

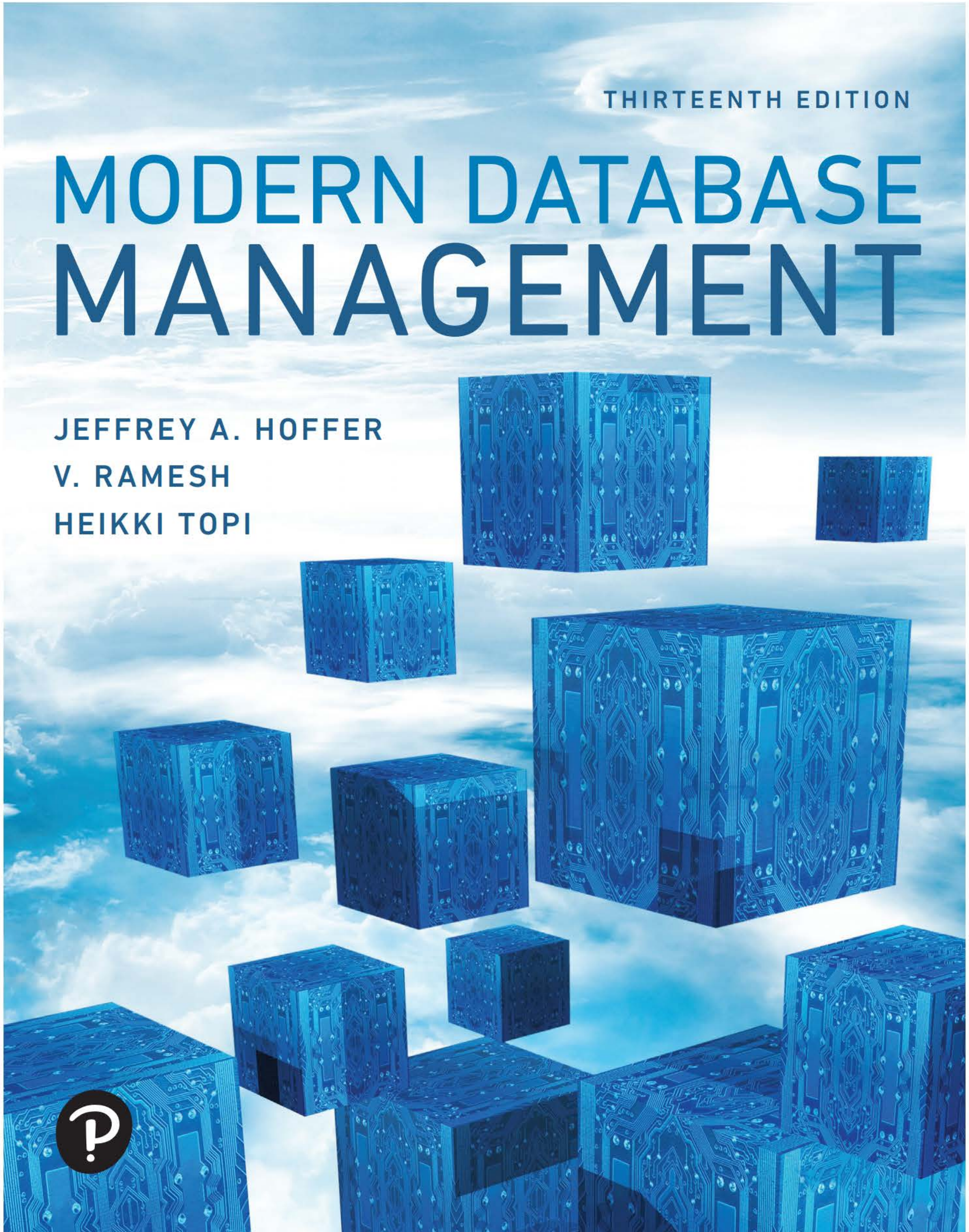
Pearson

RENTAL EDITION - RENTAL ONLY, NOT FOR SALE

THIRTEENTH EDITION

MODERN DATABASE MANAGEMENT

JEFFREY A. HOFFER
V. RAMESH
HEIKKI TOPI



Complete Listing of Chapter Opening Cases, Insight Cases, E-commerce in Action Cases, and Case Studies

CHAPTER 1 THE REVOLUTION IS JUST BEGINNING

Opening Case: Everything on Demand: The “Uberization” of E-commerce

Insight on Technology: Will Apps Make the Web Irrelevant?

Insight on Business: Startup Boot Camp

Insight on Society: Facebook and the Age of Privacy

Case Study: Pinterest: A Picture Is Worth a Thousand Words

CHAPTER 2 E-COMMERCE BUSINESS MODELS AND CONCEPTS

Opening Case: Tweet Tweet: Will Twitter Ever Find a Business Model that Works?

Insight on Society: Foursquare: Check Your Privacy at the Door

Insight on Business: Crowdfunding Takes Off

Insight on Technology: Will the Connected Car Become the Next Hot Entertainment Vehicle?

Case Study: Dollar Shave Club: From Viral Video to \$1 Billion in Just Five Years

CHAPTER 3 E-COMMERCE INFRASTRUCTURE: THE INTERNET, WEB, AND MOBILE PLATFORM

Opening Case: Voice-Controlled Intelligent Digital Assistants: Will They Revolutionize E-commerce?

Insight on Society: Government Regulation and Surveillance of the Internet

Insight on Technology: The Rise of HTML5

Insight on Business: The Apple Watch: Bringing The Internet of Things to Your Wrist

Case Study: Akamai Technologies: Attempting to Keep Supply Ahead of Demand

CHAPTER 4 BUILDING AN E-COMMERCE PRESENCE: WEBSITES, MOBILE SITES, AND APPS

Opening Case: The Wall Street Journal: Redesigning for Today’s Platforms

Insight on Business: Weebly Makes Creating Websites Easy

Insight on Society: Designing for Accessibility

Insight on Technology: Carnival Cruise Ships Go Mobile

Case Study: Dick’s Sporting Goods: Taking Control of Its E-commerce Operations

CHAPTER 5 E-COMMERCE SECURITY AND PAYMENT SYSTEMS

Opening Case: Cyberwar: MAD 2.0

Insight on Society: Equifax: Really Big Data Hacked

Insight on Technology: Think Your Smartphone Is Secure?

Insight on Business: Bitcoin

Case Study: The Mobile Payment Marketplace: Goat Rodeo

CHAPTER 6 E-COMMERCE MARKETING AND ADVERTISING CONCEPTS

Opening Case: Video Ads: Shoot, Click, Buy

Insight on Business: Are the Very Rich Different From You and Me?

Insight on Technology: The Long Tail: Big Hits and Big Misses

Insight on Society: Every Move You Take, Every Click You Make, We’ll Be Tracking You

Case Study: Programmatic Advertising: Real-Time Marketing

CHAPTER 7 SOCIAL, MOBILE, AND LOCAL MARKETING

Opening Case: Facebook: Putting Social Marketing to Work

Insight on Technology: Optimizing Social Marketing with Simply Measured

Insight on Society: Marketing to Children of the Web in the Age of Social Networks

Insight on Business: Mobile Marketing Goes 3-D

Case Study: ExchangeHunterJumper.com: Building a Brand with Social Marketing

CHAPTER 8 ETHICAL, SOCIAL, AND POLITICAL ISSUES IN E-COMMERCE

Opening Case: The Right To Be Forgotten: Europe Leads on Internet Privacy

Insight on Technology: Apple: Defender of Privacy?

Insight on Business: Internet Sales Tax Battle

Insight on Society: The Internet Drug Bazaar

Case Study: The Pirate Bay: Searching for a Safe Haven

CHAPTER 9 ONLINE RETAILING AND SERVICES

Opening Case: Blue Nile Sparkles for Your Cleopatra

E-commerce in Action: Amazon

Insight on Technology: Big Data and Predictive Marketing

Insight on Society: Phony Reviews

Insight on Business: Food on Demand: Instacart and GrubHub

Case Study: OpenTable: Your Reservation Is Waiting

CHAPTER 10 ONLINE CONTENT AND MEDIA

Opening Case: Cord Cutters and Cord Shavers: The Emerging Internet Broadcasting System (IBS)

Insight on Society: Are Millennials Really All That Different?

Insight on Business: Vox: Native Digital News

Insight on Technology: Hollywood and the Internet: Let's Cut a Deal

Case Study: Netflix: How Does This Movie End?

CHAPTER 11 SOCIAL NETWORKS, AUCTIONS, AND PORTALS

Opening Case: Social Network Fever Spreads to the Professions

Insight on Society: The Dark Side of Social Networks

Insight on Technology: Trapped Inside the Facebook Bubble?

Insight on Business: Verizon Doubles Down on Portals

Case Study: eBay Evolves

CHAPTER 12 B2B E-COMMERCE: SUPPLY CHAIN MANAGEMENT AND COLLABORATIVE COMMERCE

Opening Case: Amazon Takes on B2B with Amazon Business

Insight on Society: Where's My iPad? Supply Chain Risk and Vulnerability

Insight on Technology: Your Shoes Are in the Cloud

Insight on Business: Walmart Develops a Private Industrial Network

Case Study: Elemica: Cooperation, Collaboration, and Community

OTHER MIS TITLES OF INTEREST

Introductory MIS

Experiencing MIS, 8/e
Kroenke & Boyle ©2019

Using MIS, 10/e
Kroenke & Boyle ©2018

Management Information Systems, 15/e
Laudon & Laudon ©2018

Essentials of MIS, 13/e
Laudon & Laudon ©2019

Processes, Systems, and Information: An Introduction to MIS, 3/e
McKinney & Kroenke ©2019

Information Systems Today, 8/e
Valacich & Schneider ©2018

Introduction to Information Systems, 3/e
Wallace ©2018

Database

Hands-on Database, 2/e
Conger ©2014

Modern Database Management, 13/e
Hoffer, Ramesh & Topi ©2019

Database Concepts, 8/e
Kroenke, Auer, Vandenburg, Yoder ©2018

Database Processing, 15/e
Kroenke & Auer ©2019

Systems Analysis and Design

Modern Systems Analysis and Design, 8/e
Hoffer, George & Valacich ©2017

Systems Analysis and Design, 10/e
Kendall & Kendall ©2019

Decision Support Systems

Business Intelligence, Analytics, and Data Science, 4/e
Sharda, Delen & Turban ©2018

Business Intelligence and Analytics: Systems for Decision Support, 10/e
Sharda, Delen & Turban ©2014

Data Communications & Networking

Applied Networking Labs, 2/e
Boyle ©2014

Digital Business Networks
Dooley ©2014

Business Data Networks and Security, 11/e
Panko & Panko ©2019

Electronic Commerce

E-commerce 2018: Business. Technology. Society, 14/e
Laudon & Traver ©2019

Enterprise Resource Planning

Enterprise Systems for Management, 2/e
Motiwalla & Thompson ©2012

Project Management

Project Management: Process, Technology and Practice
Vaidyanathan ©2013

THIRTEENTH EDITION

MODERN DATABASE MANAGEMENT

THIRTEENTH EDITION

MODERN DATABASE MANAGEMENT

Jeffrey A. Hoffer
University of Dayton

V. Ramesh
Indiana University

Heikki Topi
Bentley University



330 Hudson Street, NY NY 10013

Vice President, IT & Careers: Andrew Gilfillan
Senior Portfolio Manager: Samantha Lewis
Managing Producer: Laura Burgess
Associate Content Producer: Stephany Harrington
Portfolio Management Assistant: Madeline Houpt
Director of Product Marketing: Brad Parkins
Product Marketing Manager: Heather Taylor
Product Marketing Assistant: Jesika Bethea
Field Marketing Manager: Molly Schmidt
Field Marketing Assistant: Kelli Fisher
Cover Image: VICTOR HABBICK VISIONS/Getty Images

Vice President, Product Model Management: Jason Fournier
Senior Product Model Manager: Eric Hakanson
Lead, Production and Digital Studio: Heather Darby
Digital Studio Course Producer: Jaimie Noy
Program Monitor: Danica Monzor, SPi Global
Full-Service Project Management:
Neha Bhargava, Cenveo® Publisher Services
Composition: Cenveo Publisher Services
Printer/Binder: LSC Communications
Cover Printer: Phoenix Color
Text Font: Palatino LT Pro

Credits and acknowledgments borrowed from other sources and reproduced, with permission, in this textbook appear on the appropriate page within text.

Microsoft and/or its respective suppliers make no representations about the suitability of the information contained in the documents and related graphics published as part of the services for any purpose. All such documents and related graphics are provided “as is” without warranty of any kind. Microsoft and/or its respective suppliers hereby disclaim all warranties and conditions with regard to this information, including all warranties and conditions of merchantability, whether express, implied or statutory, fitness for a particular purpose, title and non-infringement. In no event shall Microsoft and/or its respective suppliers be liable for any special, indirect or consequential damages or any damages whatsoever resulting from loss of use, data or profits, whether in an action of contract, negligence or other tortious action, arising out of or in connection with the use or performance of information available from the services.

The documents and related graphics contained herein could include technical inaccuracies or typographical errors. Changes are periodically added to the information herein. Microsoft and/or its respective suppliers may make improvements and/or changes in the product(s) and/or the program(s) described herein at any time. Partial screen shots may be viewed in full within the software version specified.

Trademarks

Microsoft® Windows®, and Microsoft Office® are registered trademarks of the Microsoft Corporation in the U.S.A. and other countries. This book is not sponsored or endorsed by or affiliated with the Microsoft Corporation.

Copyright © 2019, 2016, 2013 by Pearson Education, Inc. All rights reserved. Manufactured in the United States of America. This publication is protected by Copyright, and permission should be obtained from the publisher prior to any prohibited reproduction, storage in a retrieval system, or transmission in any form or by any means, electronic, mechanical, photocopying, recording, or likewise. For information regarding permissions, request forms and the appropriate contacts within the Pearson Education Global Rights & Permissions department, please visit www.pearsoned.com/permissions.

Acknowledgements of third party content appear on the appropriate page within the text, which constitutes an extension of this copyright page.

Unless otherwise indicated herein, any third-party trademarks that may appear in this work are the property of their respective owners and any references to third-party trademarks, logos or other trade dress are for demonstrative or descriptive purposes only. Such references are not intended to imply any sponsorship, endorsement, authorization, or promotion of Pearson’s products by the owners of such marks, or any relationship between the owner and Pearson Education, Inc. or its affiliates, authors, licensees or distributors.

Library of Congress Cataloging-in-Publication Data

Names: Hoffer, Jeffrey A., author. | Ramesh, V. (Venkataraman), author. | Topi, Heikki, author.

Title: Modern database management / Jeffrey A. Hoffer, University of Dayton, V. Ramesh, Indiana University, Heikki Topi, Bentley University.

Description: Thirteenth edition. | Boston : Pearson Education, [2017] | Includes bibliographical references and index.

Identifiers: LCCN 2017048422 | ISBN 9780134773650 (alk. paper) | ISBN 0133544613 (alk. paper)

Subjects: LCSH: Database management.

Classification: LCC QA76.9.D3 M395 2017 | DDC 005.75—dc23 LC record available at <https://lcn.loc.gov/2017048422>

10 9 8 7 6 5 4 3 2 1



ISBN 10: 0-13-477365-9
ISBN 13: 978-0-13-477365-0

To Patty, for her sacrifices, encouragement, and support for more than 35 years of being a textbook author widow. To my students and colleagues, for being receptive and critical and for challenging me to be a better teacher.

—J.A.H.

To Gayathri, for her sacrifices and patience these past 25 years. To my parents, for letting me make the journey abroad, and to my cat, Raju, who was a part of our family for more than 20 years.

—V.R.

To Anne-Louise, for her loving support, encouragement, and patience. To Leila and Saara, whose laughter and joy of life continue to teach me about what is truly important. To my teachers, colleagues, and students, from whom I continue to learn every day.

—H.T.

BRIEF CONTENTS

Part I The Context of Database Management 1

Chapter 1 The Database Environment and Development Process 3

Part II Database Analysis and Logical Design 53

Chapter 2 Modeling Data in the Organization 55

Chapter 3 The Enhanced E-R Model 115

Chapter 4 Logical Database Design and the Relational Model 153

Part III Database Implementation and Use 205

Chapter 5 Introduction to SQL 207

Chapter 6 Advanced SQL 251

Chapter 7 Databases in Applications 297

Chapter 8 Physical Database Design and Database Infrastructure 333

Part IV Advanced Database Topics 385

Chapter 9 Data Warehousing and Data Integration 387

Chapter 10 Big Data Technologies 444

Chapter 11 Analytics and Its Implications 474

Chapter 12 Data and Database Administration with Focus on Data Quality 503

Glossary of Acronyms 529

Glossary of Terms 531

Index 539

Available Online at www.pearsonhighered.com/hoffer

Chapter 13 Distributed Databases 13-1

Chapter 14 Object-Oriented Data Modeling 14-1

Appendices

Appendix A Data Modeling Tools and Notation A-1

Appendix B Advanced Normal Forms B-1

Appendix C Data Structures C-1

CONTENTS

Preface xxv

Part I The Context of Database Management 1

An Overview of Part I 1

Chapter 1 The Database Environment and Development Process 3



Learning Objectives 3

Data Matter! 4

Introduction 5

Basic Concepts and Definitions 6

Data 6

Data versus Information 7

Metadata 8

Traditional File Processing Systems 9

File Processing Systems at Pine Valley Furniture Company 9

Disadvantages of File Processing Systems 10

PROGRAM-DATA DEPENDENCE 10

DUPLICATION OF DATA 10

LIMITED DATA SHARING 10

LENGTHY DEVELOPMENT TIMES 10

EXCESSIVE PROGRAM MAINTENANCE 11

The Database Approach 11

Data Models 11

ENTITIES 11

RELATIONSHIPS 11

Relational Databases 12

Database Management Systems 13

Advantages of the Database Approach 13

PROGRAM-DATA INDEPENDENCE 13

PLANNED DATA REDUNDANCY 14

IMPROVED DATA CONSISTENCY 14

IMPROVED DATA SHARING 14

INCREASED PRODUCTIVITY OF APPLICATION DEVELOPMENT 14

ENFORCEMENT OF STANDARDS 15

IMPROVED DATA QUALITY 15

IMPROVED DATA ACCESSIBILITY AND RESPONSIVENESS 15

REDUCED PROGRAM MAINTENANCE 16

IMPROVED DECISION SUPPORT 16

CAUTIONS ABOUT DATABASE BENEFITS 16

COSTS AND RISKS OF THE DATABASE APPROACH 16

NEW, SPECIALIZED PERSONNEL 16

INSTALLATION AND MANAGEMENT COST AND COMPLEXITY 17

CONVERSION COSTS 17

NEED FOR EXPLICIT BACKUP AND RECOVERY 17

ORGANIZATIONAL CONFLICT 17

Integrated Data Management Framework 17

Components of the Database Environment 18

The Database Development Process 20

- Systems Development Life Cycle 21
 - PLANNING—ENTERPRISE MODELING 21
 - PLANNING—CONCEPTUAL DATA MODELING 21
 - ANALYSIS—CONCEPTUAL DATA MODELING 22
 - DESIGN—LOGICAL DATABASE DESIGN 23
 - DESIGN—PHYSICAL DATABASE DESIGN AND DEFINITION 23
 - IMPLEMENTATION—DATABASE IMPLEMENTATION 23
 - MAINTENANCE—DATABASE MAINTENANCE 24
- Alternative Information Systems Development Approaches 24
- Three-Schema Architecture for Database Development 25
- Managing the People Involved in Database Development 27

Evolution of Database Systems 27

- 1960s 29
- 1970s 29
- 1980s 29
- 1990s 30
- 2000 and Beyond 30

The Range of Database Applications 30

- Personal Databases 31
- Departmental Multi-Tiered Client/Server Databases 31
- Enterprise Applications 32
 - ENTERPRISE SYSTEMS 32
 - DATA WAREHOUSES 33
 - DATA LAKE 34

Developing a Database Application for Pine Valley Furniture Company 35

- Database Evolution at Pine Valley Furniture Company 36
- Project Planning 36
- Analyzing Database Requirements 37
- Designing the Database 40
- Using the Database 42
- Administering the Database 43
- Future of Databases at Pine Valley 43
 - Summary* 44 • *Key Terms* 45 • *Review Questions* 45 •
 - Problems and Exercises* 46 • *Field Exercises* 48 •
 - References* 49 • *Further Reading* 49 •
 - Web Resources* 50
- **CASE: Forondo Artist Management Excellence Inc.** 51



Part II Database Analysis and Logical Design 53

An Overview of Part II 53



Chapter 2 Modeling Data in the Organization 55

Learning Objectives 55

Introduction 55

The E-R Model: An Overview 58

- Sample E-R Diagram 58

- E-R Model Notation 60

Modeling the Rules of the Organization 61

Overview of Business Rules 62
 THE BUSINESS RULES PARADIGM 62

Scope of Business Rules 63
 GOOD BUSINESS RULES 63
 GATHERING BUSINESS RULES 64

Data Names and Definitions 64
 DATA NAMES 64
 DATA DEFINITIONS 65
 GOOD DATA DEFINITIONS 65

Modeling Entities and Attributes 67
 Entities 67
 ENTITY TYPE VERSUS ENTITY INSTANCE 67
 ENTITY TYPE VERSUS SYSTEM INPUT, OUTPUT, OR USER 67
 STRONG VERSUS WEAK ENTITY TYPES 68
 NAMING AND DEFINING ENTITY TYPES 69

Attributes 71
 REQUIRED VERSUS OPTIONAL ATTRIBUTES 71
 SIMPLE VERSUS COMPOSITE ATTRIBUTES 72
 SINGLE-VALUED VERSUS MULTIVALUED ATTRIBUTES 72
 STORED VERSUS DERIVED ATTRIBUTES 73
 IDENTIFIER ATTRIBUTE 73
 NAMING AND DEFINING ATTRIBUTES 74

Modeling Relationships 76
 Basic Concepts and Definitions in Relationships 77
 ATTRIBUTES ON RELATIONSHIPS 78
 ASSOCIATIVE ENTITIES 78

Degree of a Relationship 80
 UNARY RELATIONSHIP 81
 BINARY RELATIONSHIP 82
 TERNARY RELATIONSHIP 82

Attributes or Entity? 83

Cardinality Constraints 85
 MINIMUM CARDINALITY 85
 MAXIMUM CARDINALITY 86

Some Examples of Relationships and Their Cardinalities 86
 A TERNARY RELATIONSHIP 87

Modeling Time-Dependent Data 88

Modeling Multiple Relationships Between Entity Types 90

Naming and Defining Relationships 92

E-R Modeling Example: Pine Valley Furniture Company 93

Database Processing At Pine Valley Furniture 96
 Showing Product Information 96
 Showing Product Line Information 96
 Showing Customer Order Status 97
 Showing Product Sales 98
 Summary 99 • Key Terms 100 • Review Questions 100 •
 Problems and Exercises 101 • Field Exercises 111 •
 References 112 • Further Reading 112 •
 Web Resources 112

► CASE: Forondo Artist Management Excellence Inc. 113





Chapter 3 The Enhanced E-R Model 115

- Learning Objectives 115
- Introduction 115
- Representing Supertypes and Subtypes 116
 - Basic Concepts and Notation 117
 - AN EXAMPLE OF A SUPERTYPE/SUBTYPE RELATIONSHIP 118
 - ATTRIBUTE INHERITANCE 119
 - WHEN TO USE SUPERTYPE/SUBTYPE RELATIONSHIPS 119
 - Representing Specialization and Generalization 120
 - GENERALIZATION 120
 - SPECIALIZATION 121
 - COMBINING SPECIALIZATION AND GENERALIZATION 122
- Specifying Constraints in Supertype/Subtype Relationships 123
 - Specifying Completeness Constraints 123
 - TOTAL SPECIALIZATION RULE 123
 - PARTIAL SPECIALIZATION RULE 123
 - Specifying Disjointness Constraints 124
 - DISJOINT RULE 124
 - OVERLAP RULE 125
 - Defining Subtype Discriminators 125
 - DISJOINT SUBTYPES 125
 - OVERLAPPING SUBTYPES 126
 - Defining Supertype/Subtype Hierarchies 127
 - AN EXAMPLE OF A SUPERTYPE/SUBTYPE HIERARCHY 128
 - SUMMARY OF SUPERTYPE/SUBTYPE HIERARCHIES 128
- EER Modeling Example: Pine Valley Furniture Company 128
- Entity Clustering 132
- Packaged Data Models 135
 - A Revised Data Modeling Process with Packaged Data Models 137
 - Packaged Data Model Examples 139
 - Summary 144 • Key Terms 145 • Review Questions 145 • Problems and Exercises 146 • Field Exercises 149 • References 149 • Further Reading 150 • Web Resources 150*
 - ▶ **CASE: Forondo Artist Management Excellence Inc. 151**



Chapter 4 Logical Database Design and the Relational Model 153

- Learning Objectives 153
- Introduction 153
- The Relational Data Model 154
 - Basic Definitions 154
 - RELATIONAL DATA STRUCTURE 155
 - RELATIONAL KEYS 155
 - PROPERTIES OF RELATIONS 156
 - REMOVING MULTIVALUED ATTRIBUTES FROM TABLES 156
 - Sample Database 157
- Integrity Constraints 158
 - Domain Constraints 158
 - Entity Integrity 158
 - Referential Integrity 160

Creating Relational Tables 161
 Well-Structured Relations 162
 Transforming EER Diagrams into Relations 163
 Step 1: Map Regular Entities 164
 COMPOSITE ATTRIBUTES 164
 MULTIVALUED ATTRIBUTES 165
 Step 2: Map Weak Entities 165
 WHEN TO CREATE A SURROGATE KEY 166
 Step 3: Map Binary Relationships 167
 MAP BINARY ONE-TO-MANY RELATIONSHIPS 167
 MAP BINARY MANY-TO-MANY RELATIONSHIPS 168
 MAP BINARY ONE-TO-ONE RELATIONSHIPS 168
 Step 4: Map Associative Entities 169
 IDENTIFIER NOT ASSIGNED 169
 IDENTIFIER ASSIGNED 170
 Step 5: Map Unary Relationships 171
 UNARY ONE-TO-MANY RELATIONSHIPS 171
 UNARY MANY-TO-MANY RELATIONSHIPS 172
 Step 6: Map Ternary (and n -ary) Relationships 173
 Step 7: Map Supertype/Subtype Relationships 174
 Summary of EER-to-Relational Transformations 176
 Introduction to Normalization 176
 Steps in Normalization 177
 Functional Dependencies and Keys 177
 DETERMINANTS 179
 CANDIDATE KEYS 179
 Normalization Example: Pine Valley Furniture Company 180
 Step 0: Represent the View in Tabular Form 180
 Step 1: Convert to First Normal Form 181
 REMOVE REPEATING GROUPS 181
 SELECT THE PRIMARY KEY 182
 ANOMALIES IN 1NF 182
 Step 2: Convert to Second Normal Form 183
 Step 3: Convert to Third Normal Form 184
 REMOVING TRANSITIVE DEPENDENCIES 184
 Determinants and Normalization 185
 Step 4: Further Normalization 185
 Merging Relations 186
 An Example 186
 View Integration Problems 186
 SYNONYMS 187
 HOMONYMS 187
 TRANSITIVE DEPENDENCIES 187
 SUPERTYPE/SUBTYPE RELATIONSHIPS 188
 A Final Step for Defining Relational Keys 188
 Summary 191 • *Key Terms* 191 • *Review Questions* 191 •
 Problems and Exercises 192 • *Field Exercises* 201 •
 References 202 • *Further Reading* 202 •
 Web Resources 202
 ▶ **CASE: Forondo Artist Management Excellence Inc.** 203



Part III Database Implementation and Use 205

An Overview of Part III 205



Chapter 5 Introduction to SQL 207

Learning Objectives 207

Introduction 207

Origins of the SQL Standard 209

The SQL Environment 211

SQL Data Types 213

Defining A Database in SQL 216

Generating SQL Database Definitions 216

Creating Tables 217

Creating Data Integrity Controls 220

Changing Table Definitions 221

Removing Tables 221

Inserting, Updating, and Deleting Data 222

Batch Input 223

Deleting Database Contents 223

Updating Database Contents 224

Internal Schema Definition in RDBMSs 225

Creating Indexes 225

Processing Single Tables 226

Clauses of the SELECT Statement 226

Using Expressions 228

Using Functions 229

Using Wildcards 232

Using Comparison Operators 232

Using Null Values 233

Using Boolean Operators 233

Using Ranges for Qualification 236

Using Distinct Values 236

Using IN and NOT IN with Lists 238

Sorting Results: The ORDER BY Clause 239

Categorizing Results: The GROUP BY Clause 240

Qualifying Results by Categories: The HAVING Clause 241

Summary 243 • Key Terms 243 • Review Questions 243 •

Problems and Exercises 244 • Field Exercises 248 •

References 248 • Further Reading 249 •

Web Resources 249

▶ **CASE: Forondo Artist Management Excellence Inc. 250**



Chapter 6 Advanced SQL 251

Learning Objectives 251

Introduction 251

Processing Multiple Tables 252

Equi-Join 253

Natural Join 254

Outer Join 255

Sample Join Involving Four Tables 257

Self-Join	258
Subqueries	260
Correlated Subqueries	265
Using Derived Tables	267
Combinings Queries	267
Conditional Expressions	269
More Complicated SQL Queries	270
Tips for Developing Queries	272
Guidelines for Better Query Design	274
Using and Defining Views	275
Materialized Views	279
Triggers and Routines	279
Triggers	280
Routines and Other Programming Extensions	282
Example Routine in Oracle's PL/SQL	284
Data Dictionary Facilities	285
Recent Enhancements and Extensions to SQL	287
Analytical and OLAP Functions	287
New Temporal Features in SQL	288
Other Enhancements	288
<i>Summary</i>	289
<i>Key Terms</i>	290
<i>Review Questions</i>	290
<i>Problems and Exercises</i>	291
<i>Field Exercises</i>	294
<i>References</i>	294
<i>Further Reading</i>	295
<i>Web Resources</i>	295
▶ CASE: Forondo Artist Management Excellence Inc.	296
Chapter 7 Databases in Applications	297
Learning Objectives	297
Location, Location, Location!	297
Introduction	298
Client/Server Architectures	298
Databases in Three-Tier Applications	302
A Java Web Application	303
A Python Web Application	307
Key Considerations in Three-Tier Applications	313
Stored Procedures	313
Transactions	313
Database Connections	315
Key Benefits of Three-Tier Applications	315
Transaction Integrity	316
Controlling Concurrent Access	318
The Problem of Lost Updates	318
Serializability	319
Locking Mechanisms	319
LOCKING LEVEL	319
TYPES OF LOCKS	320
DEADLOCK	321
MANAGING DEADLOCK	321
Versioning	322





Managing Data Security in an Application Context 324

- Threats to Data Security 324
- Establishing Client/Server Security 325
 - SERVER SECURITY 326
 - NETWORK SECURITY 326
- Application Security Issues in Three-Tier Client/Server Environments 326
 - DATA PRIVACY 327
 - Summary 329 • Key Terms 329 • Review Questions 329 • Problems and Exercises 330 • Field Exercises 331 • References 331 • Further Reading 331 • Web Resources 331*
 - ▶ **CASE: Forondo Artist Management Excellence Inc. 332**

Chapter 8 Physical Database Design and Database Infrastructure 333

- Learning Objectives 333
- Introduction 334
- The Physical Database Design Process 335
 - Who Is Responsible for Physical Database Design? 335
 - Physical Database Design as a Basis for Regulatory Compliance 336
 - SOX and Databases 337
 - IT CHANGE MANAGEMENT 337
 - LOGICAL ACCESS TO DATA 337
 - IT OPERATIONS 338
 - Data Volume and Usage Analysis 338
- Designing Fields 340
 - Choosing Data Types 340
 - CODING TECHNIQUES 341
 - CONTROLLING DATA INTEGRITY 342
 - HANDLING MISSING DATA 343
- Denormalizing and Partitioning Data 343
 - Denormalization 343
 - OPPORTUNITIES FOR AND TYPES OF DENORMALIZATION 344
 - DENORMALIZE WITH CAUTION 345
 - Partitioning 347
- Designing Physical Database Files 348
 - File Organizations 350
 - HEAP FILE ORGANIZATION 350
 - SEQUENTIAL FILE ORGANIZATIONS 350
 - INDEXED FILE ORGANIZATIONS 352
 - HASHED FILE ORGANIZATIONS 353
 - Clustering Files 353
 - Designing Controls for Files 354
- Using and Selecting Indexes 354
 - Creating a Unique Key Index 354
 - Creating a Secondary (Nonunique) Key Index 355
 - When to Use Indexes 355
- Designing a Database for Optimal Query Performance 356
 - Parallel Query Processing 357
 - Overriding Automatic Query Optimization 358
- Data Dictionaries and Repositories 358

Data Dictionary	359
Repositories	359
Database Software Data Security Features	361
Views	361
Integrity Controls	362
Authorization Rules	363
User-Defined Procedures	365
Encryption	365
Authentication Schemes	365
PASSWORDS	366
STRONG AUTHENTICATION	366
Database Backup and Recovery	367
Basic Recovery Facilities	367
BACKUP FACILITIES	367
JOURNALIZING FACILITIES	368
CHECKPOINT FACILITY	368
RECOVERY MANAGER	369
Recovery and Restart Procedures	369
DISK MIRRORING	369
RESTORE/RERUN	370
BACKWARD RECOVERY	370
FORWARD RECOVERY	371
Types of Database Failure	371
ABORTED TRANSACTIONS	372
INCORRECT DATA	372
SYSTEM FAILURE	372
DATABASE DESTRUCTION	372
Disaster Recovery	373
Cloud-Based Database Infrastructure	373
Cloud-Based Models for Providing Data Management Services	373
Benefits and Downsides of Using Cloud-Based Data Management Services	374
<i>Summary</i>	375 •
<i>Key Terms</i>	376 •
<i>Review Questions</i>	377 •
<i>Problems and Exercises</i>	378 •
<i>Field Exercises</i>	382 •
<i>References</i>	383 •
<i>Further Reading</i>	383 •
<i>Web Resources</i>	383
▶ CASE: Forondo Artist Management Excellence Inc.	384



Part IV Advanced Database Topics 385

An Overview of Part IV 385

Chapter 9 Data Warehousing and Data Integration 387

Learning Objectives 387

Introduction 387

Basic Concepts of Data Warehousing 390

 A Brief History of Data Warehousing 390

 The Need for Data Warehousing 390

 NEED FOR A COMPANY-WIDE VIEW 390

 NEED TO SEPARATE OPERATIONAL AND INFORMATIONAL SYSTEMS 393

Data Warehouse Architectures 393

- Independent Data Mart Data Warehousing Environment 394
- Dependent Data Mart and Operational Data Store Architecture: A Three-Level Approach 395
- Logical Data Mart and Real-Time Data Warehouse Architecture 397
- Three-Layer Data Architecture 400
 - ROLE OF THE ENTERPRISE DATA MODEL 400
 - ROLE OF METADATA 400
- Some Characteristics of Data Warehouse Data 401
 - Status versus Event Data 401
 - Transient versus Periodic Data 402
 - An Example of Transient and Periodic Data 402
 - TRANSIENT DATA 404
 - PERIODIC DATA 404
 - OTHER DATA WAREHOUSE CHANGES 404
- The Derived Data Layer 405
 - Characteristics of Derived Data 405
 - The Star Schema 406
 - FACT TABLES AND DIMENSION TABLES 406
 - EXAMPLE STAR SCHEMA 407
 - SURROGATE KEY 408
 - GRAIN OF THE FACT TABLE 409
 - DURATION OF THE DATABASE 410
 - SIZE OF THE FACT TABLE 410
 - MODELING DATE AND TIME 411
 - Variations of the Star Schema 412
 - MULTIPLE FACT TABLES 412
 - FACTLESS FACT TABLES 413
 - Normalizing Dimension Tables 414
 - MULTIVALUED DIMENSIONS 414
 - HIERARCHIES 415
 - Slowly Changing Dimensions 417
 - Determining Dimensions and Facts 420
- Data Integration: An Overview 422
 - General Approaches to Data Integration 422
 - DATA FEDERATION 423
 - DATA PROPAGATION 423
- Data Integration for Data Warehousing: The Reconciled Data Layer 424
 - Characteristics of Data after ETL 424
 - The ETL Process 425
 - MAPPING AND METADATA MANAGEMENT 425
 - EXTRACT 426
 - CLEANSE 427
 - LOAD AND INDEX 429
- Data Transformation 430
 - Data Transformation Functions 431
 - RECORD-LEVEL FUNCTIONS 431
 - FIELD-LEVEL FUNCTIONS 432
- Data Warehouse Administration 434

The Future of Data Warehousing: Integration with Other Forms of Data Management and Analytics	434
Speed of Processing	435
Moving the Data Warehouse into the Cloud	435
Dealing with Unstructured Data	436
<i>Summary</i>	436 • <i>Key Terms</i> 437 • <i>Review Questions</i> 437 • <i>Problems and Exercises</i> 438 • <i>Field Exercises</i> 442 • <i>References</i> 442 • <i>Further Reading</i> 443 • <i>Web Resources</i> 443

Chapter 10 Big Data Technologies 444

Learning Objectives	444
Introduction	444
Moving Beyond Transactional and Data Warehousing Databases	446
Big Data	446
NoSQL	448
Classification of NoSQL DBMSs	450
KEY-VALUE STORES	450
DOCUMENT STORES	451
WIDE-COLUMN STORES	451
GRAPH-ORIENTED DATABASES	451
NoSQL Examples	451
REDIS	451
MONGODB	452
APACHE CASSANDRA	452
NEO4J	452
A NoSQL Example: MongoDB	452
DOCUMENTS	452
COLLECTIONS	454
RELATIONSHIPS	454
QUERYING MONGODB	455
Impact of NoSQL on Database Professionals	456
Hadoop	458
Components of Hadoop	459
THE HADOOP DISTRIBUTED FILE SYSTEM (HDFS)	459
MAPREDUCE	459
PIG	461
HIVE	461
HBASE	462
A Practical Introduction to Pig	462
LOADING DATA	462
TRANSFORMING DATA	463
A Practical Introduction to Hive	465
CREATING A TABLE	465
LOADING DATA INTO THE TABLE	465
PROCESSING THE DATA	466
Integrated Analytics and Data Science Platforms	466
HP HAVEN	466
TERADATA ASTER	467
IBM BIG DATA PLATFORM	469

Putting It All Together: Integrated Data Architecture 469

Summary 471 • *Key Terms* 471 • *Review Questions* 471 •
Problems and Exercises 472 • *References* 472 •
Further Reading 473 • *Web Resources* 473

Chapter 11 Analytics and Its Implications 474

Learning Objectives 474

Introduction 474

Analytics 475

Types of Analytics 475

Use of Descriptive Analytics 477

SQL OLAP QUERYING 478

OLAP TOOLS 480

DATA VISUALIZATION 482

BUSINESS PERFORMANCE MANAGEMENT AND DASHBOARDS 483

Use of Predictive Analytics 484

DATA MINING TOOLS 485

EXAMPLES OF PREDICTIVE ANALYTICS 486

Use of Prescriptive Analytics 487

Key User Tools for Analytics 488

ANALYTICAL AND OLAP FUNCTIONS 489

R 490

PYTHON 491

APACHE SPARK 492

Data Management Infrastructure for Analytics 493

Impact of Big Data and Analytics 495

Applications of Big Data and Analytics 495

BUSINESS 496

E-GOVERNMENT AND POLITICS 496

SCIENCE AND TECHNOLOGY 496

SMART HEALTH AND WELL-BEING 497

SECURITY AND PUBLIC SAFETY 497

Implications of Big Data Analytics and Decision Making 497

PERSONAL PRIVACY VERSUS COLLECTIVE BENEFITS 498

OWNERSHIP AND ACCESS 498

QUALITY AND REUSE OF DATA AND ALGORITHMS 498

TRANSPARENCY AND VALIDATION 498

CHANGING NATURE OF WORK 499

DEMANDS FOR WORKFORCE CAPABILITIES AND EDUCATION 499

Summary 499 • *Key Terms* 500 • *Review Questions* 500 •
Problems and Exercises 500 • *References* 501 •
Further Reading 502

Chapter 12 Data and Database Administration with Focus on Data Quality 503

Learning Objectives 503

Introduction 503

Overview of Data and Database Administration 505

Data Administration 505

Database Administration 506

TRADITIONAL DATABASE ADMINISTRATION	506
TRENDS IN DATABASE ADMINISTRATION	508
Evolving Data Administration Roles	510
The Open Source Movement and Database Management	511
Data Governance	512
Managing Data Quality	513
Characteristics of Quality Data	514
EXTERNAL DATA SOURCES	515
REDUNDANT DATA STORAGE AND INCONSISTENT METADATA	516
DATA ENTRY PROBLEMS	516
LACK OF ORGANIZATIONAL COMMITMENT	516
Data Quality Improvement	516
GET THE BUSINESS BUY-IN	516
CONDUCT A DATA QUALITY AUDIT	517
ESTABLISH A DATA STEWARDSHIP PROGRAM	518
IMPROVE DATA CAPTURE PROCESSES	518
APPLY MODERN DATA MANAGEMENT PRINCIPLES AND TECHNOLOGY	519
APPLY TQM PRINCIPLES AND PRACTICES	519
Summary of Data Quality	519
Data Availability	520
Costs of Downtime	520
Measures to Ensure Availability	521
HARDWARE FAILURES	521
LOSS OR CORRUPTION OF DATA	521
HUMAN ERROR	521
MAINTENANCE DOWNTIME	521
NETWORK-RELATED PROBLEMS	521
Master Data Management	521
<i>Summary</i>	<i>523</i>
<i>Key Terms</i>	<i>523</i>
<i>Review Questions</i>	<i>524</i>
<i>Problems and Exercises</i>	<i>524</i>
<i>Field Exercises</i>	<i>526</i>
<i>References</i>	<i>526</i>
<i>Further Reading</i>	<i>527</i>
<i>Web Resources</i>	<i>527</i>
<i>Glossary of Acronyms</i>	<i>529</i>
<i>Glossary of Terms</i>	<i>531</i>
<i>Index</i>	<i>539</i>

ONLINE CHAPTERS

Chapter 13 Distributed Databases 13-1

Learning Objectives 13-1

Introduction 13-1

Objectives and Trade-Offs 13-4

Options for Distributing a Database 13-6

Data Replication 13-6

SNAPSHOT REPLICATION 13-7

NEAR-REAL-TIME REPLICATION 13-8

PULL REPLICATION 13-8

DATABASE INTEGRITY WITH REPLICATION 13-8

WHEN TO USE REPLICATION 13-9

Horizontal Partitioning 13-9

Vertical Partitioning 13-10

Combinations of Operations 13-11

Selecting the Right Data Distribution Strategy 13-12

Distributed DBMS 13-13

Location Transparency 13-15

Replication Transparency 13-16

Failure Transparency 13-17

Commit Protocol 13-17

Concurrency Transparency 13-18

TIME STAMPING 13-19

Query Optimization 13-19

Evolution of Distributed DBMSs 13-22

REMOTE UNIT OF WORK 13-22

DISTRIBUTED UNIT OF WORK 13-22

DISTRIBUTED REQUEST 13-23

Summary 13-23 • Key Terms 13-24 • Review Questions 13-24 •

Problems and Exercises 13-25 • Field Exercises 13-27 •

References 13-27 • Further Reading 13-27 •

Web Resources 13-27

Chapter 14 Object-Oriented Data Modeling 14-1

Learning Objectives 14-1

Introduction 14-1

Unified Modeling Language 14-3

Object-Oriented Data Modeling 14-4

Representing Objects and Classes 14-4

Types of Operations 14-7

Representing Associations 14-7

Representing Association Classes 14-11

Representing Derived Attributes, Derived Associations,
and Derived Roles 14-12

Representing Generalization 14-13

Interpreting Inheritance and Overriding 14-18

- Representing Multiple Inheritance 14-19
- Representing Aggregation 14-19
- Business Rules 14-22
- Object Modeling Example: Pine Valley Furniture Company 14-23
 - Summary* 14-25 • *Key Terms* 14-26 • *Review Questions* 14-26 •
 - Problems and Exercises* 14-30 • *Field Exercises* 14-37 •
 - References* 14-37 • *Further Reading* 14-38 •
 - Web Resources* 14-38

Appendix A Data Modeling Tools and Notation A-1

- Comparing E-R Modeling Conventions A-1
 - Visio Professional 2016 Notation A-1
 - ENTITIES A-5
 - RELATIONSHIPS A-5
 - CA ERwin Data Modeler 9.7 Notation A-5
 - ENTITIES A-5
 - RELATIONSHIPS A-5
 - SAP Sybase PowerDesigner 16.6 Notation A-7
 - ENTITIES A-8
 - RELATIONSHIPS A-8
 - Oracle Designer Notation A-8
 - ENTITIES A-8
 - RELATIONSHIPS A-8
- Comparison of Tool Interfaces and E-R Diagrams A-8

Appendix B Advanced Normal Forms B-1

- Boyce-Codd Normal Form B-1
 - Anomalies in Student Advisor B-1
 - Definition of Boyce-Codd Normal Form (BCNF) B-2
 - Converting a Relation to BCNF B-2
- Fourth Normal Form B-3
 - Multivalued Dependencies B-5
- Higher Normal Forms B-5
 - Key Terms* B-6 • *References* B-6 •
 - Web Resources* B-6

Appendix C Data Structures C-1

- Pointers C-1
- Data Structure Building Blocks C-2
- Linear Data Structures C-4
 - Stacks C-5
 - Queues C-5
 - Sorted Lists C-6
 - Multilists C-8
- Hazards of Chain Structures C-8
- Trees C-9
 - Balanced Trees C-9
 - References* C-12

PREFACE

This text is designed for introductory courses in database management. Such a course is usually required as part of an information systems curriculum in business schools, computer technology programs, and applied computer science departments. The Association for Information Systems (AIS), the Association for Computing Machinery (ACM), and the International Federation of Information Processing Societies (IFIPS) curriculum guidelines (e.g., IS 2010 and MSIS 2016) all outline this type of database management course or the competencies a student completing the course is expected to have. Previous editions of this text have been used successfully for more than 35 years at both the undergraduate and graduate levels as well as in management and professional development programs.

WHAT'S NEW IN THIS EDITION?

This 13th edition of *Modern Database Management* updates and expands materials in areas undergoing rapid change as a result of improved managerial practices, database design tools and methodologies, and database technology. Later, we detail changes to each chapter. The themes of this 13th edition reflect the major trends in the information systems field and the skills required of modern information systems graduates. The most important changes are as follows:

- The book has been restructured in several important ways. Chapter 7 on databases in applications now also includes segments on transaction integrity, designing multi-user solutions, and application level security, bringing these important perspectives together with their context. The revised chapter on physical database design and database infrastructure (new Chapter 8) includes also coverage of database security, backup and recovery, cloud-based database solutions, and other essential database infrastructure topics. This new comprehensive structure on physical design and infrastructure is now placed after the SQL chapters. The new version of Chapter 9 integrates material on data warehousing and data integrity in a conceptually natural pairing. Recognizing the way in which analytics capabilities rely on all types of data management solutions, Chapter 11, on analytics and implications, is now separate from Chapter 10, on big data. Finally, Chapter 12 brings together data and database administration with data quality, emphasizing the essential connections between the three.
- The part structure of the book has been redesigned to be fully aligned with the new chapter structure.
- We have introduced a new overarching framework (Figure 1-5), which gives our readers a clearer overview of structure of the book and its core topic areas. The framework communicates clearly the increasing importance of informational systems (divided into Analytics–Data Warehousing and Analytics–Big Data) in addition to this book's traditional strength of transactional systems.
- Given the continued and still increasing interest in big data and analytics, we have continued to expand content in this area. The book has now separate chapters on big data technologies (Chapter 10) and analytics (Chapter 11). In addition to general coverage of NoSQL and Hadoop technologies, Chapter 10 provides also detailed examples of MongoDB, Pig, and Hive. Chapter 11 includes extended coverage of R, Python, and Apache Spark—all essential technologies for analytics professionals that allow a link between analytics and data management architectures.
- We emphasize the increasing importance of cloud-based database solutions, mobile technologies, and agile development throughout the book.
- Chapter 1 now better recognizes the broad range of enterprise level applications data management solutions enable and support, including enterprise systems, data warehouses, and data lakes.

- Chapter 7 on databases in applications now includes an extensive example demonstrating the use of Python in the context of database-driven applications.
- The instructor's manual will have more material to support the case Forondo Artist Management Excellence that was introduced in the 12th edition.

In addition to the new topics covered, specific improvements to the textbook have been made in the following areas:

- Every chapter went through significant edits to streamline coverage to ensure relevance with current technologies and eliminate redundancies.
- The entire book has been edited so that its language clearly reflects its focus on the readers as learners instead of authors as teachers
- End-of-chapter material (review questions, problems and exercises, and/or field exercises) in every chapter has been revised with new and modified questions and exercises.
- We continued to update the figures in several chapters to reflect the changing landscape of technologies that are being used in modern organizations.
- The Web Resources section in each chapter was updated to ensure that students have information on the latest database trends and expanded background details on important topics covered in the text.
- The book continues to be available through VitalSource, an innovative e-book delivery system, and as an electronic book in the Kindle format.

Also, we continue to provide on the student Companion Web site several custom-developed short videos that address key concepts and skills from different sections of the book. These videos, produced by the textbook authors, help students learn difficult material by using both the printed text and a mini-lecture or tutorial. Videos have been developed to support Chapters 1 (introduction to database), 2 and 3 (conceptual data modeling), 4 (normalization), and 6 and 7 (SQL). Look for special icons on the opening page of these chapters to call attention to these videos, and go to www.pearsonhighered.com/hoffer to find these videos.

FOR THOSE NEW TO MODERN DATABASE MANAGEMENT

Modern Database Management has been a leading text since its first edition in 1983. In spite of this market leadership position, some instructors have used other good database management texts. Why might you want to switch at this time? There are several good reasons:

- One of our goals, in every edition, has been to lead other books in coverage of the latest principles, concepts, and technologies. See what we have added for the 13th edition in "What's New in This Edition?" In the past, we have led in coverage of object-oriented data modeling and UML, Internet databases, data warehousing, and the use of CASE tools in support of data modeling. For the 13th edition, we continue this tradition by continuing to expand and improve coverage of big data and analytics, focusing on what every database student needs to understand about these topics.
- While remaining current, this text focuses on what leading practitioners say is most important for database developers. We work with many practitioners, including the professionals of the Data Management Association (DAMA) and The Data Warehousing Institute (TDWI), leading consultants, technology leaders, and authors of articles in the most widely read professional publications. We draw on these experts to ensure that what the book includes is important and covers not only important entry-level knowledge and skills but also those fundamentals and mind-sets that lead to long-term career success.
- In the 13th edition of this highly successful book, material is presented in a way that has been viewed as very accessible to students. Our methods have been refined through continuous market feedback for more than 35 years as well as through our own teaching. Overall, the pedagogy of the book is sound, and we believe that the new framework that we introduced in Chapter 1 will further strengthen our students'