

® IBM

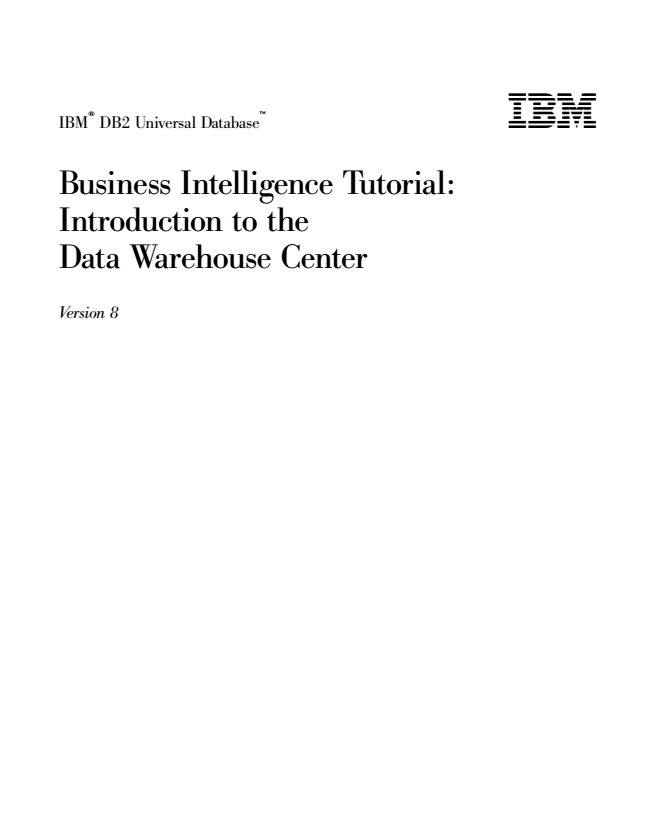
DB2 Universal Database

TM

Business Intelligence Tutorial: Introduction to the Data Warehouse Center

*Version 8*

**

**

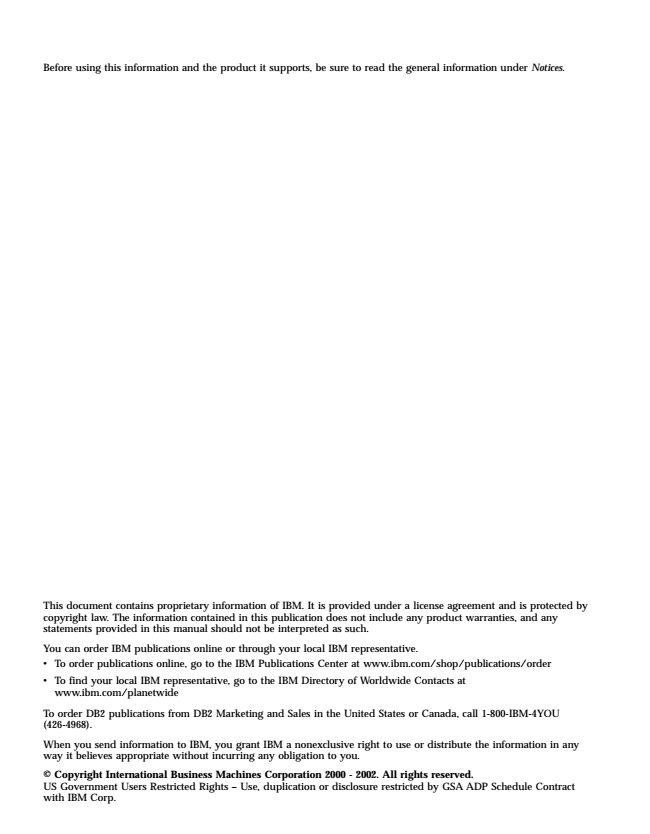
® IBM

DB2 Universal Database

TM

Business Intelligence Tutorial: Introduction to the Data Warehouse Center

*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. v To order publications online, go to the IBM Publications Center at www.ibm.com/shop/publications/order

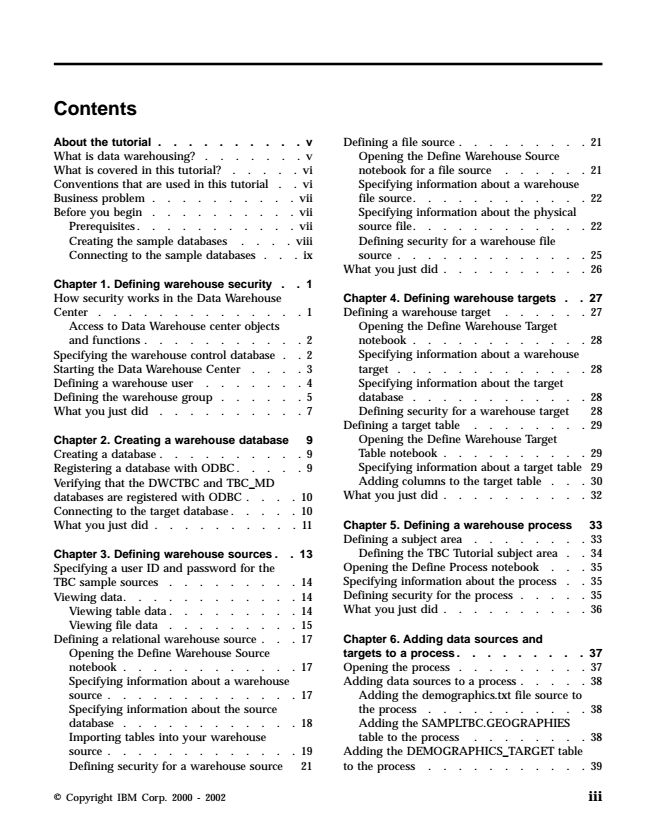
v To find your local IBM representative, go to the IBM Directory of Worldwide Contacts at

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 What is data warehousing? . . . . . . . v What is covered in this tutorial? . . . . . vi Conventions that are used in this tutorial . . vi Business problem . . . . . . . . . . vii Before you begin . . . . . . . . . . vii Prerequisites. . . . . . . . . . . vii Creating the sample databases . . . . viii Connecting to the sample databases . . . ix

Chapter 1. Defining warehouse security . . 1 How security works in the Data Warehouse Center . . . . . . . . . . . . . . 1

Access to Data Warehouse center objects and functions . . . . . . . . . . . 2 Specifying the warehouse control database . . 2 Starting the Data Warehouse Center . . . . 3 Defining a warehouse user . . . . . . . 4 Defining the warehouse group . . . . . . 5 What you just did . . . . . . . . . . 7

Chapter 2. Creating a warehouse database 9 Creating a database . . . . . . . . . . 9 Registering a database with ODBC . . . . . 9 Verifying that the DWCTBC and TBC\_MD databases are registered with ODBC . . . . 10 Connecting to the target database . . . . . 10 What you just did . . . . . . . . . . 11

Chapter 3. Defining warehouse sources . . 13 Specifying a user ID and password for the TBC sample sources . . . . . . . . . 14 Viewing data. . . . . . . . . . . . 14 Viewing table data . . . . . . . . . 14 Viewing file data . . . . . . . . . 15 Defining a relational warehouse source . . . 17

Opening the Define Warehouse Source notebook . . . . . . . . . . . . 17 Specifying information about a warehouse source . . . . . . . . . . . . . 17 Specifying information about the source database . . . . . . . . . . . . 18 Importing tables into your warehouse source . . . . . . . . . . . . . 19 Defining security for a warehouse source 21

Defining a file source . . . . . . . . . 21

Opening the Define Warehouse Source notebook for a file source . . . . . . 21 Specifying information about a warehouse file source. . . . . . . . . . . . 22 Specifying information about the physical source file. . . . . . . . . . . . 22 Defining security for a warehouse file source . . . . . . . . . . . . . 25 What you just did . . . . . . . . . . 26

**Chapter 4. Defining warehouse targets . . 27 Defining a warehouse target . . . . . . 27**

Opening the Define Warehouse Target notebook . . . . . . . . . . . . 28 Specifying information about a warehouse target . . . . . . . . . . . . . 28 Specifying information about the target database . . . . . . . . . . . . 28 Defining security for a warehouse target 28 Defining a target table . . . . . . . . 29

Opening the Define Warehouse Target Table notebook . . . . . . . . . . 29 Specifying information about a target table 29 Adding columns to the target table . . . 30 What you just did . . . . . . . . . . 32

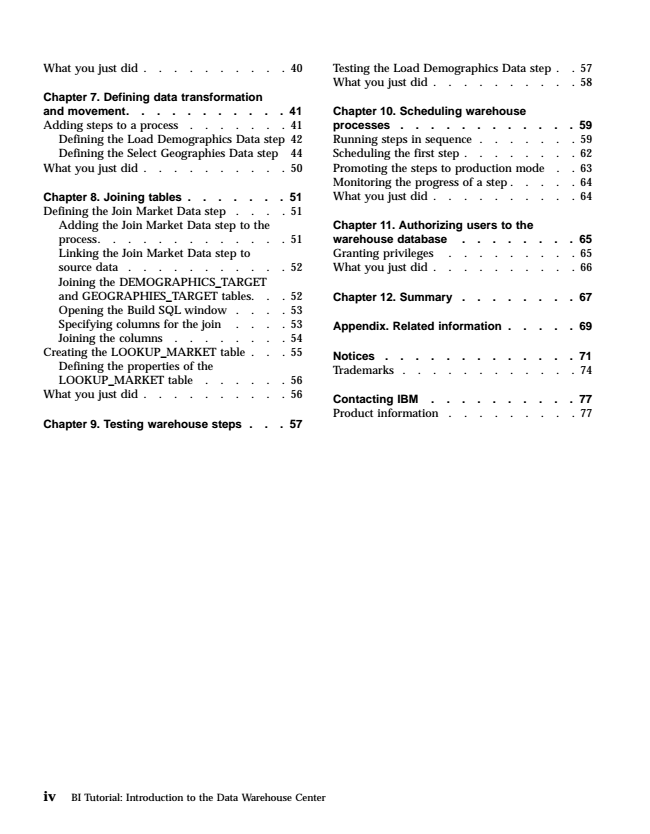
Chapter 5. Defining a warehouse process 33 Defining a subject area . . . . . . . . 33 Defining the TBC Tutorial subject area . . 34 Opening the Define Process notebook . . . 35 Specifying information about the process . . 35 Defining security for the process . . . . . 35 What you just did . . . . . . . . . . 36

Chapter 6. Adding data sources and targets to a process. . . . . . . . . 37 Opening the process . . . . . . . . . 37 Adding data sources to a process . . . . . 38

Adding the demographics.txt file source to the process . . . . . . . . . . . 38 Adding the SAMPLTBC.GEOGRAPHIES table to the process . . . . . . . . 38 Adding the DEMOGRAPHICS\_TARGET table to the process . . . . . . . . . . . 39

© Copyright IBM Corp. 2000 - 2002

**iii**

****

Testing What you just did . . . . . . . . . . 40

**the Load Demographics Data step . . 57 What you just did . . . . . . . . . . 58 Chapter 7. Defining data transformation and movement. . . . . . . . . . . 41**

Chapter 10. Scheduling warehouse Adding steps to a process . . . . . . . 41

processes . . . . . . . . . . . . 59 Defining the Load Demographics Data step 42

Running steps in sequence . . . . . . . 59 Defining the Select Geographies Data step 44

Scheduling the first step . . . . . . . . 62 What you just did . . . . . . . . . . 50

Promoting the steps to production mode . . 63 Monitoring the progress of a step . . . . . 64 Chapter 8. Joining tables . . . . . . . 51

What you just did . . . . . . . . . . 64 Defining the Join Market Data step . . . . 51

Adding the Join Market Data step to the

**Chapter 11. Authorizing users to the process. . . . . . . . . . . . . 51**

warehouse database . . . . . . . . 65 Linking the Join Market Data step to

Granting privileges . . . . . . . . . 65 source data . . . . . . . . . . . 52

What you just did . . . . . . . . . . 66 Joining the DEMOGRAPHICS\_TARGET and GEOGRAPHIES\_TARGET tables. . . 52

Chapter 12. Summary . . . . . . . . 67 Opening the Build SQL window . . . . 53 Specifying columns for the join . . . . 53

Appendix. Related information . . . . . 69 Joining the columns . . . . . . . . 54 Creating the LOOKUP\_MARKET table . . . 55

Defining the properties of the LOOKUP\_MARKET table . . . . . . 56

Notices . . . . . . . . . . . . . 71 Trademarks . . . . . . . . . . . . 74

What you just did . . . . . . . . . . 56

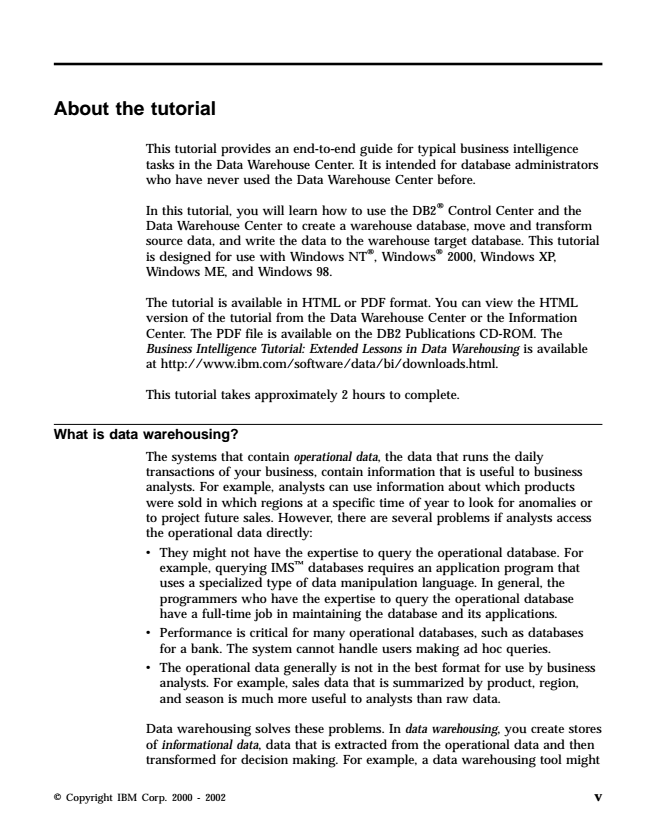
**Contacting IBM . . . . . . . . . . 77**

**Chapter 9. Testing warehouse steps . . . 57**

Product information . . . . . . . . . 77

**iv**

BI Tutorial: Introduction to the Data Warehouse Center



**About the tutorial**

This tutorial provides an end-to-end guide for typical business intelligence tasks in the Data Warehouse Center. It is intended for database administrators who have never used the Data Warehouse Center before.

In this tutorial, you will learn how to use the DB2® Control Center and the Data Warehouse Center to create a warehouse database, move and transform source data, and write the data to the warehouse target database. This tutorial is designed for use with Windows NT®, Windows® 2000, Windows XP, Windows ME, and Windows 98.

The tutorial is available in HTML or PDF format. You can view the HTML version of the tutorial from the Data Warehouse Center or the Information Center. The PDF file is available on the DB2 Publications CD-ROM. The Business Intelligence Tutorial: Extended Lessons in Data Warehousing is available at http://www.ibm.com/software/data/bi/downloads.html.

This tutorial takes approximately 2 hours to complete.

**What is data warehousing?**

The systems that contain operational data, the data that runs the daily transactions of your business, contain information that is useful to business analysts. For example, analysts can use information about which products were sold in which regions at a specific time of year to look for anomalies or to project future sales. However, there are several problems if analysts access the operational data directly: v They might not have the expertise to query the operational database. For

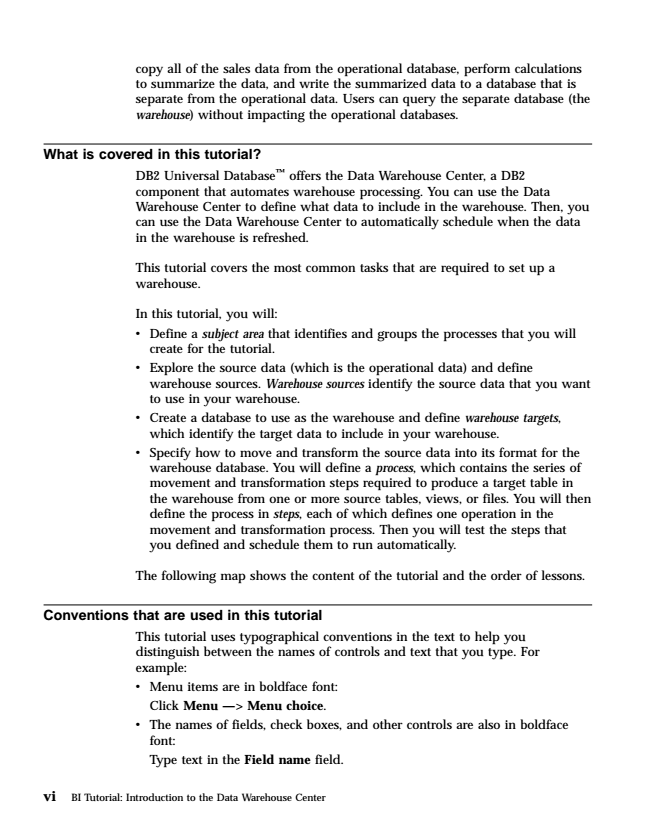
example, querying IMSTM databases requires an application program that uses a specialized type of data manipulation language. In general, the programmers who have the expertise to query the operational database have a full-time job in maintaining the database and its applications. v Performance is critical for many operational databases, such as databases

for a bank. The system cannot handle users making ad hoc queries. v The operational data generally is not in the best format for use by business analysts. For example, sales data that is summarized by product, region, and season is much more useful to analysts than raw data.

Data warehousing solves these problems. In data warehousing, you create stores of informational data, data that is extracted from the operational data and then transformed for decision making. For example, a data warehousing tool might

© Copyright IBM Corp. 2000 - 2002

**v**

****

copy all of the sales data from the operational database, perform calculations to summarize the data, and write the summarized data to a database that is separate from the operational data. Users can query the separate database (the warehouse) without impacting the operational databases.

**What is covered in this tutorial?**

DB2 Universal DatabaseTM offers the Data Warehouse Center, a DB2 component that automates warehouse processing. You can use the Data Warehouse Center to define what data to include in the warehouse. Then, you can use the Data Warehouse Center to automatically schedule when the data in the warehouse is refreshed.

This tutorial covers the most common tasks that are required to set up a warehouse.

In this tutorial, you will: v Define a subject area that identifies and groups the processes that you will

create for the tutorial. v Explore the source data (which is the operational data) and define

warehouse sources. Warehouse sources identify the source data that you want to use in your warehouse. v Create a database to use as the warehouse and define warehouse targets,

which identify the target data to include in your warehouse. v Specify how to move and transform the source data into its format for the warehouse database. You will define a process, which contains the series of movement and transformation steps required to produce a target table in the warehouse from one or more source tables, views, or files. You will then define the process in steps, each of which defines one operation in the movement and transformation process. Then you will test the steps that you defined and schedule them to run automatically.

The following map shows the content of the tutorial and the order of lessons.

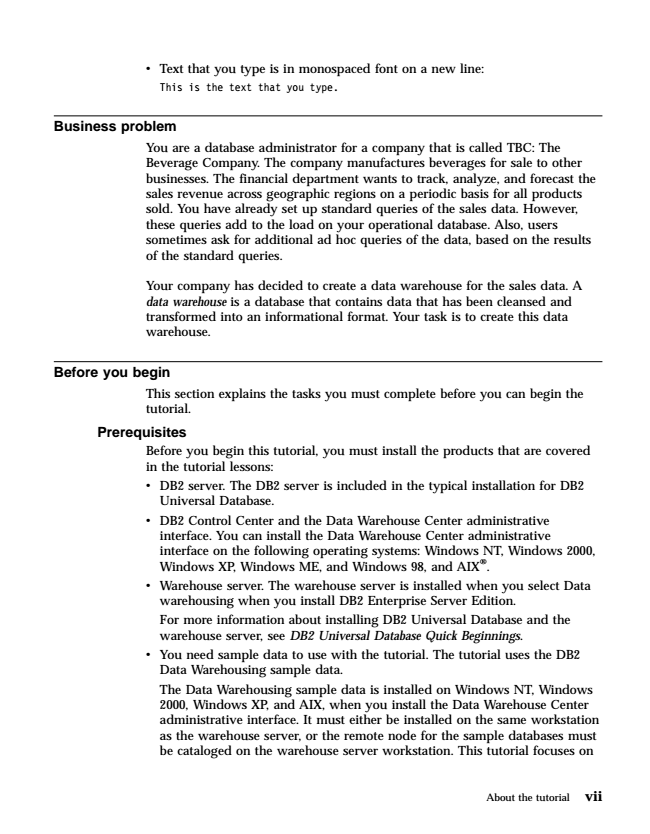
**Conventions that are used in this tutorial**

This tutorial uses typographical conventions in the text to help you distinguish between the names of controls and text that you type. For example: v Menu items are in boldface font: Click Menu —> Menu choice. v The names of fields, check boxes, and other controls are also in boldface

font: Type text in the Field name field.

**vi**

BI Tutorial: Introduction to the Data Warehouse Center



v Text that you type is in monospaced font on a new line:

This is the text that you type.

**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 geographic regions 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.

Your company has decided to create a data warehouse for the sales data. A data warehouse is a database that contains data that has been cleansed and transformed into an informational format. Your task is to create this data warehouse.

**Before you begin**

This section explains the tasks you must complete before you can begin the tutorial.

**Prerequisites**

Before you begin this tutorial, you must install the products that are covered in the tutorial lessons: v DB2 server. The DB2 server is included in the typical installation for DB2

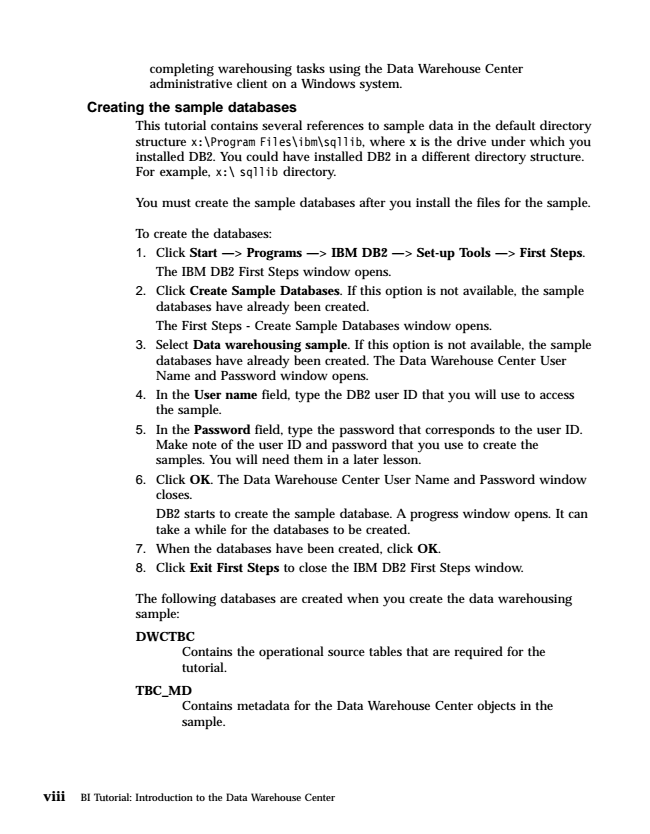
Universal Database. v DB2 Control Center and the Data Warehouse Center administrative interface. You can install the Data Warehouse Center administrative interface on the following operating systems: Windows NT, Windows 2000, Windows XP, Windows ME, and Windows 98, and AIX®. v Warehouse server. The warehouse server is installed when you select Data

warehousing when you install DB2 Enterprise Server Edition. For more information about installing DB2 Universal Database and the warehouse server, see DB2 Universal Database Quick Beginnings. v You need sample data to use with the tutorial. The tutorial uses the DB2

Data Warehousing sample data. The Data Warehousing sample data is installed on Windows NT, Windows 2000, Windows XP, and AIX, when you install the Data Warehouse Center administrative interface. It must either be installed on the same workstation as the warehouse server, or the remote node for the sample databases must be cataloged on the warehouse server workstation. This tutorial focuses on

About the tutorial

**vii**

****

completing warehousing tasks using the Data Warehouse Center administrative client on a Windows system.

**Creating the sample databases**

This tutorial contains several references to sample data in the default directory structure x:\Program Files\ibm\sqllib, where x is the drive under which you installed DB2. You could have installed DB2 in a different directory structure. For example, x:\ sqllib directory.

You must create the sample databases after you install the files for the sample.

**To create the databases: 1. Click Start —> Programs —> IBM DB2 —> Set-up Tools —> First Steps.**

The IBM DB2 First Steps window opens. 2. Click Create Sample Databases. If this option is not available, the sample

databases have already been created. The First Steps - Create Sample Databases window opens. 3. Select Data warehousing sample. If this option is not available, the sample databases have already been created. The Data Warehouse Center User Name and Password window opens. 4. In the User name field, type the DB2 user ID that you will use to access

the sample. 5. In the Password field, type the password that corresponds to the user ID.

Make note of the user ID and password that you use to create the samples. You will need them in a later lesson. 6. Click OK. The Data Warehouse Center User Name and Password window

closes. DB2 starts to create the sample database. A progress window opens. It can take a while for the databases to be created. 7. When the databases have been created, click OK. 8. Click Exit First Steps to close the IBM DB2 First Steps window.

The following databases are created when you create the data warehousing sample:

**DWCTBC**

Contains the operational source tables that are required for the tutorial.

**TBC\_MD**

Contains metadata for the Data Warehouse Center objects in the sample.

**viii**

BI Tutorial: Introduction to the Data Warehouse Center