

IBM® DB2 Universal Database™



Business Intelligence Tutorial: Extended Lessons in Data Warehousing

Version 8

IBM® DB2 Universal Database™



Business Intelligence Tutorial: Extended Lessons in Data Warehousing

Version 8

Before using this information and the product it supports, be sure to read the general information under *Notices*.

This document contains proprietary information of IBM. It is provided under a license agreement and is protected by copyright law. The information contained in this publication does not include any product warranties, and any statements provided in this manual should not be interpreted as such.

You can order IBM publications online or through your local IBM representative.

- To order publications online, go to the IBM Publications Center at www.ibm.com/shop/publications/order
- To find your local IBM representative, go to the IBM Directory of Worldwide Contacts at www.ibm.com/planetwide

To order DB2 publications from DB2 Marketing and Sales in the United States or Canada, call 1-800-IBM-4YOU (426-4968).

When you send information to IBM, you grant IBM a nonexclusive right to use or distribute the information in any way it believes appropriate without incurring any obligation to you.

© **Copyright International Business Machines Corporation 2000 - 2002. All rights reserved.**

US Government Users Restricted Rights – Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

Contents

About the tutorial	v	Viewing published objects in the Information Catalog Center	20
Before you begin	v	Updating published metadata	20
Conventions that are used in this tutorial	v	What you just did	21
Tutorial business problem	vi		
Chapter 1. Creating a star schema from within the Data Warehouse Center	1	Chapter 3. Maintaining the data warehouse	23
Defining tables for the star schema	1	Creating an index	23
Defining dimension tables.	1	Collecting table statistics	24
Creating a fact table.	3	Reorganizing a table	24
Creating the physical target tables	5	What you just did	24
Defining keys on target tables	5		
Defining a star schema	11	Chapter 4. Summary.	25
Opening the schema	11		
Adding tables to the schema	11	Appendix. Related information	27
Autojoining the tables.	12		
What you just did	12	Notices	29
		Trademarks	32
Chapter 2. Cataloging the warehouse for end users	15	Contacting IBM	35
Creating the information catalog	16	Product information	35
Selecting metadata to publish	17		

About the tutorial

This tutorial is a continuation of the *Business Intelligence Tutorial: Introduction to the Data Warehouse Center* and is written for Windows NT[®], Windows[®] 2000, Windows XP, Windows 98, and Windows ME. In this tutorial, you will learn the following tasks:

- Creating a star schema from within the DB2[®] Data Warehouse Center
- Cataloging data in the warehouse for end users
- Maintaining the data warehouse

The tutorial is available in HTML or PDF format at:

<http://www.ibm.com/software/data/bi/downloads.html>

Duration:

This tutorial takes approximately 1 hour to complete.

Before you begin

Before you begin, you must complete the *Business Intelligence Tutorial: Introduction to the Data Warehouse Center* in order to create the Data Warehouse Center objects that are necessary to complete the lessons in this tutorial.

To complete the lesson on cataloging metadata in the data warehouse for end users, the Information Catalog Center must be installed on the warehouse workstation.

Conventions that are used in this tutorial

This tutorial uses typographical conventions in the text to help you distinguish between the names of controls and text that you type. For example:

- Menu items are in boldface font:
Click **Menu** —> **Menu choice**.
- The names of fields, check boxes, and other controls are also in boldface font:
Type text in the **Field name** field.
- Text that you type is in monospaced font on a new line:
This is the text that you type.

Tutorial business problem

You are a database administrator for a company that is called TBC: The Beverage Company. The company manufactures beverages for sale to other businesses. The financial department wants to track, analyze, and forecast the sales revenue across geographies on a periodic basis for all products sold. You have already set up standard queries of the sales data. However, these queries add to the load on your operational database. Also, users sometimes ask for additional ad hoc queries of the data, based on the results of the standard queries.

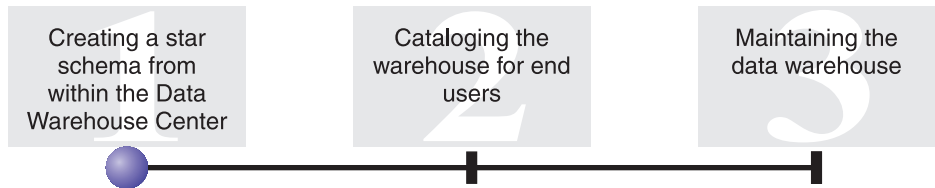
In the *Business Intelligence Tutorial: Introduction to the Data Warehouse Center*, you created a data warehouse for the sales data. In this tutorial, you will learn to maintain the warehouse that you created and work with warehouse metadata.

In this tutorial, you will create a star schema within the Data Warehouse Center. A *star schema* is a specialized design that consists of multiple dimension tables, and one fact table. *Dimension tables* describe aspects of a business. The *fact table* contains the facts about the business. In this tutorial, the star schema includes the following dimensions:

- Products
- Markets
- Scenario
- Time

The facts in the fact table include orders of the products over a period of time.

Chapter 1. Creating a star schema from within the Data Warehouse Center



In this lesson, you will create a star schema. You can use this star schema for queries in the warehouse database. You can also export the star schema to OLAP Integration Server to create an OLAP database.

This lesson shows you how to complete the following tasks:

- Define tables for the star schema
- Define a star schema

This lesson takes approximately 30 minutes to complete.

Defining tables for the star schema

This exercise shows you how to build the rest of the dimension tables and the fact table in the star schema.

When you define each table, you must define a new process for the table. Instead of defining a step for the process, you will copy the step that is defined in the sample. When you copy the step, the Data Warehouse Center copies the sources that the step uses and generates a target table.

Defining dimension tables

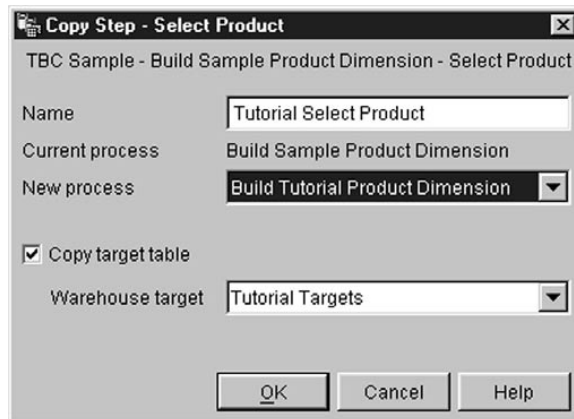
This exercise shows you how to create the rest of the dimension tables in the star schema, beginning with the Product dimension table.

To define the Product dimension table:

1. Define a new process under the TBC Tutorial subject area named:
Build Tutorial Product Dimension

Remember to give the Tutorial Warehouse Group privileges on the Security page.

2. Under the **Subject Areas** folder in the Data Warehouse Center window, expand the **TBC Sample** tree until you see the **Build Sample Product Dimension** process.
 3. Right-click the **Build Sample Product Dimension** process.
 4. Click **Open**.
The Process Model window opens.
 5. Right-click the **Select Product** step.
 6. Click **Copy**.
The Copy Step window opens.
 7. In the **Name** field, type the name for the copy of the step:
Tutorial Select Product
 8. In the **New process** list, specify the following process name:
Build Tutorial Product Dimension
- The step will be copied to the Build Tutorial Product Dimension process.
9. Verify that the **Copy target table** check box is selected.
 10. In the **Warehouse target** list, specify the **Tutorial Targets** warehouse target. The warehouse target is the same for all of the tables that you define in this lesson.



11. Click **OK**.
The step and its sources are copied to the Build Tutorial Product Dimension process. The Data Warehouse Center creates the corresponding target table.
12. Close the Process Model window.
13. Open the Build Tutorial Product Dimension process.
14. Verify that the process includes the following objects:
 - The **PRODUCT** source table

- The Tutorial Select Product step
 - The "SelectProd_T" target table
15. Rename the "SelectProd_T" target table:
 - a. In the Process Model window, right-click the "SelectProd_T" target table.
 - b. Click **Properties**.
The Properties notebook for the table opens.
 - c. In the **Table name** field, type:
LOOKUP_PRODUCT
 - d. Verify that the **Part of an OLAP schema** check box and **Dimension table** radio button are selected.
 - e. Click **OK** to save your changes and close the Properties notebook for the table.
 16. Save the process and close the Process Model window. If you receive error 3171, click **OK** and continue. This is an informational message that has no effect on the changes you have made.
 17. Repeat this procedure for the Time and Scenario dimension tables, substituting the values in the following table.

	Time dimension table	Scenario dimension table
Tutorial process:	Build Tutorial Time Dimension	Build Tutorial Scenario Dimension
Sample Process:	Build Sample Time Dimension	Build Sample Scenario Dimension
Sample step to copy:	Select Time	Select Scenario
New tutorial step name:	Tutorial Select Time	Tutorial Select Scenario
Source tables:	TIME	SCENARIO
Target table:	"SelectTime_T"	"SelectScenario_T"
New target table name:	LOOKUP_TIME	LOOKUP_SCENARIO
Warehouse target:	Tutorial Targets	Tutorial Targets

Creating a fact table

In the previous exercise, you defined the dimension tables in the star schema. In this exercise, you will define the fact table in the star schema.

To define the fact table:

1. Under the TBC Tutorial subject area, define a new process named:
Build Tutorial Fact Table

Remember to give the Tutorial Warehouse Group privileges on the Security page.

2. Under the **Subject Areas** folder in the Data Warehouse Center window, expand the **TBC Sample** tree until you see the **Build Sample Fact Table** process.
3. Right-click the **Build Sample Fact Table** process.
4. Click **Open**.
The Process Model window opens.
5. Right-click the **Fact Table Join** step.
6. Click **Copy**.
The Copy Step window opens.
7. In the **Name** field, type the name for the copy of the step:
Tutorial Fact Table Join
8. In the **New process** field, select the name of the process to which you want to copy the step:
Build Tutorial Fact Table
9. Verify that the **Copy Target Table** check box is selected.
10. In the **Warehouse target** list, specify the **Tutorial Targets** warehouse target.
11. Click **OK**.
The step and its sources are copied to the Build Tutorial Fact Table process. The Data Warehouse Center generates the corresponding target table.
12. Close the Process Model window.
13. Open the Build Tutorial Fact Table process. If it is already open, close it, then open it again.
14. Verify that the process includes the following objects:
 - The PRODUCTION_COSTS, SALES, and INVENTORY source tables
 - The Tutorial Fact Table Join step
 - The "FactTable_T" target table
15. Rename the "FactTable_T" target table:
 - a. In the Process Model window, right-click the "FactTable_T" target table.
 - b. Click **Properties**.
The Properties notebook for the table opens.
 - c. In the **Table name** field, type:
FACT_TABLE
 - d. Verify that the **Part of an OLAP schema** check box and the **Fact table** radio button are selected.

- e. Click **OK** to save your changes and close the Properties notebook for the table.

16. Save the process, and close the Process Model window.

Creating the physical target tables

The physical target tables are created when you promote steps to test or production mode.

To promote the steps that you copied to test mode:

1. Open the Build Tutorial Fact Table process.
2. Right-click the **Tutorial Fact Table Join** step, and click **Mode** → **Test**.
A progress window opens. If an error occurs when the step is processing, you will receive an error message.
3. Repeat this procedure for the following steps:

Process	Step
Build Tutorial Product Dimension	Tutorial Select Product
Build Tutorial Scenario Dimension	Tutorial Select Scenario
Build Tutorial Time Dimension	Tutorial Select Time
Build Tutorial Market Dimension	Load Demographics Data, Select Geographies Data, Join Market Data. If these steps are in production mode, demote them to test.

Defining keys on target tables

In each target table, you will select a column that can be used to uniquely identify rows in that table. This is the primary key for the table. The column that you select as a primary key must have the following qualities:

- It must always have a value. The column for a primary key cannot contain null values.
- It must have unique values. Each value in the column must be different for each row in the table.
- Its values must be stable. A value must never change to another value.

For example, the `CITY_ID` column in the `LOOKUP_MARKET` table is a good candidate for designation as a primary key. Because each city needs an identifier, no two cities can have the same identifier, and identifiers are unlikely to change.

Defining a primary key for a table is highly recommended because uniquely identifying each row speeds up row access.

You use foreign keys to define relationships between tables. In a star schema, a foreign key defines the relationship between the fact table and its associated dimension tables. The primary key of the dimension table has a corresponding foreign key in the fact table. The foreign key requires that all the values of a given column in the fact table also exist in the dimension table. For example, the CITY_ID column of the fact table might have a foreign key defined on the CITY_ID column of the LOOKUP_MARKET dimension table. This means that a row cannot exist in the fact table unless the CITY_ID exists in the LOOKUP_MARKET table.

In this exercise, you will define primary keys on the following tables: LOOKUP_MARKET, LOOKUP_TIME, LOOKUP_PRODUCT, and LOOKUP_SCENARIO. You will define corresponding foreign keys in the FACT_TABLE.

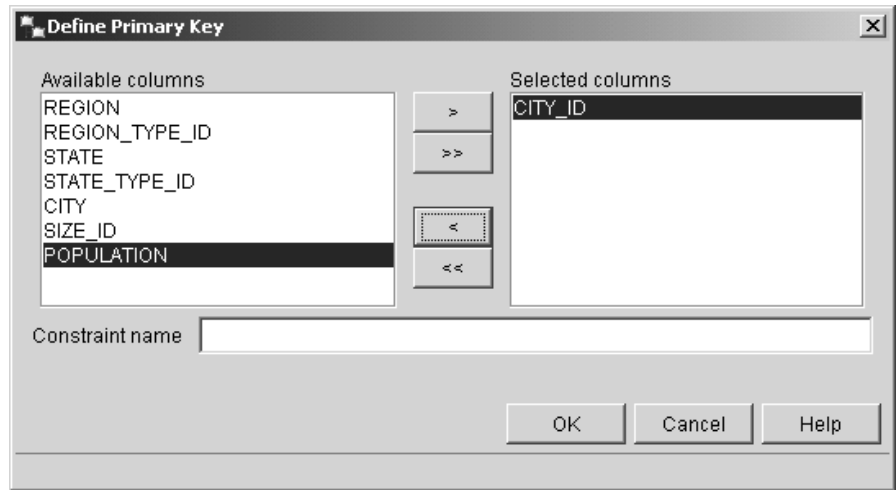
Defining a primary key

This exercise shows you how to define a primary key for the LOOKUP_MARKET, LOOKUP_TIME, LOOKUP_PRODUCT, and LOOKUP_SCENARIO target tables.

To define the primary keys:

1. Click **Start** → **Programs** → **IBM DB2** → **General Administration Tools** → **Control Center**.
The Control Center window opens.
2. Click **View** → **Refresh**.
3. Expand the tree until you see the TUTWHS database.
4. Expand the TUTWHS tree, and click on the **Tables** folder.
5. Right-click the LOOKUP_MARKET table, and click **Alter**.
The Alter Table notebook opens.
6. Click the **Keys** tab.
7. Click **Add Primary**. The Define Primary Key window opens.
8. In the **Available columns** list, click **CITY_ID**.
9. Click > to move CITY_ID to the **Selected columns** list. CITY_ID will be the primary key for the TUTWHS database.
10. Leave the **Constraint name** field empty, so that DB2 Universal Database™ will generate a constraint name for you. A primary key can be considered

a constraint, because all values in the selected column must be unique.



11. Click **OK** to save your definition.
12. Click **Close** to close the Progress window. If the change was successful, the DB2 Message window contains the following message:
The command completed successfully.
13. Click **Close** to close the DB2 Message window.

Follow the same steps to define primary keys for the other target tables.
Define:

Table	Primary key
LOOKUP_TIME	TIME_ID
LOOKUP_PRODUCT	PRODUCT_KEY
LOOKUP_SCENARIO	SCENARIO_ID

Defining a foreign key

You need to define foreign keys for the relationships between the FACT_TABLE and the other target tables.

In this exercise, you will define a foreign key in the FACT_TABLE (dependent table) based on the primary key of the LOOKUP_MARKET table (parent table).

To define the foreign keys:

1. Find the fact table in the list of tables for the TUTWHS database. Right-click the table and click **Alter**.
The Alter Table notebook opens.

2. Click the **Keys** tab.
3. Click **Add Foreign**.
The Add Foreign Key window opens.
4. In the **Table schema** list, select **IWH**.
5. In the **Table name** field, specify **LOOKUP_MARKET**, which is the parent table. The **Primary key** field displays the primary key for the **LOOKUP_MARKET** table. The steps that generate the parent table must be in test or production mode in order for the primary key to be available. A primary key must be defined for the parent table before a foreign key can be defined for the **FACT_TABLE**.
6. In the **Available columns** list, select **CITY_ID** as the column that you want defined as the foreign key.
7. Click **>** to move **CITY_ID** to the **Foreign key** list.

Add Foreign Key

Parent table

Table schema: IWH

Table name: LOOKUP_MARKET

Primary key: CITY_ID

Available columns:

- PRODUCT_KEY
- TIME_ID
- SCENARIO_ID
- TRANSDATE
- SALES
- COGS
- MARKETING
- MISC
- PAYROLL

Foreign key: CITY_ID

On delete: NO ACTION

On update: NO ACTION

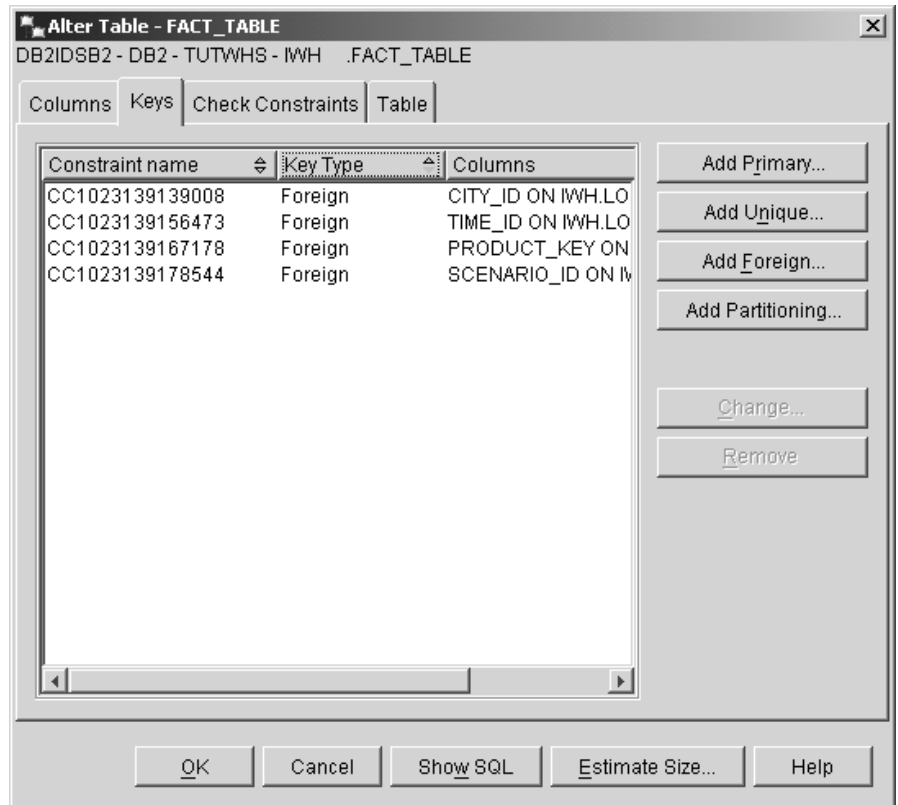
Constraint name:

OK Cancel Apply Help

8. Click **OK** to save your definitions.

Define foreign keys for the FACT_TABLE to the other target tables by repeating steps 3 through 8. Define:

Table	Foreign key
LOOKUP_TIME	TIME_ID
LOOKUP_PRODUCT	PRODUCT_KEY
LOOKUP_SCENARIO	SCENARIO_ID



9. Click **OK** to save the information in the Alter Table window.
10. Click **Close** to close the Progress window. If the change was successful, the DB2 Message window contains the following message:
The command completed successfully.
11. Click **Close** to close the DB2 Message window.

Defining foreign keys in the Data Warehouse Center

In this section, you will define foreign keys in the Data Warehouse Center.

To define foreign keys in your warehouse:

1. In the Data Warehouse Center window, expand the **Warehouse Targets** tree.
2. Expand the **Tutorial Targets** tree.
3. Click the **Tables** folder.
4. Right-click the fact table.
5. Click **Properties**.
The table Properties notebook opens.
6. Click the **Warehouse Foreign Keys** tab.
7. Remove each foreign key in the view by right-clicking the constraint name, and clicking **Remove**. These foreign key definitions were carried over when the Fact Table Join step was copied. The definitions refer to tables in the TBC Sample Targets database and must be deleted and replaced with new foreign keys that refer to tables in the Tutorial Targets database.
8. Right-click on the white space and click **Define**. The Define Warehouse Foreign Key window opens.
9. Select **IWH** in the **Object schema** list.
10. Select **LOOKUP_MARKET** in the **Object name** list. The primary key for LOOKUP_MARKET appears in the **Warehouse primary key** field.
11. Select **CITY_ID** in the **Available columns** field.
12. Click **>** to move CITY_ID into the **Warehouse foreign key columns** field.
13. In the **Constraint name** field, type:

Whse Market FK

The constraint name must be different for each foreign key.

14. Click **OK** to save your definition and close the Define Warehouse Foreign Key window.
15. Repeat steps 8 through 14 to add foreign keys for the remaining three LOOKUP tables. The constraint names for the remaining three LOOKUP tables are shown in the following table.

Table	Column	Constraint name
LOOKUP_TIME	TIME_ID	Whse Time FK
LOOKUP_PRODUCT	PRODUCT_KEY	Whse Product FK
LOOKUP_SCENARIO	SCENARIO_ID	Whse Scenario FK

16. Click **OK** to save your changes and close the table Properties window.

Defining a star schema

In this exercise, you will define a star schema that is to contain the dimension and fact tables that you defined in this tutorial.

To define a star schema:

1. From the Data Warehouse Center, right-click the **Warehouse Schemas** folder.

2. Click **Define**.

The Define Warehouse Schema notebook opens.

3. In the **Name** field, type the schema name:

Tutorial Schema

4. In the **Administrator** field, type your name as the contact for the schema.

5. In the **Description** field, type the description of the schema:

This is the TBC star schema

6. Select the **Use only one database** check box.

7. From the **Warehouse target database** list, select **TUTWHS**.

8. Click **OK** to save your changes, and close the notebook.

The star schema is added to the tree under the **Warehouse Schemas** folder.

Opening the schema

This exercise shows you how to open the schema in the Data Warehouse Center.

To open the Tutorial Schema:

1. Expand the **Warehouse Schemas** tree.
2. Right-click on **Tutorial Schema**.
3. Click **Open**.

Adding tables to the schema

In this exercise, you will add the following dimension tables and fact tables to the star schema:

To add the dimension and fact tables to the star schema:

1. Click the **Add Data** icon. 
2. Click the canvas at the spot where you want to place the tables.

The Add Data window opens.

3. Expand the Warehouse Targets tree until you see a list of tables under the **Tables** folder.

4. Select the following tables:
 - LOOKUP_MARKET
 - LOOKUP_PRODUCT
 - LOOKUP_SCENARIO
 - LOOKUP_TIME
 - FACT_TABLE
5. Click > to add the tables to the **Selected source and target tables** list.
6. Click **OK**. The tables that you selected are displayed on the window.
7. Arrange the tables in the window to place the FACT_TABLE is in the center and one dimension table in each corner.




Tip: Click **View** → **Hide Columns** to hide the table columns. This changes the icons to table icons.

Leave the Warehouse Schema Model window open for the next exercise.

Autojoining the tables

In this exercise, you will use the primary and foreign keys that you defined in “Defining keys on target tables” on page 5 to automatically join the dimension tables and the fact tables.

To autojoin the tables:

1. Click the **Save** icon in the toolbar to save your work. 
2. Select the LOOKUP_MARKET, LOOKUP_PRODUCT, LOOKUP_SCENARIO, LOOKUP_TIME, and FACT_TABLE tables.
3. Click the **Autojoin** icon in the toolbar. 
The Data Warehouse Center draws green lines between the primary keys in the dimension tables and the foreign keys in the FACT_TABLE.
4. Click the **Save** icon on the toolbar to save your work. 
The green autojoin lines become black.
5. Close the Warehouse Schema Model window.

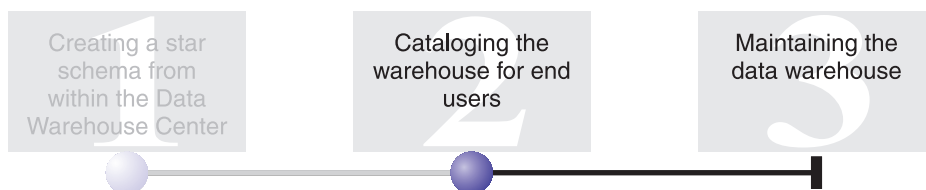
What you just did

In this lesson, you completed the following tasks:

- You defined tables for the star schema.
- You defined primary keys for the following target tables:
 - LOOKUP_MARKET

- LOOKUP_TIME
- LOOKUP_PRODUCT
- LOOKUP_SCENARIO
- You defined foreign keys for the FACT_TABLE that correspond to the primary keys.
- You created a star schema that consists of the LOOKUP_MARKET, LOOKUP_PRODUCT, LOOKUP_SCENARIO, LOOKUP_TIME, and FACT_TABLE tables.

Chapter 2. Cataloging the warehouse for end users



For this lesson, you must have the Manage Information Catalog Wizard and the Information Catalog Center installed.

This lesson shows you how to catalog the data in your data warehouse for use by end users. You catalog the data by publishing Data Warehouse Center metadata in an information catalog. An *information catalog* is the set of tables managed by the Information Catalog Center that contains business metadata that helps users identify and locate data and information available to them in the organization. Users can search the information catalog to find the tables that contain the data that they need to query.

Publishing metadata is the process of transferring metadata from the Data Warehouse Center to the Information Catalog Center. In this lesson, you will publish the metadata for the Build Tutorial Market Dimension process that you created in the *Business Intelligence Tutorial: Introduction to the Data Warehouse Center*. When you publish the process, you will publish the metadata for the following objects that are contained in the process:

- The Load Demographics Data step, its source file DEMOGRAPHICS, and its target table DEMOGRAPHICS_TARGET.
- The Select Geographies Data step, its source table GEOGRAPHIES, and its target table GEOGRAPHIES_TARGET.
- The Join Market Data step and its target table LOOKUP_MARKET. (Its source tables will be published with the other two steps.)

This lesson takes approximately 15 minutes to complete.

For more information about working with business metadata in the Information Catalog Center, see the *IBM® DB2 Warehouse Manager Information Catalog Center Tutorial* or the *IBM DB2 Warehouse Manager Information Catalog Center Administration Guide*.

Creating the information catalog

First, you must create the information catalog to hold the metadata that you publish. You can use a separate database for your information catalog, but for this tutorial, you will use the warehouse control database TBC_MD.

To create the information catalog:

1. Click **Start** —> **Programs** —> **IBM DB2** —> **Set-up tools** —> **Manage Information Catalog Wizard**.

The Manage Information Catalog wizard opens.

2. Select **Prepare an information catalog**.
3. Click **Next**.
4. Select the **In another DB2 database** radio button.
5. In the **Database name** field, type the following name:
TBC_MD
6. In the **Database schema** field, type the name of the default schema for the Information Catalog Center:
ICM
7. In the **Database user ID** field, type the user ID for the database. In this case, it is the same user ID that you used to log on to the Data Warehouse Center.
8. In the **Database password** field, type the password that corresponds to the user ID.
9. In the **Default user group** field, type the following name:
ICCUSER

This is not a Data Warehouse Center group. It is a user group that must be defined on your operating system with the privileges that you want to assign to the default user group. For this tutorial, you do not have to create the ICC Default Group on your operating system to complete the lesson, but please note that if you do not create the user group on your operating system with the desired privileges, the user group will not be able to control access to objects in the information catalog.

10. In the **Default power user group** field, type the following name:
ICCPUSER

This is not a Data Warehouse Center user group. It is a user group that must be defined on your operating system with the privileges that you want to assign to the power user group. For this tutorial, you do not have to create the ICC Default Group on your operating system to complete the lesson, but please note that if you do not create the user

group on your operating system with the desired privileges, the user group will not be able to control access to objects in the information catalog.

11. Select the **Create default object types** check box.
12. In the **Table space** field, type:
USERSPACE1
13. Click **Next**.
14. Review the information that you entered on the summary page, and click **Finish**. The Manage Information Catalog Wizard prepares your information catalog.

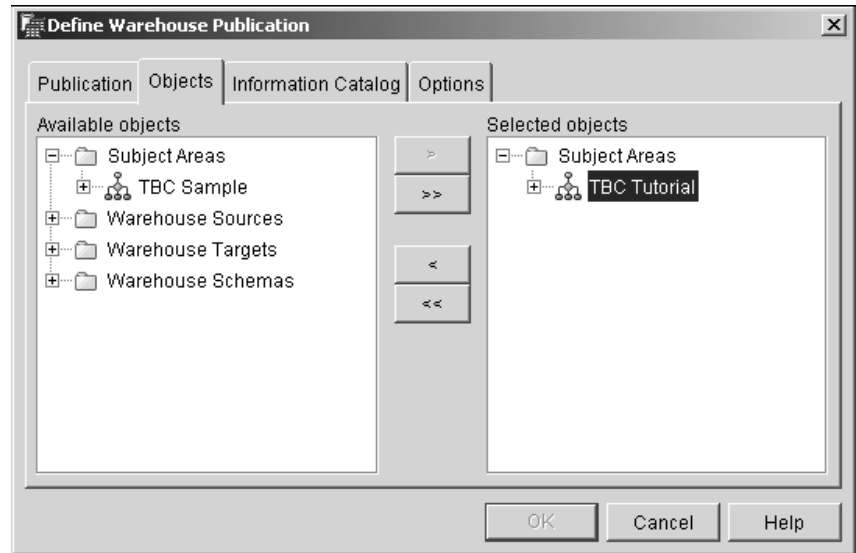
Selecting metadata to publish

After you create the metadata, you need to select the metadata that you want to publish.

To select the metadata that you want to publish:

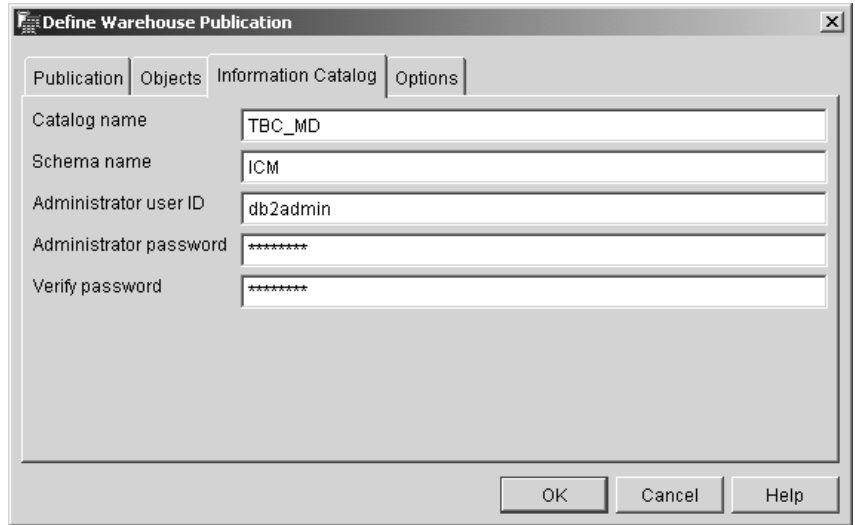
1. From the Data Warehouse Center window, right-click **Warehouse**, and click **Publish Metadata —> Data Warehouse to Information Catalog**.
The Publish Metadata - Data Warehouse to Information Catalog window opens.
2. Click **Define**.
3. In the **Name** field, type the following business name for the publication:
Published Tutorial Metadata
4. In the **Administrator** field, type your name as the contact for the publication.
5. Type a description of the publication in the **Description** field. For this tutorial, type the following sentence:
Published metadata for the Business Intelligence Tutorial.
6. Click the **Objects** tab, and specify the objects that you want to use:
 - a. In the **Available objects** list, expand the **Subject Areas** folder and click **TBC Tutorial**.
 - b. Click **>**.
The TBC Tutorial subject area moves to the **Selected objects** list.

The Define Warehouse Publication notebook opens.



7. Click the **Information Catalog** tab, and specify the following information:
 - a. In the **Catalog name** field, type:
TBC_MD

The TBC_MD database contains sample metadata.
 - b. In the **Schema name** field, type:
ICM
 - c. In the **Administrator user ID** field, type the user ID for the information catalog. In this case, it is the same user ID that you used to log on to the Data Warehouse Center.
 - d. In the **Administrator password** field, type the password that corresponds to the user ID.
 - e. In the **Verify password** field, type the password again.



8. Click the **Options** tab.
9. From the **Source to target mapping** group, verify that the **Table level** radio button is selected. This option specifies that source tables and target tables are mapped at the table level. The information catalog uses a transformation object to indicate that a mapping exists between the two tables.
10. Click **OK**.

If errors occur during the publication process, see the *IBM DB2 Universal Database Message Reference*.

If the publication step was created successfully, it is added to the list of publication steps in the Publish Metadata window. To publish the metadata, you must run the publication step.
11. Right-click the publication step in the Publish Metadata window, and click **Run**. This runs the step and copies the metadata into the information catalog. When the step completes successfully, there is a number in the status field that corresponds to the edition number for the last time the step ran. You can use this number to look up the statistics for the step in the Work in Progress window.

Leave the Publish Metadata - Data Warehouse to Information Catalog window open for the next exercise.

Viewing published objects in the Information Catalog Center

After you publish your metadata, you can view the objects in the Information Catalog Center.

To view published objects in the Information Catalog Center:

1. Click **Start** —> **Programs** —> **IBM DB2** —> **Business Intelligence Tools** —> **Information Catalog Center**.

The Information Catalog Logon window opens.

2. In the **User ID** field, verify that the user ID for your information catalog is specified.
3. In the **Password** field, type the password that corresponds to the user ID.
4. In the **Database** field, verify that TBC_MD is specified.
5. In the **Information catalog** field, verify that ICC is specified.
6. Click **OK**.

The Information Catalog Logon window closes, and you can see your published metadata in the Information Catalog Center window.

Updating published metadata

After you publish metadata, you need to update it periodically to capture the changes that are made in the Data Warehouse Center. To transfer updates of the Data Warehouse Center metadata to the information catalog, you run the step for the publication in the same way that you run any other steps in the Data Warehouse Center.

To update published metadata:

1. In the **Warehouse publications** list in the Publish Metadata - Data Warehouse to Information Catalog window, right-click **Published Tutorial Metadata**.
2. Click **Run**.
3. Go to the main Data Warehouse Center window and click **Warehouse** —> **Work in Progress**.

The Work in Progress window opens. You should see an entry for the step that is running. While the step is running, the status is Populating. When the step has completed the status is Successful.

4. Close the Work In Progress window.
5. Close the Publish Metadata window.

What you just did

In this lesson, you published to an information catalog the Data Warehouse Center metadata that you created in the tutorial. You ran the publication to update the metadata that you published.

Chapter 3. Maintaining the data warehouse



In this lesson, you will learn to define indexes, and use the RUNSTATS and REORG utilities.

The amount of maintenance that is required for a database is directly related to the amount of database activity or workload. Because this tutorial does not generate significant database activity, this lesson is primarily a guide to the DB2 tools and utilities that you can use when you are maintaining an actual warehouse database.

This lesson takes approximately 15 minutes to complete.

Creating an index

You can create an index to optimize queries for end users of the warehouse. An *index* is a set of keys, each pointing to a set of rows in a table. The index is a separate object from the table data. The database manager builds the index structure and maintains it automatically. An index gives more efficient access to rows in a table by creating a direct path to the data through the pointers that it creates.

An index is created when you define a primary key or a foreign key. For example, an index was created on the LOOKUP_MARKET table when you defined CITY_ID as its primary key.

To create additional indexes:

1. From the DB2 Control Center, expand the objects within the TUTWHS database until you see the Indexes folder.
2. Right-click the **Indexes** folder and click **Create**.
The Create Index window opens.
3. Click **Help** and follow the instructions given in the Control Center help to create your index.

Collecting table statistics

Table statistics provide information about the physical and logical characteristics of a table and its indexes. You must periodically collect these statistics, so that DB2 Universal Database can determine the best way to access your data. If extensive changes are made to the data in a table, and the last collection of statistics no longer reflects the actual table data, then data access performance can deteriorate. In general, you should update statistics if there are major changes to the data in your table.

To collect statistics about the LOOKUP_MARKET table:

1. From the DB2 Control Center, right-click the **LOOKUP_MARKET** table, and click **Run Statistics**.

The Run Statistics notebook opens.

2. Click **Help**, and use the online help to determine the level of statistics that you want to gather for the table and its indexes.
3. Click **OK** to begin collecting the table statistics.

Reorganizing a table

Reorganizing a table rearranges the table in physical storage, eliminating fragmentation and making sure that the table is stored efficiently in the database. You can also use reorganization to control the order in which the rows of a table are stored, usually according to an index.

To reorganize the LOOKUP_MARKET table:

1. From the DB2 Control Center, right-click the **LOOKUP_MARKET** table and click **Reorganize**.

The Reorganize Table notebook opens.

2. Click **Help**, and use the online help to specify the values for the notebook.
3. Click **OK** to reorganize the table immediately.

What you just did

In this lesson, you completed the following tasks:

- Created an index
- Collected statistics on a table
- Reorganized a table

Chapter 4. Summary

Congratulations! You have completed the *Business Intelligence Tutorial: Extended Lessons in Data Warehousing*. In this tutorial, you completed the following tasks:

- You defined a star schema in the Data Warehouse Center.
- You published metadata to an information catalog.
- You performed maintenance on the warehouse database to improve performance.

Appendix. Related information

This tutorial covers the most common tasks that you can accomplish with the DB2 Control Center, Data Warehouse Center, and the Information Catalog Center. For more information about related tasks, see the following resources:

Control Center

- *IBM DB2 Universal Database Quick Beginnings for DB2 Clients*
- *IBM DB2 Universal Database Quick Beginnings for DB2 Servers*
- *IBM DB2 Universal Database Administration Guide: Implementation*

Data Warehouse Center

- *IBM DB2 Universal Database Data Warehouse Center Administration Guide*
- *IBM DB2 Warehouse Manager Installation Guide*

Information Catalog Center

- *IBM DB2 Warehouse Manager Information Catalog Center Administration Guide*
- *IBM DB2 Warehouse Manager Information Catalog Center Tutorial*

Notices

IBM may not offer the products, services, or features discussed in this document in all countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
U.S.A.

For license inquiries regarding double-byte (DBCS) information, contact the IBM Intellectual Property Department in your country/region or send inquiries, in writing, to:

IBM World Trade Asia Corporation
Licensing
2-31 Roppongi 3-chome, Minato-ku
Tokyo 106, Japan

The following paragraph does not apply to the United Kingdom or any other country/region where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY, OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions; therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make

improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product, and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Licenses of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information that has been exchanged, should contact:

IBM Canada Limited
Office of the Lab Director
8200 Warden Avenue
Markham, Ontario
L6G 1C7
CANADA

Such information may be available, subject to appropriate terms and conditions, including in some cases payment of a fee.

The licensed program described in this document and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement, or any equivalent agreement between us.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems, and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements, or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

All statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

This information may contain examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious, and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information may contain sample application programs, in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM for the purposes of developing, using, marketing, or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

Each copy or any portion of these sample programs or any derivative work must include a copyright notice as follows:

© (*your company name*) (*year*). Portions of this code are derived from IBM Corp. Sample Programs. © Copyright IBM Corp. *_enter the year or years_*. All rights reserved.

Trademarks

The following terms are trademarks of International Business Machines Corporation in the United States, other countries, or both, and have been used in at least one of the documents in the DB2 UDB documentation library.

ACF/VTAM	LAN Distance
AISPO	MVS
AIX	MVS/ESA
AIXwindows	MVS/XA
AnyNet	Net.Data
APPN	NetView
AS/400	OS/390
BookManager	OS/400
C Set++	PowerPC
C/370	pSeries
CICS	QBIC
Database 2	QMF
DataHub	RACF
DataJoiner	RISC System/6000
DataPropagator	RS/6000
DataRefresher	S/370
DB2	SP
DB2 Connect	SQL/400
DB2 Extenders	SQL/DS
DB2 OLAP Server	System/370
DB2 Universal Database	System/390
Distributed Relational Database Architecture	SystemView
DRDA	Tivoli
eServer	VisualAge
Extended Services	VM/ESA
FFST	VSE/ESA
First Failure Support Technology	VTAM
IBM	WebExplorer
IMS	WebSphere
IMS/ESA	WIN-OS/2
iSeries	z/OS
	zSeries

The following terms are trademarks or registered trademarks of other companies and have been used in at least one of the documents in the DB2 UDB documentation library:

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Intel and Pentium are trademarks of Intel Corporation in the United States, other countries, or both.

Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Other company, product, or service names may be trademarks or service marks of others.

Contacting IBM

In the United States, call one of the following numbers to contact IBM:

- 1-800-237-5511 for customer service
- 1-888-426-4343 to learn about available service options
- 1-800-IBM-4YOU (426-4968) for DB2 marketing and sales

In Canada, call one of the following numbers to contact IBM:

- 1-800-IBM-SERV (1-800-426-7378) for customer service
- 1-800-465-9600 to learn about available service options
- 1-800-IBM-4YOU (1-800-426-4968) for DB2 marketing and sales

To locate an IBM office in your country or region, check IBM's Directory of Worldwide Contacts on the web at www.ibm.com/planetwide

Product information

Information regarding DB2 Universal Database products is available by telephone or by the World Wide Web at www.ibm.com/software/data/db2/udb

This site contains the latest information on the technical library, ordering books, client downloads, newsgroups, FixPaks, news, and links to web resources.

If you live in the U.S.A., then you can call one of the following numbers:

- 1-800-IBM-CALL (1-800-426-2255) to order products or to obtain general information.
- 1-800-879-2755 to order publications.

For information on how to contact IBM outside of the United States, go to the IBM Worldwide page at www.ibm.com/planetwide



Printed in U.S.A.