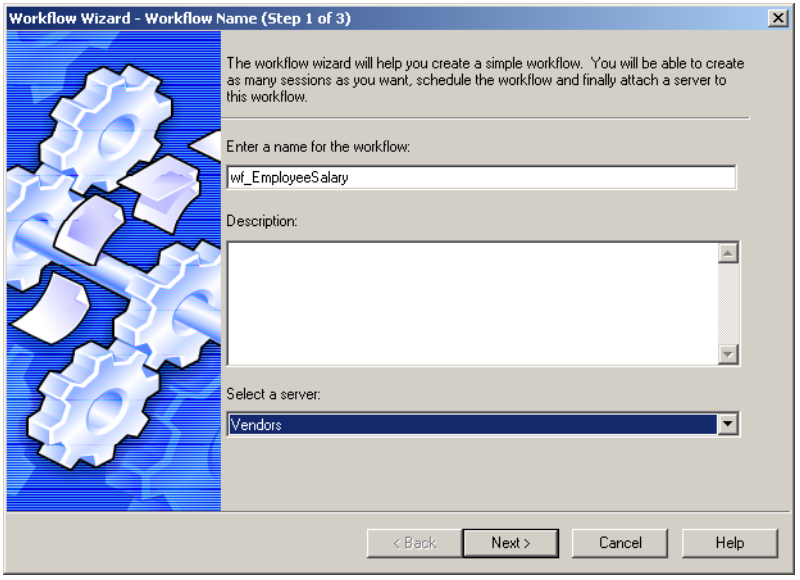
In the following steps, you create a workflow with a non-reusable session to run the mapping you just created.

Note: Before you run the workflow based on the XML mapping, you must verify that the Integration Service that runs the workflow can access the source XML file. Copy the Employees.xml file from the Tutorial folder to the $PMSourceFileDir directory for the Integration Service. Usually, this is the SrcFiles directory in the Integration Service installation directory.

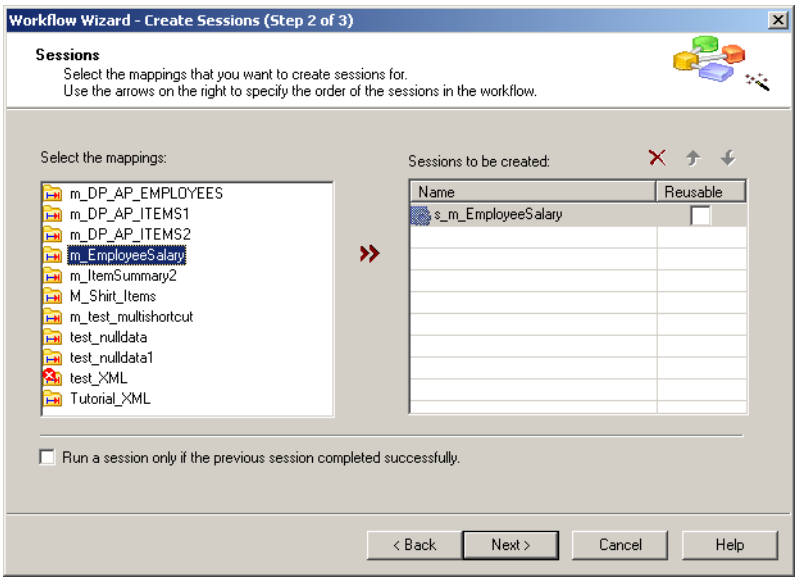
To create the workflow:

1. Open the Workflow Manager if it is not open already. Connect to the repository and open the tutorial folder.
2. Go to the Workflow Designer.
3. Click Workflows > Wizard.

   The Workflow Wizard opens.
4. Name the workflow wf_EmployeeSalary and select a service on which to run the workflow. Then click Next.



5. Select the m_EmployeeSalary mapping to create a session.



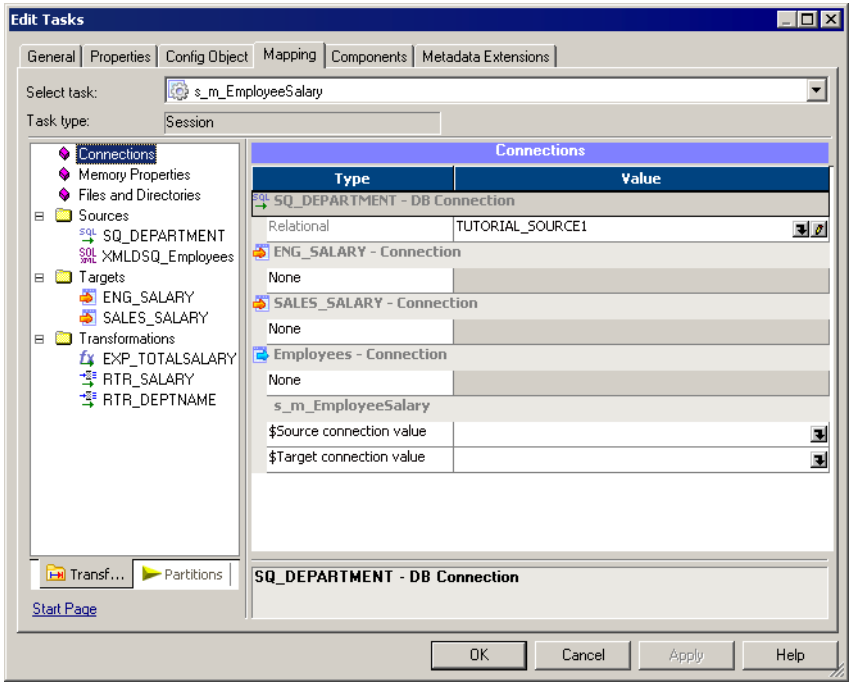
Note: The Workflow Wizard creates a session called s_m_EmployeeSalary.

6. Click Next.
7. Click Run on demand and click Next.

   The Workflow Wizard displays the settings you chose.
8. Click Finish to create the workflow.

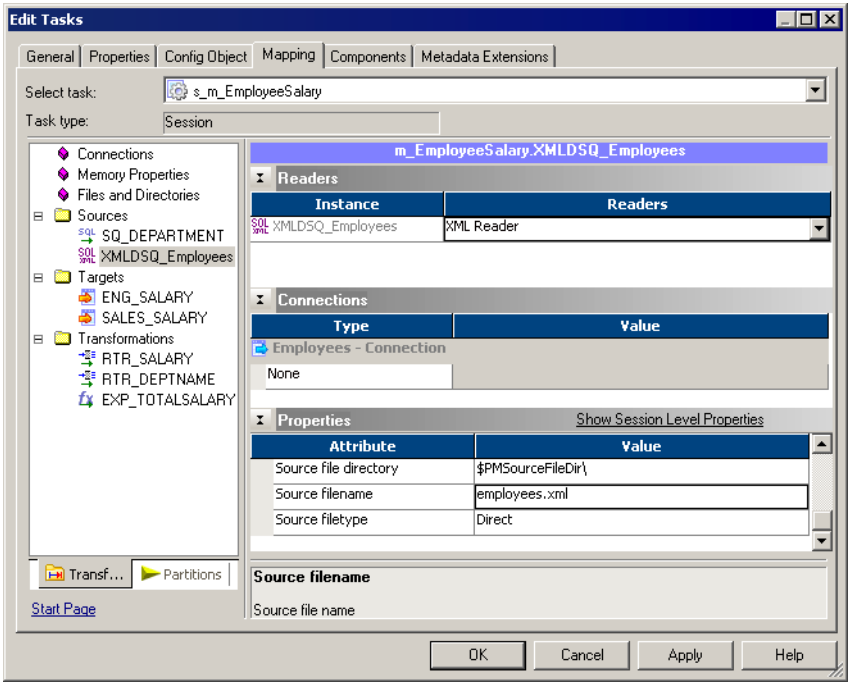
   The Workflow Wizard creates a Start task and session. You can add other tasks to the workflow later.
9. Click Repository > Save to save the new workflow.
10. Double-click the s_m_EmployeeSalary session to open it for editing.
11. Click the Mapping tab.

12. Select the connection for the SQ_DEPARTMENT Source Qualifier transformation.

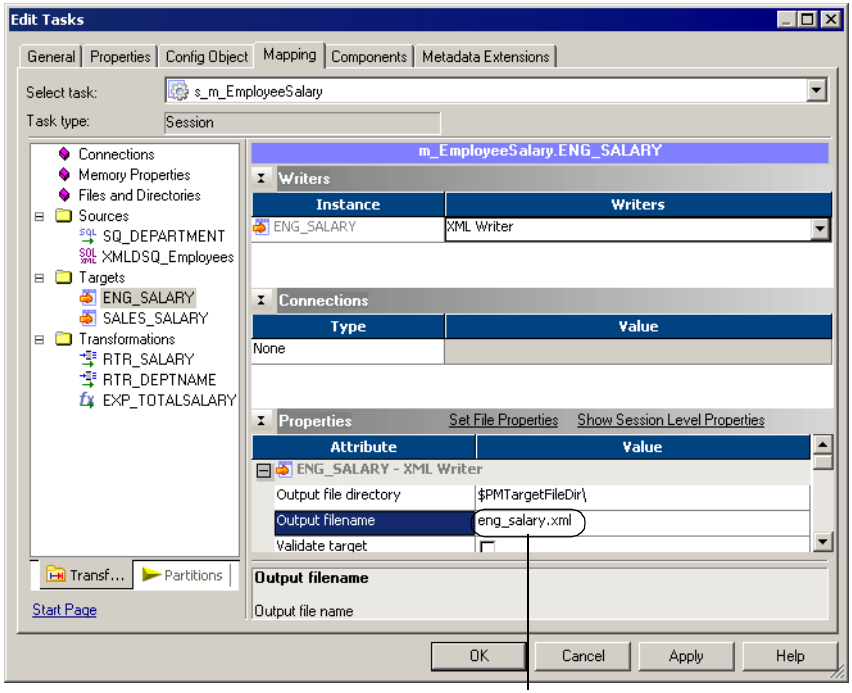


13. Select the XMLDSQ_Employees source on the Mapping tab.

14. Verify that the Employees.xml file is in the specified source file directory.



15. Click the ENG_SALARY target instance on the Mapping tab and verify that the output file name is eng_salary.xml.



Edit the output file name.

16. Click the SALES_SALARY target instance on the Mapping tab and verify that the output file name is sales_salary.xml.
17. Click OK to close the session.
18. Click Repository > Save.
19. Run and monitor the workflow.

    The Integration Service creates the eng_salary.xml and sales_salary.xml files.

Figure 7-3 shows the ENG_SALARY.XML file that the Integration Service creates when it runs the session:

Figure 7-3. ENG_SALARY.XML Output

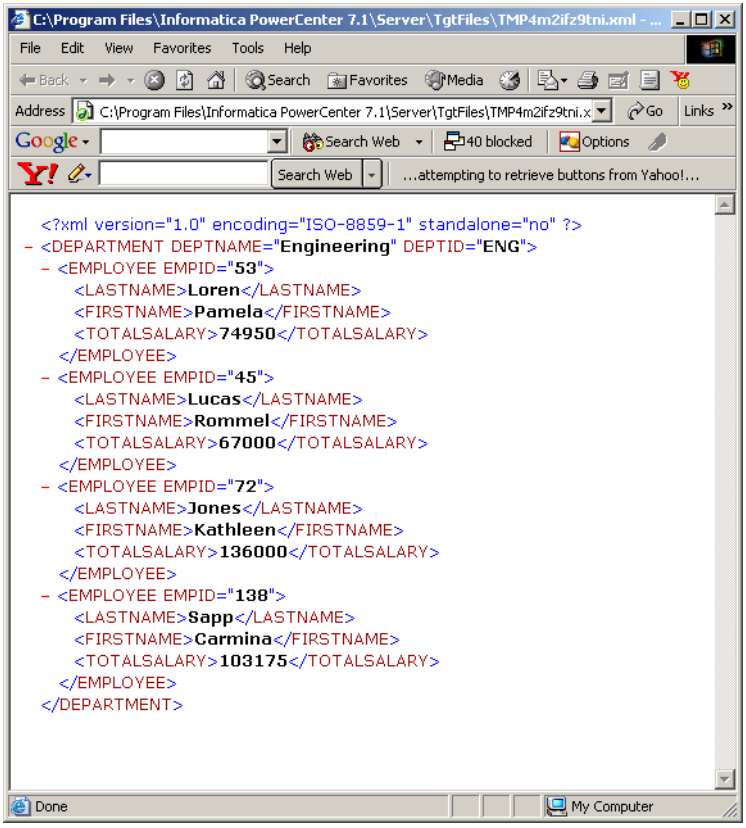
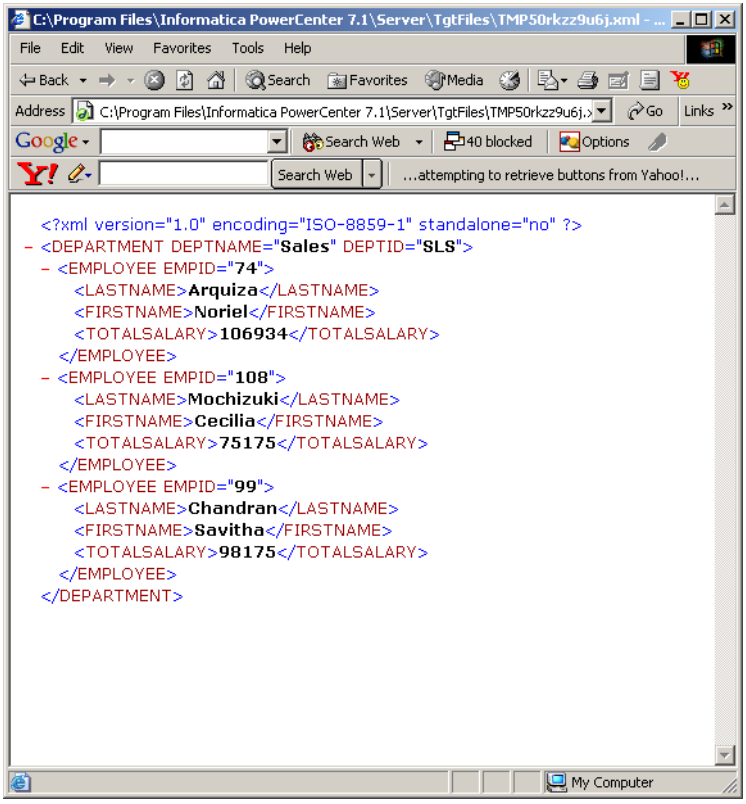


Figure 7-4 shows the SLS_SALARY.XML file that the Integration Service creates when it runs the session:

Figure 7-4. SLS_SALARY.XML Output



Appendix A

Naming Conventions

This appendix provides suggested naming conventions for PowerCenter repository objects.

Suggested Naming Conventions

The following naming conventions appear throughout the PowerCenter documentation and client tools. Use the following naming convention when you design mappings and create sessions.

Transformations

Table A-1 lists the recommended naming convention for all transformations:

Table A-1. Naming Conventions for Transformations

Transformation	Naming Convention
Aggregator	AGG_*TransformationName*
Application Source Qualifier	ASQ_*TransformationName*
Custom	CT_*TransformationName*
Expression	EXP_*TransformationName*
External Procedure	EXT_*TransformationName*
Filter	FIL_*TransformationName*
Java	JTX_*TransformationName*
Joiner	JNR_*TransformationName*
Lookup	LKP_*TransformationName*
MQ Source Qualifier	SQ_MQ_*TransformationName*
Normalizer	NRM_*TransformationName*
Rank	RNK_*TransformationName*
Router	RTR_*TransformationName*
Sequence Generator	SEQ_*TransformationName*
Sorter	SRT_*TransformationName*
Stored Procedure	SP_*TransformationName*
Source Qualifier	SQ_*TransformationName*
Transaction Control	TC_*TransformationName*
Union	UN_*TransformationName*
Update Strategy	UPD_*TransformationName*
XML Generator	XG_*TransformationName*
XML Parser	XP_*TransformationName*
XML Source Qualifier	XSQ_*TransformationName*

Targets

The naming convention for targets is: T_*TargetName*.

Mappings

The naming convention for mappings is: m_*MappingName*.

Mapplets

The naming convention for mapplets is: mplt_*MappletName*.

Sessions

The naming convention for sessions is: s_*MappingName*.

Worklets

The naming convention for worklets is: wl_*WorkletName*.

Workflows

The naming convention for workflows is: wf_*WorkflowName*.