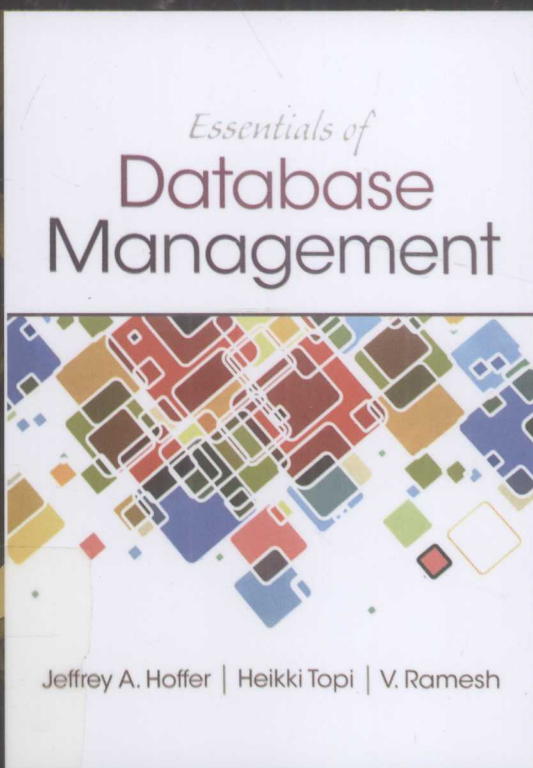


PEARSON

数据库管理 基础教程

杰弗里 A. 霍弗 (Jeffrey A. Hoffer)
[美] 海基·托皮 (Heikki Topi) 著 岳丽华 张怡文 等译
拉梅什·文卡塔拉曼 (Ramesh Venkataraman)

Essentials of Database Management



机械工业出版社
China Machine Press

计 算 机 科 学 丛 书

数据库管理 基础教程

杰弗里 A. 霍弗 (Jeffrey A. Hoffer)

[美] 海基·托皮 (Heikki Topi)

著 岳丽华 张怡文 等译

拉梅什·文卡特拉曼 (Ramesh Venkataraman)

Essentials of Database Management

Database Management

Jeffrey A. Hoffer | Heikki Topi | V. Ramesh

00991639



机械工业出版社
China Machine Press

图书在版编目 (CIP) 数据

数据库管理基础教程 / (美) 霍弗 (Hoffer, J. A.), (美) 托皮 (Topi, H.), (美) 拉梅什 (Ramesh, V.) 著; 岳丽华等译. —北京: 机械工业出版社, 2016.1

(计算机科学丛书)

书名原文: Essentials of Database Management

ISBN 978-7-111-52623-0

I. 数… II. ①霍… ②托… ③拉… ④岳… III. 数据库系统 - 教材 IV. TP311.13

中国版本图书馆 CIP 数据核字 (2016) 第 010376 号

本书版权登记号: 图字: 01-2013-6489

Authorized translation from the English language edition, entitled *Essentials of Database Management*, 978-0-13-340568-2 by Jeffrey A. Hoffer, Heikki Topi, V. Ramesh, published by Pearson Education, Inc., Copyright © 2014.

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage retrieval system, without permission from Pearson Education, Inc.

Chinese simplified language edition published by Pearson Education Asia Ltd., and China Machine Press Copyright © 2016.

本书中文简体字版由 Pearson Education (培生教育出版集团) 授权机械工业出版社在中华人民共和国境内 (不包括中国台湾地区和中国香港、澳门特别行政区) 独家出版发行。未经出版者书面许可, 不得以任何方式抄袭、复制或节录本书中的任何部分。

本书封底贴有 Pearson Education (培生教育出版集团) 激光防伪标签, 无标签者不得销售。

本书是在已出版了 11 版的《现代数据库管理》(Modern Database Management, MDM) 教材基础之上, 为满足那些不需要深入讨论数据库技术高级内容的课程而编写的数据库管理导论教材。全书共分四部分, 第一部分 (第 1 章) 介绍数据库环境和开发过程, 第二部分 (第 2、3 章) 介绍组织中的数据建模和增强的 E-R 模型, 第三部分 (第 4、5 章) 介绍逻辑数据库设计和关系模型、物理数据库设计和性能, 第四部分 (第 6 ~ 9 章) 介绍 SQL、高级 SQL、数据库应用开发和数据仓库。此外, 为便于读者学习, 还提供了大量难度不同的经过测试的复习题、问题和实践材料等。

本书适合作为工商学院、信息学院等相关专业数据库导论课程的教材。

出版发行: 机械工业出版社 (北京市西城区百万庄大街 22 号 邮政编码 100037)

责任编辑: 迟振春

责任校对: 董纪丽

印刷: 北京文昌阁彩色印刷有限责任公司

版次: 2016 年 3 月第 1 版第 1 次印刷

开本: 185mm × 260mm 1/16

印张: 24.5

书号: ISBN 978-7-111-52623-0

定价: 79.00 元

凡购本书, 如有缺页、倒页、脱页, 由本社发行部调换

客服热线: (010) 88378991 88361066

投稿热线: (010) 88379604

购书热线: (010) 68326294 88379649 68995259

读者信箱: hzsj@hzbook.com

版权所有·侵权必究

封底无防伪标均为盗版

本书法律顾问: 北京大成律师事务所 韩光 / 邹晓东

文艺复兴以来，源远流长的科学精神和逐步形成的学术规范，使西方国家在自然科学的各个领域取得了垄断性的优势；也正是这样的优势，使美国在信息技术发展的六十多年间名家辈出、独领风骚。在商业化的进程中，美国的产业界与教育界越来越紧密地结合，计算机学科中的许多泰山北斗同时身处科研和教学的最前线，由此而产生的经典科学著作，不仅肇划了研究的范畴，还揭示了学术的源变，既遵循学术规范，又自有学者个性，其价值并不会因年月的流逝而减退。

近年，在全球信息化大潮的推动下，我国的计算机产业发展迅猛，对专业人才的需求日益迫切。这对计算机教育界和出版界都既是机遇，也是挑战；而专业教材的建设在教育战略上显得举足轻重。在我国信息技术发展时间较短的现状下，美国等发达国家在其计算机科学发展的几十年间积淀和发展的经典教材仍有许多值得借鉴之处。因此，引进一批国外优秀计算机教材将对我国计算机教育事业的发展起到积极的推动作用，也是与世界接轨、建设真正的世界一流大学的必由之路。

机械工业出版社华章公司较早意识到“出版要为教育服务”。自1998年开始，我们就将工作重点放在了遴选、移译国外优秀教材上。经过多年的不懈努力，我们与Pearson, McGraw-Hill, Elsevier, MIT, John Wiley & Sons, Cengage等世界著名出版公司建立了良好的合作关系，从他们现有的数百种教材中甄选出Andrew S. Tanenbaum, Bjarne Stroustrup, Brian W. Kernighan, Dennis Ritchie, Jim Gray, Alfred V. Aho, John E. Hopcroft, Jeffrey D. Ullman, Abraham Silberschatz, William Stallings, Donald E. Knuth, John L. Hennessy, Larry L. Peterson等大师名家的一批经典作品，以“计算机科学丛书”为总称出版，供读者学习、研究及珍藏。大理石纹理的封面，也正体现了这套丛书的品位和格调。

“计算机科学丛书”的出版工作得到了国内外学者的鼎力相助，国内的专家不仅提供了中肯的选题指导，还不辞劳苦地担任了翻译和审校的工作；而原书的作者也相当关注其作品在中国的传播，有的还专门为其书的中译本作序。迄今，“计算机科学丛书”已经出版了近百个品种，这些书籍在读者中树立了良好的口碑，并被许多高校采用为正式教材和参考书籍。其影印版“经典原版书库”作为姊妹篇也被越来越多实施双语教学的学校所采用。

权威的作者、经典的教材、一流的译者、严格的审校、精细的编辑，这些因素使我们的图书有了质量的保证。随着计算机科学与技术专业学科建设的不断完善和教材改革的逐渐深化，教育界对国外计算机教材的需求和应用都将步入一个新的阶段，我们的目标是尽善尽美，而反馈的意见正是我们达到这一终极目标的重要帮助。华章公司欢迎老师和读者对我们的工作提出建议或给予指正，我们的联系方式如下：

华章网站：www.hzbook.com

电子邮件：hzjsj@hzbook.com

联系电话：(010) 88379604

联系地址：北京市西城区百万庄南街1号

邮政编码：100037



华章教育

华章科技图书出版中心

当今大数据时代，数据处理是各行各业须臾不可或缺的重要需求。数据库技术作为 20 世纪 60 年代发展起来的数据管理与处理技术，仍然是该领域非常重要的技术。目前，国内大学几乎所有专业都有关于数据库内容的课程。但是，除了计算机科学与技术 and 软件工程专业需要将数据库作为系统软件学习外，一般信息处理专业更关注数据建模和数据库应用技术。Jeffrey A. Hoffer 等人编著的这本书不失为一本较好满足更多信息处理专业需求的教材。

本书是作者在已出版了 11 版的《现代数据库管理》(Modern Database Management, MDM) 教材基础之上，为满足那些不需要深入讨论数据库技术高级内容的课程而编写的数据库管理导论教材。这本新教材不是对《现代数据库管理》教材的简单裁减，而是吸收了 30 年来使用《现代数据库管理》教材的教师、学生的经验和建议后重新编写的，特别适合工商学院、管理学院、信息学院、计算机技术专业等使用。该教材不仅包括传统数据库系统课程关注的数据库设计和 SQL 主题，而且介绍了数据库应用开发和数据仓库主题，更通过实例强调了现代信息系统中数据组成的发展，以及数据资源的管理，并对数据建模给出了更多的叙述和讨论。与已有的数据库教材比较，该书每一章后面不仅有与教学内容相关的复习题，还提供了与实际组织、企业信息处理需求相关的丰富多彩的问题与练习，这些对于学生利用课程内容解决实际问题有很好的帮助。

本书主要由岳丽华和张怡文负责翻译和审校，参加翻译的还有张晓翔、王倩、李璐、桑若新、徐娇等。

限于水平，译文中难免有错误与不足之处，欢迎读者批评指正。

译者

2015 年 12 月

很高兴向读者介绍这本新的有关数据库管理的教材。这本书从概念和技术上讲述了大多数数据和数据库管理课程的核心内容。我们特别关注数据库开发生命周期中较深入的问题，从使用（增强的）实体-联系模型的概念数据建模开始，经过逻辑层面的关系建模，到使用结构化查询语言（SQL）的数据库实现。基于最新的教学法和数据管理方向的技术发展，我们逐步深入地介绍这个核心领域。

本书以《现代数据库管理》（Modern Database Management, MDM，目前该书已出版到第 11 版）为基础，去掉了其中深入和广泛讨论的高级内容，目标是作为数据库管理导论课程教材。实现这个目标并不意味本书不注重质量和严密性。使本书更加成功的几个外部因素是：特别是在工商学校，其课程体系中有较少的信息系统课程，这也就意味着所有独立领域的论题（包括数据库管理）都必须丢弃一些内容。数据库导论课程的非专业比例在上升，那些核心内容之外的高级技术知识对他们来说没有意义。很多综合课程中数据库模块的授课时数不足一个完整学期，本教材非常适合这样的课程使用。

与市场上的数据库教材比较，本教材具有如下优点：

- 建立在不断发展的领先教材（《现代数据库管理》）的坚实基础之上，讨论了最新的数据库管理核心议题。
- 所有内容都保持了概念的严密性。
- 非常敏锐地关注了集成的数据库开发周期。
- 为专业设计者和开发者（而不是终端用户）提供了工具和技术。
- 提供了大量难度不同的经过测试的复习题和实践材料。
- 结合了作者多年在不同大学的教学经验以及在主导产业和课程组的研究成果。

本教材特别适合作为工商学院、信息学院、计算机技术专业及应用计算机科学系的信息系统或信息技术课程体系的一部分使用。信息系统协会（Association for Information Systems, AIS）、计算机协会（Association for Computing Machinery, ACM）和信息处理协会国际联盟（International Federation of Information Processing Societies, IFIPS）的课程体系指南中都列出了这种类型的数据库管理课程。例如，本书涵盖了 IS 2010^① 建议的绝大多数数据库核心内容，IS 2010 是最新的信息系统本科生课程体系。另外，除了 4 年制的本科专业外，该教材还可以满足社区学院信息和计算机技术专业教学大纲和研究生基础课程中数据管理模块的需求。如上面所讨论的，本书也适用于非 IS 专业学生，这些学生的兴趣是广泛的信息利用领域。例如，该教材可以为研究生或本科生层次的商业分析专业学生奠定良好的数据管理课程基础。

本书特色

- 本书覆盖了最新的原理、概念和技术，系统、全面、详细地介绍了关键主题。从《现

① Topi, Heikki; Valacich, Joseph S.; Wright, Ryan T.; Kaiser, Kate; Nunamaker, Jr., Jay F.; Sipior, Janice C.; and de Vreede, Gert Jan (2010) "IS 2010: Curriculum Guidelines for Undergraduate Degree Programs in Information Systems," Communications of the Association for Information Systems: Vol. 26, Article 18.

代数据库管理》一书中获得的经验和专门知识,使得本书虽是第1版,但有很坚实的基础。数据库开发周期核心内容是数据库环境,以及与组织数据管理相关的上下文主题,如基于互联网应用的数据库开发、现代数据仓库概念、新的基础设施技术(如云计算)。

- 本书重点介绍领先实践者所说的对于数据库开发者和资深数据库用户来说最重要的内容。我们与很多实践者一起工作,包括数据管理协会(Data Management Association, DAMA)和数据仓库研究所(The Data Warehousing Institute, TDWI)的专家、高级咨询师、技术主管以及具有广泛读者的专业出版物的专栏作家。我们提及这些专家是为了保证本书所覆盖的内容都是重要的,不仅包括重要的入门知识和技能,而且包括引领职业生涯成功的基本原则和思维模式。
- 本书是从方便学生学习的角度组织编写的,书中的内容、复习题、问题与练习都直接得益于MDM 30年来持续的市场反馈。总而言之,本书的教学法是合理的,我们使用了很多图来清晰地说明重要的概念和技术,并且使用了最新的符号。本书的组织是灵活的,因此可以按照学生的情况安排书中章节的顺序。我们还为本书增补了可以动手实践学习的数据集,并使用新的媒体资源使那些更有挑战的议题更具吸引力。
- 你可能乐意在课程中较早地讲授SQL,该教材适合这样做。首先,第6章和第7章逐渐深入地介绍SQL这一数据库领域的核心技术。其次,在前面的章节中包含了很多SQL例子。再次,已有很多教师在较早的课程中成功地使用了这两章SQL内容。虽然逻辑上这两章是本书实现部分,但很多教师是在第1章之后就开始讲授这两章的内容,或者并行地与前面其他几章一起讲授。最后,SQL是本书通篇使用的语言,例如,第8章中关于Web应用与关系数据库的联系的讲解以及第9章中的联机分析处理。
- 本书有最新的内容补充和配套网站。
- 本书是为现代信息系统课程体系的一部分编写的,重点关注商业系统开发。书中所包含的内容和讨论的议题都是对其他典型课程中原理的加强,如系统分析和设计、网络、Web站点设计与开发、MIS原理以及计算机程序设计等。本书更强调现代信息系统中数据库组件的开发以及数据资源的管理。因此,本书理论与实践结合,支持课程项目和其他实际动手的课堂练习,鼓励学生将贯穿专业课程体系所学的概念与数据库概念相关联。本书的两位作者是发展全球信息系统本科生和研究生层面样本课程体系的主要参与者,他们在本书中体现了该课程体系的意图。

与《现代数据库管理》之间的主要区别

这两本书之间的最主要区别是,MDM第10~14章和附录A~C的内容都不包括在本教材中。但是一些关键内容,如数据质量和集成(MDM第10章)以及数据和数据库管理(MDM第11章)在本书第1~9章中有介绍。分布式数据库(MDM第12章)、面向对象数据建模(MDM第13章)以及利用关系数据库提供对象持久性(MDM第14章)在本书中完全没有涉及。另外,第1~9章按照更好地满足本书读者需求的方式组织,包括如下几点改变:

- 缩减了如下内容:
 - 业务规则

- 时间相关数据建模
- 实时数据仓库
- 缓变维
- SQL 联机分析处理查询
- 删除了如下内容：
 - 实体簇
 - 超过 3NF 的范式
 - 企业密钥
 - 数据卷和使用分析
 - 分区
 - 簇文件
 - 特定的专用 SQL 命令
 - 视图上的某些特有内容
 - 某个数据库管理系统上的专有内容
 - PHP
 - 无事实的事实表

我们的重要意图是让 MDM 和本教材同时持续发展，作为课本它们是紧密相连的，但是又各有清晰的思路和目标读者。

章节安排

下面分章给出每一章的目标及其核心内容。

第一部分 数据库管理的上下文

第 1 章 数据库环境和开发过程

本章讨论组织中数据库的作用 (role)，简要介绍本书其他章节的主要内容。在对数据存储和检索相关的基本术语进行简要介绍后，将传统的文件处理系统和现代数据库技术做了很好的比较。然后，介绍数据库环境的核心组成，以及数据库应用的范围，这些应用都是当前组织正在使用的，如个人、二层、多层和企业级应用。企业级数据库的描述包括已经成为部分企业资源规划系统和数据仓库的数据库。数据库技术的演化，从早期数据库文件到现代的对象关系技术，都有所呈现。接下来，在结构化生命周期、原型系统、敏捷方法学上下文中讨论数据库开发过程，表达方式与 Hoffer、George 和 Valacich 编写的系统分析教材保持一致。本章还讨论了数据库发展中的重要问题和理解数据库体系结构与技术的框架（包括三级模式体系结构）。审阅者通常认为该章对于学生学习系统分析和设计课程很有益处。

第二部分 数据库分析

第 2 章 组织中的数据建模

本章介绍用实体-联系 (E-R) 模型实现概念数据建模。章名强调了选择实体-联系模型的理由：那些影响数据库设计的业务规则要无二义性地记录下来。多个小节详细地解释了

如何命名和定义数据模型的元素，这是无二义开发 E-R 图的基本要求。本章中的例子从简单到复杂，最后给出一个描述 Pine Valley 家具公司的综合 E-R 图。

第 3 章 增强型 E-R 模型

本章讨论几个高级 E-R 数据模型结构。特别地，全面讨论了超类型 / 子类型联系，并且给出了一个描述 Pine Valley 家具公司的扩展 E-R 数据模型的完整例子。

第三部分 数据库设计

第 4 章 逻辑数据库设计和关系模型

本章描述将概念数据模型转换成关系数据模型的过程，讨论如何将新关系合并到已有的规范化数据库中，提供了规范化的概念基础和实践过程，并且强调了在规范化中使用函数依赖和决定因子的重要性。本章还讨论了外键，强调了关系数据模型的基本概念，以及在逻辑设计过程中数据库设计者的任务。

第 5 章 物理数据库设计和性能

本章描述完成一个有效的数据库设计的基本步骤，重点放在数据库设计和实现方面，这些都是现代数据库环境中需要数据库专家专门控制的。本章另一个重点是改进数据库性能，特别是参照 Oracle 和其他数据库管理系统 (DBMS) 中的技术改进数据库处理性能。不同的索引类型是数据库技术中提高查询处理速度的技术。

第四部分 实现

第 6 章 SQL 导论

本章介绍在大多数 DBMS 中使用的 SQL (SQL:1999)，同时还介绍最新版 SQL (SQL:2008) 中的一些变动。本章和下一章都是有关 SQL 的内容。本章包括 SQL 代码举例，大多使用 SQL:1999 和 SQL:2008 的语法，还有的使用 Oracle 11g 与微软 SQL Server 的语法。MySQL 的某些独有特征和动态视图在此也有介绍。本章介绍建立和维护数据库的 SQL 命令与单表查询编程。另外，还介绍了双表编程、IS NULL/IS NOT NULL、更多的嵌入函数、导出表以及聚合函数和 GROUP BY 子句。本章仍然以 Pine Valley 家具公司为例来解释各种查询和查询结果。

第 7 章 高级 SQL

本章继续介绍 SQL，内容包括多表查询、事务完整性、数据字典、触发器和存储过程 (它们之间的区别有清晰的解释)，以及与其他编程语言结合的嵌入式 SQL，还讨论了 OUTER JOIN 命令。本章使用标准 SQL。本章介绍了如何在导出表中存储查询结果，以及如何使用 CAST 命令转换不同数据类型。为了解释什么时候使用 EXISTS (NOT EXISTS) 和 IN (NOT IN)，还介绍了自连接。本章还包括 SQL 中最复杂和功能最强的子查询和相关子查询。

第 8 章 数据库应用开发

本章讨论最新的客户端 / 服务器体系结构与应用、中间件以及在当代数据库环境中的数据库访问等概念，介绍创建二层和三层应用的通用技术。本章给出了一些应用实例，说明如

何在流行的编程语言（如 Java、VB.NET、ASP.NET 和 JSP）中访问数据库。本章还介绍了可扩展标记语言（XML）以及相关数据存储和检索技术，内容包括 XML 模式基础、XQuery 和 XSLT。本章以 Web 服务的综述、相关标准和技术以及如何利用它们在基于 Web 的应用中无缝地、安全地移动数据结束。本章还简单介绍了面向服务的体系结构（SOA）。

第 9 章 数据仓库

本章介绍数据仓库的基本概念，解释为什么数据仓库被认为是很多组织保持竞争优势的决定性技术，以及为实现数据仓库所需要的数据库设计步骤和结构。本章还复习了决定维模型需求的最佳实践，介绍了正在兴起的列式数据库技术，这些都是专门为数据仓库应用而开发的技术。本章包括了一些支持数据管理的重要新方法，它们都是目前的热点议题，如大数据概念以及与此相关联的工具和技术。本章还提供了一些使用 SQL 和称作微策略的 BI 工具的数据集市实践性练习。这些都可以从 Teradata 大学网络上得到。附加的议题有可选择的数据仓库体系结构和适于数据仓库的维数据模型（或星模式）。有关如何处理缓变维数据也有深入讨论，并定义了操作数据存储、独立性、相关性、逻辑数据集市和各种联机分析处理（OLAP）格式。用户接口包括 OLAP、数据可视化、商业成果管理和仪表盘，以及数据挖掘等。数据质量和治理也在本章有介绍。

教学法

本教材包括了广泛的技术资源以支持各种教学方法。各章都独立成篇，这样可以按教师喜欢的方式以不同的次序使用。如前面已讨论的，某些教师可能喜欢在课程中较早地介绍 SQL 和可以动手练习的数据库作业，而不是按本书的章节顺序上课。另一些教师可能喜欢在课程最后才介绍物理数据库设计和数据库实现。

每章中都包含几个标准特征，以支持教和学按某种一致的方式使用。

1) 学习目标放在每一章的开始，预先列出了该章中学习者将要掌握的主要概念和技能，也为学习者提供了很好的有助于其准备练习和考试复习的综述。

2) 引言和总结部分给出了每一章的主要概念，以及与相关章的连接，为学生提供了有关课程的概念框架。

3) 每一章都给出了**关键术语**、**复习题**、**问题与练习**等。

- **关键术语** 为学生检查对重要概念、基本事实和有意义议题的掌握提供方便的参考。
- **复习题** 覆盖了本章的关键内容，为学生掌握本章的概念提供了一个综合工具。
- **问题与练习** 难度逐渐增加，方便教师和学生从中发现合适的题目。重点关注每一章中涉及技能的系统开发。在多个章节里，提供了本教材使用的大量数据集。
- **Web 资源** 每一章都提供了包含本章补充内容的一组网站的有效 URL。这些 Web 网站含有在线发表文档、提供商、电子文献、工业标准组织和很多其他资源。利用这些网站学生和教师可以找到本书出版后不断出现的最新产品信息、便于更深入学习的背景资料和撰写研究论文的资源。

鼓励教师灵活使用本书，以适应自己的课程体系和学生职业需求。本书模块化的特点、覆盖内容广泛、大量的解释以及对高级议题和最新话题的涵盖使得定制很方便。对当前文献和网站的很多引用，能帮助教师拓展补充的阅读列表，或扩展超出本教材内容的课堂讨论。

配套网站 (www.pearsonhighered.com/hoffer) ^①

本书提供一个可用的综合且灵活的技术支持工具箱以增强教学和学习体验。

对学生

下面是为学生提供的在线资源：

- Web 资源模块包括网站链接，可以帮助学生进一步考察数据库管理的内容。
- 按首字母缩写排序给出的完整术语表。
- 与数据集相关的网站链接。虽然我们提供的数据集的格式很容易在大学的计算机上或学生自己的 PC 上装载，但教师并不需要负责支持本地数据集。应用服务提供商（见 www.teradatauniversitynetwork.com）提供了到 SQL 编码环境的瘦客户端接口。
- 提供配套的数据库。建立了两个版本的 Pine Valley 家具公司案例。第一个版本是为了匹配本书的举例。第二个版本具有更多的数据和表以及样本表单、报告和 Visual Basic 代码块。可是，这个版本并不完整，学生可以将缺失的表、附加的表单、报告和模块补齐。数据库以多种格式（ASCII 表、Oracle 脚本以及微软 Access）提供，但是格式在这两个版本中有变化。还提供了某些数据库文档。两个版本的 PVFC 数据库也在 Teradata 大学网络中心给出。
- 多个读者开发的短视频。这些视频针对书中不同章节中的关键概念和能力要求，有助于学生对这些难理解内容的学习。这些视频集成了练习和短的讲座。

对教师

下面是为教师提供的在线资源：

- 教师手册逐章给出了教学目标、课堂思路和对复习题、问题与练习的解答。
- 测试题库和测试生成（TestGen），包括一组多选题、判断题和简短的答案，根据难易程度排列，按照书中页码（指英文书页码，与书中页边标注的页码一致）和标题引用。测试题库按 Microsoft Word 格式提供，且是计算机化的测试生成。测试生成是 PC/Mac 兼容的，且预装了全部的测试题库。你可以手工或随机地查看测试题，然后抽取得到一份测验。你也可以根据需要增加或修改测试题库的题目。
- PPT，给出了关键术语和概念。教师可以通过增加或编辑定制需要的 PPT。
- 影像库按教材的章节汇集，包括所有的图片、表格、截屏，能够用来增强课堂教学效果和完善 PPT。

致谢

每一本书的完成除了作者外都得到了一组人的帮助，深深地感谢支持和参与本书第 1 版出版工作的人员。我们特别感谢执行编辑 Bob Horan 的指导以及他热诚地与我们共同确立本书的创意，从而使该书既不同于 MDM 也不同于其他同类教材，我们相信本书的读者对象更广泛。同样重要的是来自世界各地同行的贡献，他们用过 MDM，为我们提供了高价值的反馈意见，这些意见使得我们可以持续地修订 MDM，并促使我们编写此书。我们还特别感谢

① 关于本书教辅资源，用书教师可向培生教育出版集团北京代表处申请，电话：010-5735 5169/5735 5171，电子邮件：service.cn@pearson.com。——编辑注

那些帮助梳理本书核心思路且给出了指导和反馈的同事，他们是：Indiana 大学的 Hillol Bala，Arizona 大学的 Sue Brown，Bentley 大学的 Monica Garfield，Indiana 大学的 Vijay Khatri，Massachusetts 大学 Amherst 分校的 Ryan Wright，以及 Bentley 大学的 Jennifer Xu。

我们也感谢本书的审阅者给出有深刻见解的详细建议，他们是：Gonzaga 大学的 Jason Chen，Minnesota 大学 Morris 分校的 Jinzhu Gao，American 大学的 Rick Gibson，Franklin 大学的 William Hochstettler，Michigan-Dearborn 大学的 Barbara Klein，Marist 学院的 Alan Labouseur，加州理工大学 Pomona 分校的 Lara Preiser-Huoy，Alabama 大学的 Uzma Raja。

我们也感谢 Pearson 出版社的工作人员对这本书出版过程中的支持和指导。除了执行编辑 Bob Horan 之外，还要感谢高层执行项目主管 Kelly Loftus，他总是确保每件事情完美；产品项目主管 Jane Bonnell；执行市场主管 Anne Fahlgren；编辑助理 Kaylee Rotella。我们还要特别感谢 PreMediaGlobal 的 Katy Gabel，他对产品生产过程的监控非常优秀。

最后，对我们的配偶致以最诚挚的感谢，她们不仅接受而且支持这项新任务。特别是 Patty Hoffer，她见证了 11 版 MDM 和本教材的编写过程，长达 30 多年。诚挚地感谢 Anne Louise Klaus，她支持 Heikki 参与本书写作项目。同样，非常感谢 Gayathri Mani 的持续支持和理解。这本书有她们的忍耐、鼓励和爱，书中的任何错误和不足都是我们的责任。

Jeffrey A. Hoffer

Heikki Topi

V. Ramesh

出版者的话

译者序

前言

第一部分 数据库管理的上下文

第 1 章 数据库环境和开发过程	2
1.1 基本概念和定义	3
1.1.1 数据	4
1.1.2 数据与信息	4
1.1.3 元数据	5
1.2 传统文件处理系统	6
1.3 数据库方法	7
1.3.1 数据模型	7
1.3.2 关系数据库	8
1.3.3 数据库管理系统	9
1.3.4 数据库方法的优点	9
1.3.5 数据库优点的警告	11
1.3.6 数据库方法的代价和风险	12
1.4 数据库环境的组成	12
1.5 数据库开发过程	13
1.5.1 系统开发生命周期	14
1.5.2 信息系统开发方法的选择	17
1.5.3 数据库开发的三级模式 体系结构	18
1.6 数据库系统的演化	19
1.7 数据库应用范围	21
1.7.1 个人数据库	22
1.7.2 二层客户端/服务器数据库	22
1.7.3 多层客户端/服务器数据库	22
1.7.4 企业级应用	23
1.8 Pine Valley 家具公司数据库 应用开发	24
1.8.1 Pine Valley 家具公司 数据库的演化	26
1.8.2 项目规划	26

1.8.3 数据库需求分析	27
1.8.4 数据库设计	28
1.8.5 使用数据库	31
1.8.6 管理数据库	32
1.8.7 Pine Valley 数据库的未来	32
总结	32
关键术语	33
复习题	34
问题与练习	35
参考文献	37
扩展阅读	37
Web 资源	37

第二部分 数据库分析

第 2 章 组织中的数据建模	41
2.1 E-R 模型概述	42
2.1.1 E-R 图示例	42
2.1.2 E-R 模型符号	44
2.2 组织规则建模	45
2.3 实体和属性建模	48
2.3.1 实体	48
2.3.2 属性	51
2.4 联系建模	56
2.4.1 联系的基本概念和定义	57
2.4.2 联系的度	59
2.4.3 属性或实体	62
2.4.4 基数约束	64
2.4.5 联系和基数的例子	65
2.4.6 时间相关的数据建模	66
2.4.7 实体类型间的多元联系建模	67
2.4.8 命名和定义联系	68
2.5 E-R 建模实例: Pine Valley 家具公司	70
2.6 Pine Valley 家具公司的数据库处理	72
2.6.1 显示产品信息	72

2.6.2 显示产品线信息	73	4.2.1 域约束	122
2.6.3 显示客户订单状态	73	4.2.2 实体完整性	122
总结	74	4.2.3 参照完整性	123
关键术语	75	4.2.4 创建关系表	124
复习题	76	4.2.5 完整结构化关系	125
问题与练习	77	4.3 将 EER 图转换成关系	126
参考文献	85	4.3.1 步骤 1: 映射常规实体	127
扩展阅读	86	4.3.2 步骤 2: 映射弱实体	128
Web 资源	86	4.3.3 步骤 3: 映射二元联系	130
第 3 章 增强型 E-R 模型	87	4.3.4 步骤 4: 映射关联实体	132
3.1 超类型和子类型的表示	88	4.3.5 步骤 5: 映射一元联系	134
3.1.1 基本概念和符号	88	4.3.6 步骤 6: 映射三元 (和 n 元) 联系	135
3.1.2 特化和泛化表示	91	4.3.7 步骤 7: 映射超类型 / 子类型 联系	137
3.2 在超类型 / 子类型联系中 声明约束	94	4.3.8 将 EER 转换成关系总结	138
3.2.1 声明完全性约束	94	4.4 规范化简介	138
3.2.2 声明分离性约束	95	4.4.1 规范化步骤	139
3.2.3 定义子类型鉴别子	97	4.4.2 函数依赖和键	139
3.2.4 定义超类型 / 子类型层次结构	98	4.5 规范化实例: Pine Valley 家具公司	141
3.3 EER 建模实例: Pine Valley 家具公司	100	4.5.1 步骤 0: 表示表格中的视图	142
3.4 打包数据模型	103	4.5.2 步骤 1: 转换成第一范式	142
3.4.1 利用打包数据模型改进 数据建模过程	104	4.5.3 步骤 2: 转换成第二范式	144
3.4.2 预打包数据模型总结	106	4.5.4 步骤 3: 转换成第三范式	145
总结	106	4.5.5 决定因子和规范化	146
关键术语	107	4.6 合并关系	146
复习题	107	4.6.1 例子	147
问题与练习	108	4.6.2 视图集成问题	147
参考文献	112	总结	149
扩展阅读	112	关键术语	149
Web 资源	112	复习题	150
		问题与练习	151
		参考文献	158
		扩展阅读	159
		Web 资源	159
第三部分 数据库设计		第 5 章 物理数据库设计和性能	160
第 4 章 逻辑数据库设计和关系模型	117	5.1 物理数据库设计过程	161
4.1 关系数据模型	118	5.2 设计字段	162
4.1.1 基本定义	118	5.3 去规范化数据	165
4.1.2 样本数据库	120		
4.2 完整性约束	122		

5.4 设计物理数据库文件	168
5.4.1 文件组织	169
5.4.2 对文件设计控制	174
5.5 使用和选择索引	174
5.5.1 创建唯一键索引	174
5.5.2 创建二级 (非唯一) 键索引	175
5.5.3 何时使用索引	175
5.6 为优化查询性能设计数据库	176
总结	176
关键术语	177
复习题	178
问题与练习	178
参考文献	181
扩展阅读	181
Web 资源	181

第四部分 实现

第 6 章 SQL 导论	185
6.1 SQL 标准的起源	186
6.2 SQL 环境	188
6.3 用 SQL 定义数据库	192
6.3.1 生成 SQL 数据库定义	192
6.3.2 创建表	193
6.3.3 建立数据完整性控制	195
6.3.4 修改表定义	195
6.3.5 删除表	196
6.4 插入、更新和删除数据	196
6.4.1 删除数据库内容	197
6.4.2 更新数据库内容	198
6.5 RDBMS 的内模式定义	198
6.6 单表操作	199
6.6.1 SELECT 语句中的子句	199
6.6.2 使用表达式	201
6.6.3 使用函数	202
6.6.4 使用通配符	204
6.6.5 使用比较运算符	205
6.6.6 使用 NULL 值	205
6.6.7 使用布尔运算符	206
6.6.8 使用范围限制	208

6.6.9 使用 DISTINCT	208
6.6.10 在列表中使用 IN 和 NOT IN	210
6.6.11 结果排序: ORDER BY 子句	211
6.6.12 结果分类: GROUP BY 子句	212
6.6.13 限制分类结果: HAVING 子句	213
6.6.14 使用和定义视图	214
总结	216
关键术语	217
复习题	217
问题与练习	218
参考文献	222
扩展阅读	222
Web 资源	222
第 7 章 高级 SQL	224
7.1 多表处理	224
7.1.1 等值连接	225
7.1.2 自然连接	227
7.1.3 外连接	228
7.1.4 四表连接示例	229
7.1.5 自连接	231
7.1.6 子查询	232
7.1.7 关联子查询	238
7.1.8 使用导出表	239
7.1.9 联合查询	239
7.2 创建查询的技巧	241
7.3 确保事务完整性	244
7.4 数据字典	245
7.5 触发器和例程	246
7.5.1 触发器	247
7.5.2 例程	248
7.5.3 Oracle PL/SQL 中的例程示例	249
7.6 嵌入式 SQL 和动态 SQL	251
总结	252
关键术语	253
复习题	253
问题与练习	254

参考文献	257	9.2.1 独立的数据集市数据 仓库环境	295
扩展阅读	257	9.2.2 相关数据集市和操作型数据存储 体系结构: 一个三层方法	296
Web 资源	258	9.2.3 逻辑数据集市和实时数据 仓库体系结构	298
第 8 章 数据库应用开发	259	9.2.4 三层数据体系结构	300
8.1 客户端 / 服务器体系结构	259	9.3 数据仓库数据的一些特征	301
8.2 二层体系结构中的数据库	261	9.3.1 状态数据与事件数据	301
8.2.1 VB.NET 例子	262	9.3.2 临时数据与定期数据	302
8.2.2 Java 例子	264	9.3.3 临时数据与定期数据例子	302
8.3 三层体系结构	265	9.4 派生数据层	304
8.4 Web 应用组件	266	9.4.1 派生数据的特征	304
8.5 三层应用中的数据库	268	9.4.2 星模式	305
8.5.1 JSP Web 应用	268	9.4.3 缓变维	312
8.5.2 ASP.NET 例子	272	9.4.4 决定维和事实	313
8.6 三层应用中的关键因素	273	9.5 大数据和列式数据库	315
8.6.1 存储过程	273	9.6 NoSQL	316
8.6.2 事务	273	9.7 用户接口	317
8.6.3 数据库连接	275	9.7.1 元数据的作用	317
8.6.4 三层应用的主要优点	275	9.7.2 SQL OLAP 查询	318
8.6.5 云计算和三层应用	276	9.7.3 联机分析处理工具	318
8.7 可扩展标记语言	276	9.7.4 数据可视化	321
8.7.1 存储 XML 文档	278	9.7.5 商业成果管理和仪表盘	321
8.7.2 检索 XML 文档	279	9.7.6 数据挖掘工具	322
8.7.3 显示 XML 数据	281	9.8 数据治理与数据质量	323
8.7.4 XML 和 Web 服务	282	9.8.1 数据治理	324
总结	284	9.8.2 管理数据质量	324
关键术语	285	9.8.3 数据质量的特征	325
复习题	286	总结	326
问题与练习	286	关键术语	327
参考文献	287	复习题	328
扩展阅读	287	问题与练习	329
Web 资源	287	参考文献	333
第 9 章 数据仓库	289	扩展阅读	334
9.1 数据仓库的基本概念	291	Web 资源	334
9.1.1 数据仓库简史	291	缩略词	335
9.1.2 数据仓库的需求	292	术语表	337
9.2 数据仓库体系结构	294	索引	343

数据库管理的上下文

第一部分包括第 1 章。

第 1 章给出了贯穿本书的基本数据库概念和定义。数据库管理作为一个活跃的、具有挑战性的和正在发展的领域，为信息学院学生创造了巨大的就业机会。数据库已成为我们日常生活的必需品，并且是商业运作的更重要的核心。数据库可以是仅仅在个人数字助理（PDA）和智能手机中存储的通信信息，也可以是支持企业级信息系统的非常大的数据库。数十年前，数据库就已成为数据存储的中心点。近年来发展起来的客户关系管理和网上购物是两个数据库相关活动的例子。数据仓库的发展为管理者更深入和广泛地分析历史数据提供了机会，而且它还将持续得到关注。

本部分首先定义数据、数据库、元数据、数据库管理系统（DBMS）、数据仓库和其他有关的术语。通过仔细地规划数据库的使用，我们将数据库与早期的文件管理系统进行对比，并描述其几个重要的优点。

第 1 章还描述了数据库分析、设计、实现和管理所遵循的一般步骤，用示意图说明了数据库开发过程如何融进整个信息系统开发过程中。然后，讨论了数据库开发中的结构化生命周期和原型方法。我们引入企业数据建模，该模型设定了组织数据库的范围和一般内容。这一步常常是数据库开发中的第一步。本章介绍模式以及三级模式体系结构的概念，这是现代数据库系统中的主要方法。本章介绍数据库环境的主要成分和应用类型以及两层、多层和企业数据库。企业数据库包含用于支持企业资源规划系统和数据仓库的内容。最后，讨论在数据库开发项目中涉及的主要人员的职责。引入 Pine Valley 家具公司案例来解释数据库管理的原则和概念，这个案例将贯穿全书。