

普通高等教育




计算机系列规划教材

数据库管理 与数据分析

(SQL Server 2012+Tableau)

★ 郭进 吕峻闽 徐鸿雁 主编 ★

★ 陈婷 陈昌平 罗文佳 罗丹 副主编 ★

 中国工信出版集团



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
<http://www.phei.com.cn>



普通高等教育计算机系列规划教材

数据库管理与数据分析

(SQL Server 2012+Tableau)

郭 进 吕峻闽 徐鸿雁 主 编

陈 婷 陈昌平 罗文佳 罗 丹 副主编

电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

本书集数据管理和数据分析为一体,以微软公司最新数据库产品 SQL Server 2012 为数据平台,在简述传统数据库理论的基础上,首先认识数据,掌握数据库设计原理,将数据以关系化方式存储到数据库中,然后基于 SQL Server 2012 平台进行数据库管理操作,最后基于 Tableau 平台进行数据分析和展示,结合案例数据进行商业智能化分析,提供翔实可视化的智能分析报告。同时,本书设计练习实验,强化实践教学和综合应用,并给出案例数据的完整分析和系统开发过程,有利于读者对照学习。

本书清晰描述了数据库管理和数据分析的主要过程,重点和难点突出,注重对实际技能的培养。本书既可作为计算机应用和信息管理等相关专业(计算机科学与技术或信息管理与信息系统等专业)的教材,也可供数据库开发技术人员使用。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有,侵权必究。

图书在版编目(CIP)数据

数据库管理与数据分析:SQL Server 2012+Tableau/郭进,吕峻闽,徐鸿雁主编. —北京:电子工业出版社,2017.9
普通高等教育计算机系列规划教材

ISBN 978-7-121-32189-4

I. ①数… II. ①郭… ②吕… ③徐… III. ①关系数据库系统—可视化软件—高等学校—教材
IV. ①TP311.138

中国版本图书馆 CIP 数据核字(2017)第 161181 号

策划编辑:徐建军(xujj@phei.com.cn)

责任编辑:谭丽莎

印 刷:天津嘉恒印务有限公司

装 订:天津嘉恒印务有限公司

出版发行:电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

开 本:787×1092 1/16 印张:17.75 字数:454 千字

版 次:2017 年 9 月第 1 版

印 次:2018 年 1 月第 2 次印刷

定 价:45.00 元

凡所购买电子工业出版社图书有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系,联系及邮购电话:(010) 88254888, 88258888。

质量投诉请发邮件至 zlts@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式:(010) 88254570。

本书编委会成员

(按拼音排序)

陈昌平 陈 婷 陈 婷 陈小宁 高玲玲

龚轩涛 郭 进 何臻祥 黄纯国 靳紫辉

李长松 李 化 刘 强 罗 丹 罗文佳

吕峻闽 马 明 汤来锋 王 强 王书伟

魏雨东 夏钰红 肖 忠 徐鸿雁 杨大友

姚一永 银 梅 袁 勋 张诗雨

前言

Preface

纵观国内，大数据已经形成产业规模，并上升到国家战略层面，大数据技术和应用呈现纵深发展趋势。面向大数据的云计算技术、大数据计算框架等不断推出，新型大数据挖掘方法和算法大量出现，大数据新模式、新业态层出不穷，传统产业开始利用大数据实现转型升级。数据在纵横两向的普及和应用进一步体现了数据的价值。而数据存储与管理与数据分析为大数据的应用提供了必备的基础支撑，具体涉及数据的采集、筛选、存储、分析、展示，每部分都是在数据利用过程中的重要环节，通过数据分析和解释进一步呈现了数据本质，挖掘了数据价值，最后利用大数据处理技术对用户决策等提供依据。

本书主要提供数据科学的思维方法，积极探索对非 IT 专业大学生进行数据分析技能的培养途径，结合多年来的教学和项目应用实践经验，以 Microsoft 的 SQL Server 2012 数据库数据存储和管理为平台，结合兼具建模和展示多功能的 Tableau 工具进行后期分析，详略结合，重点突出，既汲取原有部分数据库教学资源，又增添了案例数据分析和展示的介绍，从内容和形式上有所创新。

全书共分为 14 章，内容涵盖数据库理论介绍和 SQL Server 2012 数据库产品的详细安装方法，数据库和数据表的创建、修改和查询，T-SQL 语言的使用方法、存储过程等高级应用。同时以 SQL Server 2012 为后台数据源，使用 Tableau 对案例数据进行详细的分析，根据数据特性，采用不同的分析方式和图表展示方法，以及地图和仪表盘，案例丰富，细节翔实，可供读者边学习边实践，以方便读者快速、全面地掌握数据分析技术。

本书由郭进、吕峻闽、徐鸿雁任主编，陈婷、陈昌平、罗文佳、罗丹任副主编并负责编写相应各章节。参加本书编写的还有陈婷、陈小宁、高玲玲、龚轩涛、何臻祥、靳紫辉、李化、李长松、刘强、马明、汤来锋、王强、王书伟、魏雨东、夏钰红、肖忠、姚一永、银梅、张诗雨等。同时，西南财经大学天府学院信息技术教学中心和现代技术中心的各位老师为本书提供了许多帮助，在此，编者对以上人员致以最诚挚的谢意！在编写本书的过程中参考了相关的图书和资料，在此也对这些资料的相关作者深表感谢。

为了方便教师教学，本书配有电子教学课件，请有此需要的教师登录华信教育资源网 (www.hxedu.com.cn) 注册后免费进行下载，如有问题可在网站留言板留言或与电子工业出版社联系 (E-mail: hxedu@phei.com.cn)。

由于编者水平有限，加之时间仓促，书中难免有不妥之处，敬请读者批评指正，以便在今后的修订中不断改进。

编者

目录

Contents

第 1 章 SQL Server 2012 简介	(1)
1.1 SQL Server 2012 数据库概述	(1)
1.1.1 SQL Server 2012 基本功能	(1)
1.1.2 SQL Server 2012 管理工具	(2)
1.2 数据库基础知识	(4)
1.2.1 数据库相关概念	(4)
1.2.2 数据库系统的体系结构	(5)
1.2.3 数据库技术的研究领域	(6)
第 2 章 数据库设计	(8)
2.1 数据库设计概述	(8)
2.1.1 数据库设计方法	(9)
2.1.2 数据库设计步骤	(9)
2.1.3 数据库三级模式	(11)
2.2 需求分析	(12)
2.2.1 需求分析任务	(12)
2.2.2 需求分析方法	(13)
2.3 概念结构模型设计	(15)
2.3.1 概念结构模型	(15)
2.3.2 概念结构模型设计方法	(16)
2.3.3 概念结构模型设计步骤	(17)
2.4 逻辑结构设计	(19)
2.4.1 E-R 图向关系模型的转换	(20)
2.4.2 关系模式的规范化	(20)
2.4.3 关系模式的改进	(21)
2.5 物理结构设计	(21)
2.5.1 物理结构设计内容	(21)
2.5.2 关系模式存取	(24)
2.5.3 评价物理结构	(24)

2.6	数据库实施	(25)
2.6.1	数据库的实施	(25)
2.6.2	数据库试运行	(25)
2.6.3	数据库的维护	(26)
第3章	数据库基本管理操作	(27)
3.1	数据库操作	(27)
3.1.1	创建数据库	(28)
3.1.2	修改数据库	(31)
3.1.3	删除数据库	(35)
3.2	数据表操作	(37)
3.2.1	创建表	(38)
3.2.2	修改表	(41)
3.2.3	删除表	(43)
3.3	数据更新	(44)
3.3.1	添加记录	(44)
3.3.2	修改记录	(46)
3.3.3	删除记录	(48)
3.4	单表查询	(50)
3.4.1	查询简单列	(50)
3.4.2	查询简单行	(54)
3.4.3	条件查询	(55)
3.4.4	聚合函数	(59)
3.4.5	分组查询	(60)
3.4.6	对查询结果排序	(61)
3.5	多表查询	(62)
3.5.1	连接查询	(62)
3.5.2	集合查询	(64)
3.5.3	嵌套查询	(66)
第4章	数据库管理	(69)
4.1	数据库恢复	(69)
4.1.1	数据库恢复概述	(69)
4.1.2	数据库故障类型	(69)
4.1.3	数据库恢复技术	(70)
4.1.4	数据库镜像	(71)
4.2	数据库并发控制	(71)
4.2.1	数据库并发控制概述	(71)
4.2.2	数据库活锁和死锁	(73)
4.3	数据库备份和还原	(75)
4.3.1	备份数据库	(75)
4.3.2	还原数据库	(79)

第 5 章 数据库应用	(83)
5.1 索引	(83)
5.1.1 索引概述	(83)
5.1.2 索引类型	(83)
5.1.3 创建索引	(84)
5.1.4 删除索引	(85)
5.2 视图	(86)
5.2.1 视图概述	(86)
5.2.2 创建视图	(86)
5.2.3 修改视图	(89)
5.2.4 使用视图	(90)
5.2.5 删除视图	(94)
5.3 存储过程	(95)
5.3.1 存储过程概述	(95)
5.3.2 使用存储过程	(95)
5.3.3 管理存储过程	(99)
5.4 触发器	(102)
5.4.1 触发器概述	(102)
5.4.2 创建触发器	(102)
5.4.3 管理触发器	(105)
5.5 事务	(107)
5.5.1 事务概述	(107)
5.5.2 事务定义	(107)
第 6 章 Tableau 简介	(110)
6.1 数据可视化基本概念	(110)
6.1.1 什么是数据可视化	(110)
6.1.2 数据可视化的应用	(110)
6.2 Tableau 介绍	(111)
6.2.1 Tableau 的产品简介	(111)
6.2.2 Tableau 的主要特性	(112)
6.3 Tableau 数据连接	(114)
6.3.1 文件数据源连接	(115)
6.3.2 数据库连接	(116)
第 7 章 Tableau 基础操作	(118)
7.1 排序	(118)
7.2 分层	(119)
7.3 分组	(122)
7.4 参数设置	(123)
7.5 可视化图表	(124)
7.5.1 条形图	(124)

7.5.2	线形图	(130)
7.5.3	饼图	(132)
7.5.4	复合图	(134)
7.5.5	嵌套条形图	(135)
7.5.6	热图	(138)
7.5.7	动态图	(140)
7.5.8	突显图	(142)
7.5.9	散点图	(144)
7.5.10	气泡图	(148)
7.5.11	甘特图	(150)
7.5.12	靶标图	(151)
7.5.13	瀑布图	(153)
7.5.14	直方图	(156)
7.5.15	帕累托图	(157)
第 8 章	绘制地图	(163)
8.1	地图简介	(163)
8.2	创建地图	(163)
8.2.1	分配地理角色	(163)
8.2.2	绘制符号地图	(165)
8.2.3	绘制填充地图	(167)
8.2.4	绘制混合地图	(168)
第 9 章	动态仪表板	(171)
9.1	仪表板	(171)
9.1.1	工作区	(171)
9.1.2	对象	(172)
9.1.3	布局	(173)
9.2	创建新的仪表板	(174)
9.2.1	新建仪表板	(174)
9.2.2	调整格式并添加内容	(175)
9.2.3	添加交互操作	(176)
第 10 章	客户数据分析案例	(180)
10.1	零售业客户细分	(180)
10.1.1	销售情况时间序列分析	(180)
10.1.2	客户偏好分析	(182)
10.1.3	区域消费分析	(182)
10.1.4	零售业客户细分仪表板	(183)
10.2	网站客户细分	(185)
10.2.1	访问量区域分析	(185)
10.2.2	访问时间序列分析	(186)
10.2.3	访问目标完成情况分析	(186)

10.2.4	网站客户细分仪表盘	(188)
10.3	游戏客户细分	(189)
10.3.1	客户属性分析	(189)
10.3.2	游戏类型细分	(191)
10.3.3	游戏进程分析	(192)
10.3.4	客户区域分布	(194)
10.3.5	游戏客户细分仪表盘	(196)
第 11 章	教育发展分析案例	(197)
11.1	教育发展分析指标	(197)
11.2	教育水平评估	(197)
11.2.1	学校教育水平评估	(197)
11.2.2	区域教育水平评估	(207)
第 12 章	销售数据分析案例	(214)
12.1	分析说明	(214)
12.2	帕累托图	(215)
12.3	区域产品利润率分析	(216)
12.4	物流时间与运输成本分析	(216)
12.5	利润额与运输成本对比	(217)
12.6	产品类别销量比较	(219)
12.7	区域利润额与销售额分析	(220)
12.8	区域利润额分析	(220)
12.9	平均物流时间分析	(221)
12.10	各省订单种类及数量分析	(222)
12.11	结论说明	(224)
第 13 章	篮球运动员数据分析案例	(225)
13.1	篮球运动员价值分析指标	(225)
13.2	篮球运动员技术价值评估	(225)
13.2.1	球员技术价值评估	(225)
13.2.2	球场位置球员技术价值评估	(233)
第 14 章	上机实验指导	(235)
14.1	实验指导一——SQL Server 2012 安装	(235)
14.2	实验指导二——数据库设计项目	(248)
14.3	实验指导三——数据库操作及 SQL 命令	(248)
14.4	实验指导四——表操作及 SQL 命令	(250)
14.5	实验指导五——数据更新及 SQL 命令	(251)
14.6	实验指导六——单表查询及 SQL 命令	(253)
14.7	实验指导七——多表查询及 SQL 命令	(254)
14.8	实验指导八——数据库备份和还原	(255)
14.9	实验指导九——索引	(255)
14.10	实验指导十——视图	(256)

14.11	实验指导十一——存储过程	(259)
14.12	实验指导十二——触发器	(260)
14.13	实验指导十三——VB/SQL 数据库开发	(262)
14.14	实验指导十四——Tableau 安装和数据连接	(264)
14.15	实验指导十五——Tableau 基本图表一	(264)
14.16	实验指导十六——Tableau 基本图表二	(265)
14.17	实验指导十七——Tableau 仪表盘设计	(265)
附录 A	SQL 命令查询	(266)

第1章

SQL Server 2012 简介

1.1 SQL Server 2012 数据库概述

1.1.1 SQL Server 2012 基本功能

SQL Server 是由 Microsoft 开发和推广的关系数据库管理系统 (DBMS)，它最初是由 Microsoft、Sybase 和 Ashton-Tate 三家公司共同开发的，并于 1988 年推出了第一个 OS/2 版本。Microsoft SQL Server 近年来不断更新版本，1996 年，Microsoft 推出了 SQL Server 6.5 版本；1998 年，SQL Server 7.0 版本和用户见面；SQL Server 2000 是 Microsoft 公司于 2000 年推出的；2012 年 3 月推出了 SQL Server 2012；目前最新版本为 SQL Server 2016，于 2016 年 6 月发布。考虑到资源占用和教学使用的有限功能，本书采用 SQL Server 2012 版本。

2012 年 3 月 7 日消息，Microsoft 正式发布最新的 SQL Server 2012 RTM (Release-to-Manufacturing) 版本，面向公众的版本将于 4 月 1 日发布。Microsoft 此次版本发布的口号是用“大数据”来替代“云”的概念，Microsoft 对 SQL Server 2012 的定位是帮助企业处理每年大量的数据 (Z 级别) 增长。

来自 Microsoft 商业平台事业部的副总裁 Ted Kummert 称：SQL Server 2012 更加具备可伸缩性、可靠性及前所未有的高性能；而 Power View 为用户对数据的转换和勘探提供强大的交互操作能力，并协助做出正确的决策。即将推出 3 个主要版本和很多新特征，同时 Microsoft 也透露了 SQL Server 2012 的价格和版本计划，其中增加一个新的智能商业包。

SQL Server 2012 的主要版本包括新的商务智能版本，增加了 Power View 数据查找工具和数据质量服务，企业版本则提高安全性、可用性。从大数据到 StreamInsight 复杂事件处理，再到新的可视化数据和分析工具等，都将成为 SQL Server 2012 最终版本的一部分。

SQL Server 2012 的新增功能如下。

(1) AlwaysOn: 这个功能将数据库的镜像提到了一个新的高度。用户可以针对一组数据

库而不是一个单独的数据库做灾难恢复。

(2) Windows Server Core 支持: Windows Server Core 是命令行界面的 Windows, 使用 DOS 和 PowerShell 来做用户交互。它的资源占用更少, 更安全, 支持 SQL Server 2012。

(3) Columnstore 索引: 这是 SQL Server 独有的功能。它们是为数据仓库查询设计的只读索引。数据被组织成扁平化的压缩形式存储, 极大地减少了 I/O 和内存使用。

(4) 自定义服务器权限数据库管理员 (DBA) 可以创建数据库的权限, 但不能创建服务器的权限。例如, DBA 想要一个开发组拥有某台服务器上所有数据库的读写权限, 他必须手动完成这个操作。但是 SQL Server 2012 支持针对服务器的权限设置。

(5) 增强的审计功能: 所有的 SQL Server 版本都支持审计。用户可以自定义审计规则, 记录一些自定义的时间和日志。

(6) BI 语义模型: 这个功能是用来替代 “Analysis Services Unified Dimensional Model” 的。这是一种支持 SQL Server 所有 BI 体验的混合数据模型。

(7) Sequence Objects: 使用 Oracle 的人一直想要这个功能。一个序列 (sequence) 就是根据触发器的自增值而实现的。SQL Server 有一个类似的功能, identity columns, 但是用对象实现了。

(8) 增强的 PowerShell 支持: 所有的 Windows 和 SQL Server 管理员都应该认真地学习 PowerShell 的技能。Microsoft 正在大力开发服务器端产品对 PowerShell 的支持。

(9) 分布式回放 (Distributed Replay): 这个功能类似于 Oracle 的 Real Application Testing 功能。不同的是 SQL Server 企业版自带了这个功能, 而使用 Oracle 时, 你还需额外购买这个功能。这个功能可以让你记录生产环境的工作状况, 然后在另外一个环境重现这些工作状况。

(10) PowerView: 这是一个强大的自主 BI 工具, 可以让用户创建 BI 报告。

(11) SQL Azure 增强: 这和 SQL Server 2012 没有直接关系, 但是 Microsoft 确实对 SQL Azure 做了一个关键改进, 如 Reprint Service, 备份到 Windows Azure。Azure 数据库的上限提高到了 150G。

(12) 大数据支持: 这是最重要的一点, 虽然放在了最后。在 PASS (Professional Association for SQL Server) 会议上, Microsoft 宣布了与 Hadoop 的提供商 Cloudera 的合作。提供了 Linux 版本的 SQL Server ODBC 驱动。主要的合作内容是微软开发 Hadoop 的连接器, 也就是 SQL Server 也跨入了 NoSQL 领域。

1.1.2 SQL Server 2012 管理工具

1. SQL Server 管理环境

SQL Server 管理环境 (SQL Server Management Studio, SSMS) 是用于访问、配置、控制、管理和开发 SQL Server 2012 所有组件的一种集成环境, 它可以管理从 SQL 服务器到 SQL 数据库之间所有的基本操作。同时, SSMS 将大量的图形工具和丰富的脚本编辑器组合在一起, 实现了各种技术级别的开发人员和管理员对 SQL Server 的访问。

SSMS 将早期版本的 SQL Server 中所包含的企业管理器、查询分析器和 Analysis Manager 功能整合到单一的环境中。除此之外, SSMS 还可以协同 SQL Server 的所有组件一起工作, 如 Integration Services、Reporting Services 和 SQL Server Compact 3.5 SP1。数据库管理员可以获得功能齐全的统一实用工具, 开发人员也可以获得熟悉的体验。

2. SQL Server 2012 配置管理器

SQL Server 2012 配置管理器为 SQL Server 服务、服务器协议、客户端协议和客户端别名提供基本的配置管理。它可以通过执行“配置工具”→“SQL Server 2012 配置管理器”打开，如图 1-1 所示，也可以通过在命令提示下输入 sqlservemanager.msc 命令打开。



图 1-1 打开 SQL Server 2012 配置管理器

3. SQL Server Profiler

SQL Server Profiler 提供了一个用于 SQL 跟踪的图形用户界面，用于监视数据库引擎实例或 Analysis Services 实例。它可以帮助捕获关于每个数据库事件的数据，并将其保存到文件或表中，以供日后分析。SQL Server Profiler 分析界面如图 1-2 所示。

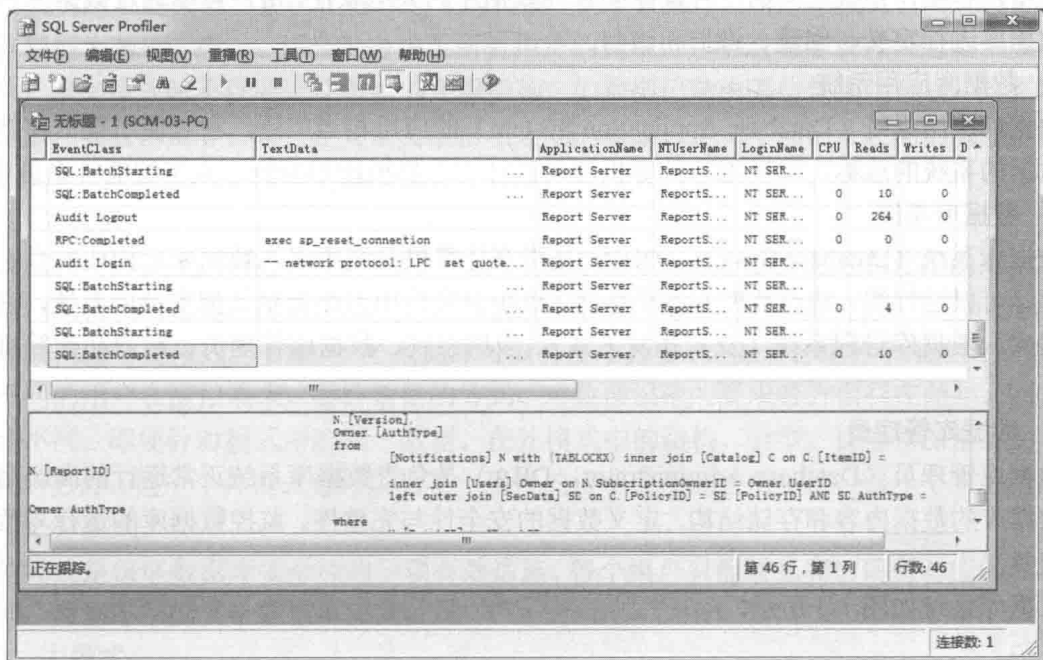


图 1-2 SQL Server Profiler 分析界面

4. 数据库引擎优化顾问

数据库引擎优化顾问 (Database Engine Tuning Advisor) 是协助用户创建索引、索引视图和分区的最佳组合, 它可以帮助用户分析工作负荷、提出创建高效率索引的建议等。

5. 命令提示实用工具

除了上述的图形化管理工具, SQL Server 2012 还提供了可以从命令提示符中运行的工具, 如 sqlcmd.exe、osql.exe、bcp.exe、dtexec.exe、dtutil.exe、rsconfig.exe、sqlwb.exe、tablediff.exe 等。

1.2 数据库基础知识

1.2.1 数据库相关概念

1. 数据

数据 (Data) 是数据库中存储的基本对象, 是人们用来描述信息的可识别的符号。数据具有多种表现形式, 它与传统意义上理解的数据不同, 可以是数字、文字、图形、图像、声音、动画等, 它们都可以经过数字化后存入计算机。

2. 数据库

数据库 (Data Base, DB) 是长期存储在计算机内的、有组织的, 并且可以共享的大量数据的集合, 它将数据按照一定的数据模型组织、描述和存储, 具有较小冗余度、较高数据独立性和易扩展性、可以被各种用户共享等特点。

3. 数据库管理系统

数据库管理系统 (Database Management System, DBMS) 对数据库进行统一的管理和控制, 是位于用户和操作系统之间的一种操纵和管理数据库的大型软件, 用户可以通过数据库管理系统对数据库进行定义、创建、维护和访问。

4. 数据库应用系统

数据库应用系统 (Database Application System) 的应用相当广泛, 它是使用数据库技术管理其数据的系统的总称, 可以用于计算机辅助设计、计算机图形分析、人工智能等系统中。

5. 数据库系统

数据库系统 (Database System, DBS) 是指在计算机系统中引入数据库后构成的系统, 是一个实际可以运行的, 按照数据库方法存储、维护并向应用系统提供数据支持的系统, 它一般由数据库、数据库管理系统 (及其开发工具)、应用系统、数据库管理员、用户和计算机硬件构成。

6. 数据库管理员

数据库管理员 (Database Administrator, DBA) 是负责数据库系统正常运行的高级用户, 决定数据库的数据内容和存储结构, 定义数据的安全性与完整性, 监控数据库的运行与数据的重组恢复。

数据库系统如图 1-3 所示。

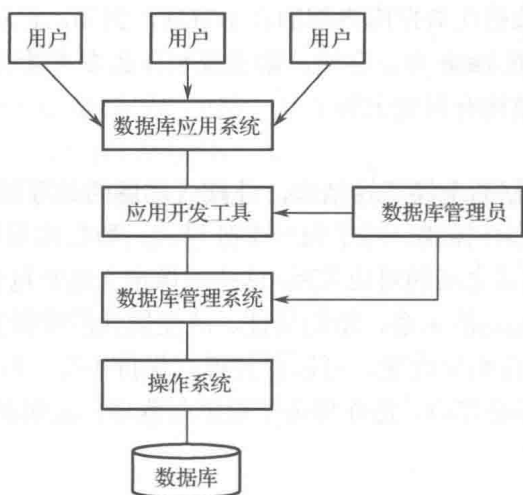


图 1-3 数据库系统

1.2.2 数据库系统的体系结构

数据库系统的体系结构分为三级模式与二级映像。

数据库系统的三级模式结构是指数据库系统由外模式、模式和内模式三级构成。数据库管理系统在这三级模式之间提供了两级映像：外模式/模式映像和模式/内模式映像。

1. 模式

模式 (Schema) 也称概念模式，是对数据库中全部数据的逻辑结构和特征的描述，是所有用户的公共数据视图。它是数据库系统模式结构的中间层，既不涉及数据的物理存储细节和硬件环境，也不涉及具体的应用程序及所使用的应用开发工具和高级程序设计语言。

模式实际上是数据库数据在概念级上的视图，一个数据库只有一个模式。模式通常以某一种数据模型为基础，统一、综合地考虑了所有用户的需求，并将这些需求有机地结合成一个逻辑整体。定义模式时不仅要定义数据的逻辑结构，如数据记录由哪些数据项构成，数据项的名称、类型、取值范围等，而且还要定义数据项之间的联系，定义不同记录之间的联系，以及定义与数据有关的完整性、安全性等要求。

2. 外模式

外模式 (External Schema) 也称用户模式，它是数据库用户能够看见和使用的局部数据的逻辑结构和特征的描述，是数据库用户的数据视图，即个别用户涉及的数据的逻辑结构。

外模式通常是模式的子集，一个数据库可以有多个外模式。由于它是各个用户的数据视图，如果不同的用户在应用需求、看待数据的方式、对数据保密的要求等方面存在差异，则其外模式描述不同。即使针对模式中的同一数据，在外模式中的结构、类型、长度和保密级别等都可以不同。另外，同一个外模式也可以为某一用户的多个应用系统所使用，但一个应用程序只能使用一个外模式。

外模式是保证数据库安全性的一项有效措施。每个用户只能看见和访问所对应的外模式中的数据，数据库中的其余数据不可见。

3. 内模式

内模式 (Internal Schema) 也称存储模式，一个数据库只有一个内模式，它是数据物理结

构和存储方式的描述,是数据在数据库内部的表示方式。例如,记录的存储方式是顺序存储、按照 B 树结构存储还是按照 hash 方法存储;索引按照什么方式组织;数据是否压缩存储,是否加密;数据的存储记录结构有何规定等。

4. 外模式/模式映像

外模式/模式映像是数据的全局逻辑结构,外模式描述的是数据的局部逻辑结构,对应于同一个模式可以有任意多个外模式。对于每一个外模式,数据库系统都有一个外模式/模式映像,它定义了该外模式与模式之间的对应关系,这些映像定义通常包含在各自外模式的描述中。

当模式改变时(如增加新的关系、新的属性、改变属性的数据类型等),由数据库管理员对各个外模式/模式映像进行相应改变,可以使外模式保持不变。应用程序是根据数据的外模式编写的,因此应用程序不必修改,进而保证了数据与程序的逻辑独立性,简称数据的逻辑独立性。

5. 模式/内模式映像

数据库中只有一个模式,也只有一个内模式,因此模式/内模式映像是唯一的,它定义了数据全局逻辑结构与存储结构之间的对应关系。例如,它说明逻辑记录和字段在内部是如何表示的。该映像定义通常包含在模式描述中。当数据库的存储结构改变(如选用了另一种存储结构)时,由数据库管理员对模式/内模式映像进行相应改变即可,可以使模式保持不变,从而使得应用程序也不必改变,保证了数据与程序的物理独立性,简称数据的物理独立性。

在数据库的三级模式结构中,数据库模式即全局逻辑结构是数据库的中心与关键,它独立于数据库的其他层次。因此,设计数据库模式结构时应首先确定数据库的逻辑模式。

数据库的内模式依赖于它的全局逻辑结构,但独立于数据库的用户视图即外模式,也独立于具体的存储设备。它将全局逻辑结构中所定义的数据结构及其联系按照一定的物理存储策略进行组织,以达到较好的时间效率和空间效率。

数据库的外模式面向具体的应用程序,它定义在逻辑模式之上,但独立于存储模式和存储设备。当应用需求发生较大变化,相应外模式不能满足其视图要求时,该外模式就得进行相应改动,因此设计外模式时应充分考虑到应用的扩充性。

1.2.3 数据库技术的研究领域

数据库学科的研究范围主要包括以下三个领域。

1. 数据库管理系统软件的研制

数据库管理系统是数据库系统的基础。数据库管理系统的研制包括研制数据库管理系统本身及以数据库管理系统为核心的一组相互联系的软件系统,包括工具软件和中间件。研制的目标是提高系统的性能和用户的生产率。

2. 数据库设计

数据库设计的研究包括以下内容:

- (1) 数据库的设计方法、设计工具和设计理论的研究;
- (2) 数据模型和数据建模的研究;
- (3) 计算机辅助数据库设计及其软件系统的研究;
- (4) 数据库设计规范和研究等。