

IBM DB2 通用数据库



商业智能教程

版本 7

IBM DB2 通用数据库



商业智能教程

版本 7

在使用本资料以及它支持的产品之前，务必阅读第149页的『注意事项』中的一般信息。

此文档包含 IBM 的专利信息。此文档通过许可证协议提供并受版权法保护。此出版物中包含的信息不包括任何产品保证，且本手册提供的任何声明都不应如此解释。

通过 IBM 代表或在您本地服务的 IBM 分支机构，或者通过致电 1-800-879-2755（在美国）或 1-800-IBM-4YOU（在加拿大）来订购出版物。

当您发送信息给 IBM 后，即授予 IBM 非专有权，IBM 对于您所提供的任何信息，有权利以任何她认为适当的方式使用或散发，而不必对您承担任何责任。

© Copyright International Business Machines Corporation 2000. All rights reserved.

目录

关于本教程.	v
教程中的商业问题	v
开始之前	vi
本教程中使用的约定	viii
相关信息	viii
与 IBM 联系	ix
产品信息	ix

第1部分 数据入库. 1

第1章 有关数据入库	3
什么是数据入库?	3
课程概述	3

第2章 创建仓库数据库	5
创建数据库	5
向 ODBC 注册数据库	6
连接目标数据库	8
您刚才做了什么	9

第3章 浏览源数据	11
查看表数据	11
查看文件数据	12
您刚才所执行的操作	13

第4章 定义仓库安全性	15
重新初始化仓库控制数据库	16
启动数据仓库中心	16
定义仓库用户	17
定义仓库组	19
您刚才做了什么	21

第5章 定义主题区	23
定义“TBC 教程”主题区	23
您刚才做了什么	24

第6章 定义仓库源	25
定义关系仓库源	25
定义文件源	29
您刚才做了什么	33

第7章 定义仓库目标	35
----------------------	----

定义仓库目标.	35
您刚才做了什么	39

第8章 定义数据变换和移动	41
定义进程	41
打开进程	42
将表添加至进程	42
将步骤添加至进程	43
定义“装入人口统计”步骤	44
定义“选择地理”步骤.	47
定义“连接市场数据”步骤	53
您刚才做了什么	58
定义星型模式的其余步骤（可选）	59
您刚才做了什么	61

第9章 测试仓库步骤	63
测试“装入人口统计数据”步骤.	63
提升星型模式中的其余步骤（可选）	65
您刚才做了什么	65

第10章 调度仓库进程	67
指定步骤按顺序运行	67
调度第一个步骤	68
将步骤提升至生产方式.	70
您刚才所执行的操作	70

第11章 定义目标表的关键字	71
定义主关键字.	71
定义外部关键字	73
您刚才所执行的操作	75

第12章 维护数据仓库	77
创建索引	77
收集表统计信息	77
重组表	79
监控数据库	79
您刚才所执行的操作	80

第13章 授予用户使用仓库数据库的权限	81
授予特权	81
您刚才所执行的操作	81

第14章 编目仓库中的数据以供最终用户使用	83	第21章 连接和编辑维度表	119
创建信息目录	83	编辑维度表	120
选择要发布的元数据	84	您刚才做了什么	120
更新已发布的元数据	86		
您刚才所执行的操作	86	第22章 定义分层结构	121
		创建分层结构	121
第15章 使用商业元数据	87	预览分层结构	122
打开信息目录	87	您刚才做了什么	123
浏览主题	87		
搜索信息目录	89	第23章 预览和保存 OLAP 模型	125
创建对象集合	91	您刚才做了什么	127
启动程序	92		
创建“程序”对象	93	第24章 启动 OLAP 元轮廓	129
从“文件”对象启动程序	96	启动“元轮廓辅助程序”	129
您刚才所执行的操作	97	连接至源数据库	130
		您刚才做了什么	131
第16章 从“数据仓库中心”内创建星型模式	99	第25章 选择维和成员	133
定义星型模式	99	您刚才做了什么	134
打开模式	100		
将表添加至模式	100	第26章 设置特性	135
自动连接表	100	设置维特性	135
调出星型模式	101	设置成员特性	136
您刚才做了什么	102	检查帐户特性	137
		您刚才做了什么	138
第17章 摘要	103	第27章 设置过滤器	139
		复查过滤器	140
第2部分 多维数据分析	105	您刚才做了什么	140
		第28章 创建 OLAP 应用程序	141
第18章 关于多维分析	107	您刚才做了什么	142
什么是多维分析?	107		
课程概述	107	第29章 探究 Starter Kit 的其余部分	143
		探究“OLAP 模型”界面	143
第19章 启动 OLAP 模型	109	探究“OLAP 元轮廓”界面	144
启动“OLAP 集成服务器”桌面	109	探究“管理管理器”	144
连接至 OLAP 目录	109	您刚才做了什么	145
启动“模型辅助程序”	111		
您刚才做了什么	111	第3部分 附录及附属资料	147
		注意事项	149
第20章 选择事实表并创建维	113	注册商标	151
选择事实表	113		
创建时间维	114		
创建标准维	115		
您刚才做了什么	117		

关于本教程

本教程提供了典型的商业智能任务的端对端的指导。它分为两个主要章节：

数据入库

学习本节中的课程以了解如何使用“DB2®控制中心”和“数据仓库中心”创建仓库数据库、移动和变换源数据以及将数据写入仓库目标数据库。完成本节可能需要大约两小时的时间。

多维数据分析

学习本节中的课程以了解如何通过联机分析处理（OLAP）技术来使用 OLAP Starter Kit 对关系数据执行多维分析。完成本节可能需要大约一小时的时间。

本教程有 HTML 和 PDF 两种格式。您可以从“数据仓库中心”、OLAP Starter Kit 或“信息中心”查看本教程的 HTML 版本。PDF 文件在“DB2 出版物”CD-ROM 上。

教程中的商业问题

您是某公司的数据库管理员，该公司称为 TBC: The Beverage Company。该公司制造要销售给其他行业的饮料。财务部门想要以时间为基础来跟踪、分析和预测各地区所有售出产品的销售收入。您已经建立了销售数据的标准查询。但是，这些查询增加了您的操作数据库的负荷。而且，根据标准查询的结果，用户有时请求额外的特别数据查询。

您的公司决定为销售数据创建一个数据仓库。数据仓库是一个包含数据的数据库，这些数据已经过提炼且变换为资料性格式。您的任务是创建此数据仓库。

您计划对仓库使用星型模式设计。星型模式是由多个维度表和一个事实表组成的专门设计。维度表描述了业务的各个方面。事实表包含有关该业务的事实或度量。在本教程中，星型模式包括下列维：

- 产品
- 市场
- 方案
- 时间

事实表中的事实包括一段时间内的产品订购情况。

本教程中的“数据入库”部分告诉您如何定义此星型模式。

您的下一个任务是创建用来分析数据的 OLAP 应用程序。首先创建一个 OLAP 模型和元轮廓，然后用它们创建应用程序。本教程中的“多维分析”部分告诉您如何创建 OLAP 应用程序。

开始之前

开始之前，必须安装想要使用的教程中的章节所涉及的产品：

- 对于“数据入库”一节，必须安装“DB2 控制中心”，它包括“数据仓库中心”管理接口。您可在下列操作系统上安装“数据仓库中心”管理接口：Windows NT[®]、95、98、AIX[®]和“Solaris 操作环境”。还必须安装 DB2 服务器和仓库服务器，它们包括在“DB2 通用数据库[®]”的典型安装中。但是，必须将仓库服务器安装在 Windows NT 上。

如果将 DB2 服务器安装在不同于仓库服务器或“数据仓库中心”管理接口的工作站上，则必须将 DB2 CAE 安装在同一工作站上或安装在“数据仓库中心”管理接口上。

有关安装“DB2 通用数据库”的详情，参见您所使用的操作系统的 *DB2 通用数据库快速入门*。

如果您有“DB2 仓库管理程序”，则您可以选择安装“信息目录管理程序”。如果您没有“DB2 仓库管理程序”，则跳过第83页的『第14章 编目仓库中的数据以供最终用户使用』和第87页的『第15章 使用商业元数据』。

有关安装“DB2 仓库管理程序”的详情，请参阅 *DB2 仓库管理程序安装指南*。

- 对于“多维数据分析”一节，必须安装 DB2 和 OLAP Starter Kit。OLAP 客户机仅支持 Windows。

必须安装本教程。在 DB2 Windows 版中，您可将教程作为典型安装的一部分来进行安装。在 DB2 AIX 版或“Solaris 操作环境”版中，您可以将本教程与文档安装在一起。

您需要样本数据来配合教程使用。本教程使用“DB2 数据入库”样本数据和 OLAP 样本数据。

当您安装教程时，“数据入库”样本数据仅安装在 Windows NT 上。它必须安装在与仓库管理程序相同的工作站上，或样本数据库的远程节点必须在管理程序工作站上进行编目。

您可以在 Windows NT、AIX 和“Solaris 操作环境”上安装 OLAP 样本数据。它必须安装在与“OLAP 集成服务器”服务器相同的工作站上，或样本数据库的远程节点必须在服务器工作站上进行编目。

此教程引用了 X:\sqllib 目录下的几个样本数据，其中 X 是安装有 DB2 的驱动器。如果使用缺省目录结构，则数据安装在 X:\Program Files\sqllib 而不是 X:\sqllib 下。

在安装了用于样本的文件后，必须创建样本数据库。要创建该数据库：

1. 打开“第一步骤”窗口。
2. 单击**创建样本数据库**。
“创建样本数据库”窗口打开。
3. 选择“**数据入库**”样本校验框、**OLAP** 样本校验框，或两者，这取决于您想要使用教程的哪些部分。
4. 单击**确认**。
5. 如果您正在安装“数据入库”样本，会有一个窗口打开，用来输入用于存取该样本的 DB2 用户 ID 和口令。
 - a. 输入想要使用的用户 ID 和口令。记下用户 ID 和口令，因为在后面的课程中定义安全性时将需要它们。
 - b. 单击**确认**。

DB2 开始创建样本数据库。进度窗口打开。当创建了数据库时，单击**确认**。

如果是在 Windows NT 上安装样本，则将自动对 ODBC 注册数据库。如果是在 AIX 或“Solaris 操作系统”上安装样本，则必须人工地对 ODBC 注册数据库。有关在 AIX 或“Solaris 操作系统”上注册数据库的详情，参见您所使用的操作系统的 *DB2 通用数据库快速入门*。

如果选择了“数据入库”样本，则创建下列数据库：

DWCTBC

包含教程的“数据入库”一节所必需的操作源表。

TBC_MD

包含样本中“数据仓库中心”对象的元数据。

如果选择了 OLAP 样本，则创建下列数据库：

TBC 包含教程的“多维数据分析”一节所必需的提炼过的和已变换的表。

TBC_MD

包含样本中 OLAP 对象的元数据。

如果同时选择“数据入库”和 OLAP 样本，则 TBC_MD 数据库将同时包含样本中的“数据仓库中心”和 OLAP 对象的元数据。

在开始教程之前，验证是否可连接至样本数据库：

1. 启动“DB2 控制中心”：
 - 在 Windows NT 上，单击**开始—>程序—> DB2 Windows 版—>控制中心**。
 - 在 AIX 或 Sun Solaris 上，输入以下命令：

```
db2jstrt 6790
db2cc 6790b
```
2. 展开目录树，直到看到下列样本数据库之一为止：DWCTBC、TBC 或 TBC_MD。
3. 用鼠标右键单击该数据库的名称并单击**连接**。
“连接”窗口打开。
4. 在**用户 ID** 字段中，输入您用于创建样本的用户 ID。
5. 在**口令**字段中，输入用来创建样本的口令。
6. 单击**确认**。
“DB2 控制中心”连接至该数据库。

本教程中使用的约定

本教程在文本中使用印刷体约定来帮助您将控件名与您输入的文本区别开来。例如：

- 菜单项为粗体：
单击**菜单—>菜单选项**。
- 字段、校验框和按钮的名称也为粗体：
在**字段**字段中输入文本。
- 输入的文本是新行上的示例字体：
这是您输入的文本。

相关信息

本教程包括可使用“DB2 控制中心”、“数据仓库中心”和 OLAP Starter Kit 完成的最常见的任务。有关相关任务的详情，参见下列文档：

控制中心

- “DB2 控制中心”联机帮助
- “客户机配置辅助”联机帮助
- “事件监控程序”联机帮助
- 您的操作系统的 *DB2 通用数据库快速入门*
- *DB2 Warehouse Manager Installation Guide*

- *DB2 通用数据库 SQL 入门*
- *DB2 Universal Database SQL Reference*
- *DB2 通用数据库管理指南—实现*

数据仓库中心

- “数据仓库中心” 联机帮助
- *DB2 通用数据库数据仓库中心管理指南*

OLAP Starter Kit

- *OLAP Setup and User's Guide*
- *OLAP Model User's Guide*
- *OLAP Metaoutline User's Guide*
- *OLAP Administrator's Guide*
- *OLAP 电子表格附件用户指南 1-2-3 版*
- *OLAP 电子表格附件用户指南 Excel 版*

与 IBM 联系

如果有技术问题，请在与“DB2 客户支持中心”联系之前复查并执行 *Troubleshooting Guide* 所建议的操作。本指南对您可以收集哪些信息以使“DB2 客户支持中心”更好地为您服务提出了建议。

要获取信息或订购任何“DB2 通用数据库”产品，与当地分支机构的 IBM 代表联系，或与任何特许 IBM 软件经销商联系。

您如果住在美国，请致电下列其中一个号码：

- 1-800-237-5511，可获得客户支持
- 1-888-426-4343，可了解所提供的服务项目

产品信息

您如果住在美国，请致电下列其中一个号码：

- 1-800-IBM-CALL (1-800-426-2255) 或 1-800-3IBM-OS2 (1-800-342-6672)，可订购产品或获取一般信息。
- 1-800-879-2755，可订购出版物。

<http://www.ibm.com/software/data/>

DB2 万维网网页提供关于新闻、产品说明、培训计划等等的当前 DB2 信息。

<http://www.ibm.com/software/data/db2/library/>

“DB2 产品和服务技术库”可供您访问常见问题、修订、书籍以及最新的 DB2 技术资料。

注：此资料可能只有英文版。

<http://www.elink.ibm.com/pbl/pbl/>

International Publications Ordering Web 站点提供关于如何订购书籍的信息。

<http://www.ibm.com/education/certify/>

IBM Web 站点中的“专业认证程序”提供各种 IBM 产品（包括 DB2）的认证测试信息。

[ftp.software.ibm.com](ftp://software.ibm.com)

以匿名形式注册。可在目录 /ps/products/db2 中找到有关 DB2 和许多其他产品的演示程序、修订、信息和工具。

comp.databases.ibm-db2, bit.listserv.db2-l

这些 Internet 新闻组可供用户来讨论使用 DB2 产品的经验。

On Compuserve: GO IBMDB2

输入此命令来访问 IBM DB2 系列论坛。这些论坛支持所有的 DB2 产品。

有关如何在美国以外的地区与 IBM 联系的信息，参见 *IBM Software Support Handbook* 的附录 A。要存取此文档，访问以下 Web 页面：
<http://www.ibm.com/support/>，然后选择该页面底部附近的 IBM Software Support Handbook 链接。

注：在某些国家，IBM 特许经销商应与他们的经销商支持机构联系，而不是与“IBM 支持中心”联系。

第1部分 数据入库

第1章 有关数据入库

在本节中，您将获得此教程中的数据入库及数据入库任务的概述。

什么是数据入库？

包含操作数据（该数据运行日常的商业事务）的系统中包含对业务分析员有用的信息。例如，分析员可使用有关在一年中的哪个时间、哪个区域销售了哪种产品的信息来查找异常情况或规划未来的销售。但是，如果分析员直接存取操作数据，会有许多问题：

- 他们可能没有查询操作数据库的专门知识。例如，查询 IMS 数据库需要使用一种专门类型的数据操作语言的应用程序。一般说来，那些具有查询操作数据库专门知识的程序员担任着维护该数据库及其应用程序的专职工作。
- 对许多操作数据库（如银行的数据库）来说，性能非常重要。系统不能对执行特别查询的用户进行处理。
- 一般来说，操作数据的格式并非供业务分析员使用的最佳格式。例如，对分析员来说，按产品、地区和季节汇总的销售数据比原始数据有用得多。

数据入库解决了下列问题。在数据入库中，将创建资料性数据的存储库，该数据是从操作数据抽取的，然后变换为用于最终用户决策的形式。例如，数据入库工具可能从操作数据库复制所有的销售数据，执行计算以汇总数据，并将汇总数据写入与操作数据不同的数据库中。最终用户可查询该独立的数据库（仓库），而不会影响操作数据库。

课程概述

“DB2 通用数据库”提供了“数据仓库中心”，它是一个自动执行仓库处理的 DB2 部件。您可使用“数据仓库中心”来定义要包括在仓库内的数据。然后，可使用“数据仓库中心”来自动调度该仓库中的数据刷新。

本教程包括了设置一个仓库所需的最常用的任务。

在本教程中，您将：

- 定义一个主题区，它标识您将为教程创建的进程并对其进行分组。
- 开发源数据（它是操作数据）并定义仓库源。仓库源标识您想要在仓库中使用的源数据。

- 创建要用作仓库的数据库并定义仓库目标，这些仓库目标标识了要包括在仓库中的目标数据。
- 指定移动源数据并将其变换为用于仓库数据库的格式的方式。将定义一个进程，它包含从一个或多个源表、视图或文件生成仓库中的目标表所需一系列移动和变换的步骤。然后，将该进程分为几个步骤，每一步定义移动和变换进程中的一个操作。接着，将测试您定义的步骤，并调度它们来自动地运行。
- 通过定义安全性和监控数据库使用情况来管理仓库。
- 如果已经安装了“DB2 仓库管理程序”程序包，则创建仓库中数据的信息目录。信息目录是一个包含商业元数据的数据库，帮助用户在组织结构中识别并找出他们可用的数据和信息。该仓库的最终用户可搜索目录来确定要查询的那些表。
- 定义仓库中数据的星型模式模型。星型模式是由多个维度表（它描述业务的各个方面）和一个事实表（它包含有关该业务的事实）组成的专门设计。例如，如果您制造软饮料，则某些维度表为产品、市场和时间。事实表可能包含有关按季节划分的每个区域订购的产品的交易信息。
- 您可连结事实表和维度表以将维度表中的细节与订购信息组合在一起。例如，您可能将产品维与事实表连结在一起以添加有关如何将每个产品包装以符合订单要求的信息。