



DATABASE SYSTEMS

Design, Implementation,
and Management

13e

Carlos Coronel | Steven Morris



Australia • Brazil • Mexico • Singapore • United Kingdom • United States

Brief Contents

Preface, xv

Text Features, xx

Additional Features, xxii

Acknowledgments, xxiv

Part 1: Database Concepts 1

1. Database Systems, 2

2. Data Models, 34

Part 2: Design Concepts 67

3. The Relational Database Model, 68

4. Entity Relationship (ER) Modeling, 113

5. Advanced Data Modeling, 167

6. Normalization of Database Tables, 199

Part 3: Advanced Design and Implementation 243

7. Introduction to Structured Query Language (SQL), 244

8. Advanced SQL, 359

9. Database Design, 439

Part 4: Advanced Database Concepts 481

10. Transaction Management and Concurrency Control, 482

11. Database Performance Tuning and Query Optimization, 515

12. Distributed Database Management Systems, 553

13. Business Intelligence and Data Warehouses, 589

14. Big Data and NoSQL, 657

Part 5: Databases and the Internet 691

15. Database Connectivity and Web Technologies, 692

Part 6: Database Administration 733

16. Database Administration and Security, 734

Glossary, 782

Index, 793

The following appendices are included on the Instructor and Student Companion Sites at www.cengagebrain.com.

- Appendix A1:** Designing Databases with Visio Professional 2010: A Tutorial
- Appendix A2:** Designing Databases with Visio 2013: A Tutorial
- Appendix B:** The University Lab: Conceptual Design
- Appendix C:** The University Lab: Conceptual Design Verification, Logical Design, and Implementation
- Appendix D:** Converting an ER Model into a Database Structure
- Appendix E:** Comparison of ER Modeling Notations
- Appendix F:** Client/Server Systems
- Appendix G:** Object-Oriented Databases
- Appendix H:** Unified Modeling Language (UML)
- Appendix I:** Databases in Electronic Commerce
- Appendix J:** Web Database Development with ColdFusion
- Appendix K:** The Hierarchical Database Model
- Appendix L:** The Network Database Model
- Appendix M:** MS Access Tutorial
- Appendix N:** Creating a New Database Using Oracle 12c
- Appendix O:** Data Warehouse Implementation Factors
- Appendix P:** Working with MongoDB
- Appendix Q:** Working with Neo4j

Contents

Preface, xv

Text Features, xx

Additional Features, xxii

Acknowledgments, xxiv

Part 1: Database Concepts 1

Chapter 1: Database Systems 2

- 1-1** Why Databases? 3
- 1-2** Data versus Information 4
- 1-3** Introducing the Database 6
 - 1-3a Role and Advantages of the DBMS 7
 - 1-3b Types of Databases 9
- 1-4** Why Database Design Is Important 12
- 1-5** Evolution of File System Data Processing 15
 - 1-5a Manual File Systems 15
 - 1-5b Computerized File Systems 15
 - 1-5c File System Redux: Modern End-User Productivity Tools 18
- 1-6** Problems with File System Data Processing 18
 - 1-6a Structural and Data Dependence 19
 - 1-6b Data Redundancy 20
 - 1-6c Data Anomalies 21
- 1-7** Database Systems 21
 - 1-7a The Database System Environment 22
 - 1-7b DBMS Functions 24
 - 1-7c Managing the Database System: A Shift in Focus 27
- 1-8** Preparing for Your Database Professional Career 28
 - Summary 30 • Key Terms 31 • Review Questions 31 • Problems 32

Chapter 2: Data Models 34

- 2-1** Data Modeling and Data Models 35
- 2-2** The Importance of Data Models 36
- 2-3** Data Model Basic Building Blocks 36
- 2-4** Business Rules 38
 - 2-4a Discovering Business Rules 38
 - 2-4b Translating Business Rules into Data Model Components 39
 - 2-4c Naming Conventions 40
- 2-5** The Evolution of Data Models 40
 - 2-5a Hierarchical and Network Models 40
 - 2-5b The Relational Model 42
 - 2-5c The Entity Relationship Model 44
 - 2-5d The Object-Oriented Model 47
 - 2-5e Object/Relational and XML 48
 - 2-5f Emerging Data Models: Big Data and NoSQL 49
 - 2-5g Data Models: A Summary 53
- 2-6** Degrees of Data Abstraction 54
 - 2-6a The External Model 57
 - 2-6b The Conceptual Model 58
 - 2-6c The Internal Model 59
 - 2-6d The Physical Model 60
 - Summary 61 • Key Terms 62 • Review Questions 62 • Problems 63

Part 2: Design Concepts 67

Chapter 3: The Relational Database Model 68

- 3-1 A Logical View of Data 69**
 - 3-1a Tables and Their Characteristics 69
- 3-2 Keys 72**
 - 3-2a Dependencies 72
 - 3-2b Types of Keys 73
- 3-3 Integrity Rules 76**
- 3-4 Relational Algebra 78**
 - 3-4a Formal Definitions and Terminology 78
 - 3-4b Relational Set Operators 79
- 3-5 The Data Dictionary and the System Catalog 87**
- 3-6 Relationships within the Relational Database 89**
 - 3-6a The 1:M Relationship 89
 - 3-6b The 1:1 Relationship 91
 - 3-6c The M:N Relationship 93
- 3-7 Data Redundancy Revisited 97**
- 3-8 Indexes 99**
- 3-9 Codd's Relational Database Rules 100**
 - Summary 102 • Key Terms 103 • Review Questions 103 • Problems 106

Chapter 4: Entity Relationship (ER) Modeling 113

- 4-1 The Entity Relationship Model 114**
 - 4-1a Entities 114
 - 4-1b Attributes 114
 - 4-1c Relationships 120
 - 4-1d Connectivity and Cardinality 121
 - 4-1e Existence Dependence 122
 - 4-1f Relationship Strength 123
 - 4-1g Weak Entities 125
 - 4-1h Relationship Participation 127
 - 4-1i Relationship Degree 131
 - 4-1j Recursive Relationships 133
 - 4-1k Associative (Composite) Entities 136
- 4-2 Developing an ER Diagram 138**
- 4-3 Database Design Challenges: Conflicting Goals 146**
 - Summary 150 • Key Terms 151 • Review Questions 151 • Problems 154 • Cases 159

Chapter 5: Advanced Data Modeling 167

- 5-1 The Extended Entity Relationship Model 168**
 - 5-1a Entity Supertypes and Subtypes 168
 - 5-1b Specialization Hierarchy 169
 - 5-1c Inheritance 170
 - 5-1d Subtype Discriminator 172
 - 5-1e Disjoint and Overlapping Constraints 172
 - 5-1f Completeness Constraint 174
 - 5-1g Specialization and Generalization 175
- 5-2 Entity Clustering 175**
- 5-3 Entity Integrity: Selecting Primary Keys 176**
 - 5-3a Natural Keys and Primary Keys 177
 - 5-3b Primary Key Guidelines 177
 - 5-3c When to Use Composite Primary Keys 177
 - 5-3d When to Use Surrogate Primary Keys 179
- 5-4 Design Cases: Learning Flexible Database Design 180**
 - 5-4a Design Case 1: Implementing 1:1 Relationships 181
 - 5-4b Design Case 2: Maintaining History of Time-Variant Data 182
 - 5-4c Design Case 3: Fan Traps 185
 - 5-4d Design Case 4: Redundant Relationships 186
 - Summary 187 • Key Terms 187 • Review Questions 188 • Problems 189 • Cases 190

Chapter 6: Normalization of Database Tables 199

- 6-1 Database Tables and Normalization 200**
- 6-2 The Need for Normalization 200**
- 6-3 The Normalization Process 203**
 - 6-3a Conversion to First Normal Form (1NF) 205
 - 6-3b Conversion to Second Normal Form (2NF) 209
 - 6-3c Conversion to Third Normal Form (3NF) 211
- 6-4 Improving the Design 213**
- 6-5 Surrogate Key Considerations 217**
- 6-6 Higher-Level Normal Forms 218**
 - 6-6a The Boyce-Codd Normal Form 219
 - 6-6b Fourth Normal Form (4NF) 222
- 6-7 Normalization and Database Design 224**
- 6-8 Denormalization 227**
- 6-9 Data-Modeling Checklist 230**
 - Summary 232 • Key Terms 233 • Review Questions 233 • Problems 235

Part 3: Advanced Design and Implementation 243

Chapter 7: Introduction to Structured Query Language (SQL) 244

- 7-1 Introduction to SQL 245**
 - 7-1a Data Types 245
 - 7-1b SQL Queries 247
 - 7-1c The Database Model 248
- 7-2 Basic SELECT Queries 249**
- 7-3 SELECT Statement Options 250**
 - 7-3a Using Column Aliases 251
 - 7-3b Using Computed Columns 253
 - 7-3c Arithmetic Operators: The Rule of Precedence 254
 - 7-3d Date Arithmetic 255
 - 7-3e Listing Unique Values 255
- 7-4 FROM Clause Options 256**
 - 7-4a Natural Join 257
 - 7-4b JOIN USING Syntax 259
 - 7-4c JOIN ON Syntax 260
 - 7-4d Common Attribute Names 261
 - 7-4e Outer Joins 261
 - 7-4f Cross Join 264
 - 7-4g Joining Tables with an Alias 264
 - 7-4h Recursive Joins 265
- 7-5 ORDER BY Clause Options 266**
- 7-6 WHERE Clause Options 269**
 - 7-6a Selecting Rows with Conditional Restrictions 269
 - 7-6b Using Comparison Operators on Character Attributes 271
 - 7-6c Using Comparison Operators on Dates 272
 - 7-6d Logical Operators: AND, OR, and NOT 273
 - 7-6e Old-Style Joins 275
 - 7-6f Special Operators 276
- 7-7 Aggregate Processing 281**
 - 7-7a Aggregate Functions 281
 - 7-7b Grouping Data 285
 - 7-7c HAVING Clause 288
- 7-8 Subqueries 290**
 - 7-8a WHERE Subqueries 292
 - 7-8b IN Subqueries 293
 - 7-8c HAVING Subqueries 294
 - 7-8d Multirow Subquery Operators: ALL and ANY 294
 - 7-8e FROM Subqueries 295
 - 7-8f Attribute List Subqueries 296
 - 7-8g Correlated Subqueries 298
- 7-9 SQL Functions 302**
 - 7-9a Date and Time Functions 302
 - 7-9b Numeric Functions 306

- 7-9c String Functions 307
 - 7-9d Conversion Functions 309
 - 7-10 Relational Set Operators 311**
 - 7-10a UNION 311
 - 7-10b UNION ALL 313
 - 7-10c INTERSECT 314
 - 7-10d EXCEPT (MINUS) 315
 - 7-10e Syntax Alternatives 316
 - 7-11 Crafting SELECT Queries 317**
 - 7-11a Know Your Data 317
 - 7-11b Know the Problem 317
 - 7-11c Build One Clause at a Time 318
- Summary 319 • Key Terms 321 • Review Questions 321 • Problems 323

Chapter 8: Advanced SQL 359

- 8-1 Data Definition Commands 360**
 - 8-1a Starting Database Model 360
 - 8-1b Creating the Database 361
 - 8-1c The Database Schema 362
 - 8-1d Data Types 362
- 8-2 Creating Table Structures 366**
 - 8-2a CREATE TABLE command 366
 - 8-2b SQL Constraints 370
 - 8-2c Create a Table with a SELECT Statement 373
 - 8-2d SQL Indexes 374
- 8-3 Altering Table Structures 375**
 - 8-3a Changing a Column's Data Type 376
 - 8-3b Changing a Column's Data Characteristics 376
 - 8-3c Adding a Column 377
 - 8-3d Adding Primary Key, Foreign Key, and Check Constraints 377
 - 8-3e Dropping a Column 378
 - 8-3f Deleting a Table from the Database 378
- 8-4 Data Manipulation Commands 379**
 - 8-4a Adding Table Rows 379
 - 8-4b Inserting Table Rows with a SELECT Subquery 381
 - 8-4c Saving Table Changes 382
 - 8-4d Updating Table Rows 383
 - 8-4e Deleting Table Rows 385
 - 8-4f Restoring Table Contents 386
- 8-5 Virtual Tables: Creating a View 387**
 - 8-5a Updatable Views 388
- 8-6 Sequences 391**
- 8-7 Procedural SQL 396**
 - 8-7a Triggers 401
 - 8-7b Stored Procedures 411
 - 8-7c PL/SQL Processing with Cursors 416
 - 8-7d PL/SQL Stored Functions 418
- 8-8 Embedded SQL 419**
 - Summary 423 • Key Terms 425 • Review Questions 425 • Problems 426 • Cases 433

Chapter 9: Database Design 439

- 9-1 The Information System 440**
- 9-2 The Systems Development Life Cycle 442**
 - 9-2a Planning 442
 - 9-2b Analysis 443
 - 9-2c Detailed Systems Design 444
 - 9-2d Implementation 444
 - 9-2e Maintenance 445
- 9-3 The Database Life Cycle 445**
 - 9-3a The Database Initial Study 445
 - 9-3b Database Design 450
 - 9-3c Implementation and Loading 451
 - 9-3d Testing and Evaluation 454

- 9-3e Operation 456
- 9-3f Maintenance and Evolution 457
- 9-4 Conceptual Design 457**
 - 9-4a Data Analysis and Requirements 459
 - 9-4b Entity Relationship Modeling and Normalization 461
 - 9-4c Data Model Verification 464
 - 9-4d Distributed Database Design 467
- 9-5 DBMS Software Selection 467**
- 9-6 Logical Design 468**
 - 9-6a Map the Conceptual Model to the Logical Model 468
 - 9-6b Validate the Logical Model Using Normalization 470
 - 9-6c Validate Logical Model Integrity Constraints 470
 - 9-6d Validate the Logical Model against User Requirements 471
- 9-7 Physical Design 471**
 - 9-7a Define Data Storage Organization 472
 - 9-7b Define Integrity and Security Measures 472
 - 9-7c Determine Performance Measures 473
- 9-8 Database Design Strategies 473**
- 9-9 Centralized versus Decentralized Design 474**
 - Summary 477 • Key Terms 477 • Review Questions 477 • Problems 478

Part 4: Advanced Database Concepts 481

Chapter 10: Transaction Management and Concurrency Control 482

- 10-1 What Is a Transaction? 483**
 - 10-1a Evaluating Transaction Results 484
 - 10-1b Transaction Properties 487
 - 10-1c Transaction Management with SQL 488
 - 10-1d The Transaction Log 489
- 10-2 Concurrency Control 490**
 - 10-2a Lost Updates 490
 - 10-2b Uncommitted Data 491
 - 10-2c Inconsistent Retrievals 492
 - 10-2d The Scheduler 493
- 10-3 Concurrency Control with Locking Methods 495**
 - 10-3a Lock Granularity 496
 - 10-3b Lock Types 498
 - 10-3c Two-Phase Locking to Ensure Serializability 500
 - 10-3d Deadlocks 500
- 10-4 Concurrency Control with Time Stamping Methods 502**
 - 10-4a Wait/Die and Wound/Wait Schemes 502
- 10-5 Concurrency Control with Optimistic Methods 503**
- 10-6 ANSI Levels of Transaction Isolation 504**
- 10-7 Database Recovery Management 506**
 - 10-7a Transaction Recovery 506
- Summary 510 • Key Terms 511 • Review Questions 511 • Problems 512

Chapter 11: Database Performance Tuning and Query Optimization 515

- 11-1 Database Performance-Tuning Concepts 516**
 - 11-1a Performance Tuning: Client and Server 517
 - 11-1b DBMS Architecture 518
 - 11-1c Database Query Optimization Modes 520
 - 11-1d Database Statistics 521
- 11-2 Query Processing 522**
 - 11-2a SQL Parsing Phase 523
 - 11-2b SQL Execution Phase 524
 - 11-2c SQL Fetching Phase 525
 - 11-2d Query Processing Bottlenecks 525
- 11-3 Indexes and Query Optimization 526**
- 11-4 Optimizer Choices 528**
 - 11-4a Using Hints to Affect Optimizer Choices 530

- 11-5 SQL Performance Tuning 531**
 - 11-5a Index Selectivity 531
 - 11-5b Conditional Expressions 533
- 11-6 Query Formulation 534**
- 11-7 DBMS Performance Tuning 536**
- 11-8 Query Optimization Example 538**
 - Summary 546 • Key Terms 547 • Review Questions 547 • Problems 548

Chapter 12: Distributed Database Management Systems 553

- 12-1 The Evolution of Distributed Database Management Systems 554**
- 12-2 DDBMS Advantages and Disadvantages 556**
- 12-3 Distributed Processing and Distributed Databases 556**
- 12-4 Characteristics of Distributed Database Management Systems 559**
- 12-5 DDBMS Components 560**
- 12-6 Levels of Data and Process Distribution 561**
 - 12-6a Single-Site Processing, Single-Site Data 561
 - 12-6b Multiple-Site Processing, Single-Site Data 562
 - 12-6c Multiple-Site Processing, Multiple-Site Data 563
- 12-7 Distributed Database Transparency Features 564**
- 12-8 Distribution Transparency 565**
- 12-9 Transaction Transparency 568**
 - 12-9a Distributed Requests and Distributed Transactions 568
 - 12-9b Distributed Concurrency Control 571
 - 12-9c Two-Phase Commit Protocol 571
- 12-10 Performance and Failure Transparency 573**
- 12-11 Distributed Database Design 575**
 - 12-11a Data Fragmentation 575
 - 12-11b Data Replication 578
 - 12-11c Data Allocation 580
- 12-12 The CAP Theorem 581**
- 12-13 C. J. Date's 12 Commandments for Distributed Databases 583**
 - Summary 584 • Key Terms 585 • Review Questions 585 • Problems 586

Chapter 13: Business Intelligence and Data Warehouses 589

- 13-1 The Need for Data Analysis 590**
- 13-2 Business Intelligence 590**
 - 13-2a Business Intelligence Architecture 592
 - 13-2b Business Intelligence Benefits 596
 - 13-2c Business Intelligence Evolution 597
 - 13-2d Business Intelligence Technology Trends 600
- 13-3 Decision Support Data 601**
 - 13-3a Operational Data versus Decision Support Data 601
 - 13-3b Decision Support Database Requirements 604
- 13-4 The Data Warehouse 606**
 - 13-4a Data Marts 609
 - 13-4b Twelve Rules That Define a Data Warehouse 609
- 13-5 Star Schemas 609**
 - 13-5a Facts 610
 - 13-5b Dimensions 610
 - 13-5c Attributes 611
 - 13-5d Attribute Hierarchies 613
 - 13-5e Star Schema Representation 615
 - 13-5f Performance-Improving Techniques for the Star Schema 616
- 13-6 Online Analytical Processing 620**
 - 13-6a Multidimensional Data Analysis Techniques 620
 - 13-6b Advanced Database Support 622
 - 13-6c Easy-to-Use End-User Interfaces 622
 - 13-6d OLAP Architecture 622
 - 13-6e Relational OLAP 625
 - 13-6f Multidimensional OLAP 627
 - 13-6g Relational versus Multidimensional OLAP 627

- 13-7 Data Analytics 628**
 - 13-7a Data Mining 629
 - 13-7b Predictive Analytics 631
 - 13-8 SQL Analytic Functions 632**
 - 13-8a The ROLLUP Extension 633
 - 13-8b The CUBE Extension 634
 - 13-8c Materialized Views 636
 - 13-9 Data Visualization 639**
 - 13-9a The Need for Data Visualization 640
 - 13-9b The Science of Data Visualization 642
 - 13-9c Understanding the Data 644
- Summary 645 • Key Terms 646 • Review Questions 647 • Problems 648

Chapter 14: Big Data and NoSQL 657

- 14-1 Big Data 658**
 - 14-1a Volume 660
 - 14-1b Velocity 661
 - 14-1c Variety 662
 - 14-1d Other Characteristics 663
 - 14-2 Hadoop 664**
 - 14-2a HDFS 665
 - 14-2b MapReduce 667
 - 14-2c Hadoop Ecosystem 669
 - 14-3 NoSQL 672**
 - 14-3a Key-Value Databases 673
 - 14-3b Document Databases 674
 - 14-3c Column-Oriented Databases 675
 - 14-3d Graph Databases 677
 - 14-3e Aggregate Awareness 679
 - 14-4 NewSQL Databases 680**
 - 14-5 Working with Document Databases Using MongoDB 680**
 - 14-5a Importing Documents in MongoDB 682
 - 14-5b Example of a MongoDB Query Using find() 683
 - 14-6 Working with Graph Databases Using Neo4j 684**
 - 14-6a Creating Nodes in Neo4j 685
 - 14-6b Retrieving Node Data with MATCH and WHERE 686
 - 14-6c Retrieving Relationship Data with MATCH and WHERE 686
- Summary 688 • Key Terms 689 • Review Questions 690

Part 5: Databases and the Internet 691

Chapter 15: Database Connectivity and Web Technologies 692

- 15-1 Database Connectivity 693**
 - 15-1a Native SQL Connectivity 694
 - 15-1b ODBC, DAO, and RDO 695
 - 15-1c OLE-DB 697
 - 15-1d ADO.NET 699
 - 15-1e Java Database Connectivity (JDBC) 703
- 15-2 Database Internet Connectivity 704**
 - 15-2a Web-to-Database Middleware: Server-Side Extensions 705
 - 15-2b Web Server Interfaces 707
 - 15-2c The Web Browser 708
 - 15-2d Client-Side Extensions 709
 - 15-2e Web Application Servers 710
 - 15-2f Web Database Development 711
- 15-3 Extensible Markup Language (XML) 715**
 - 15-3a Document Type Definitions (DTD) and XML Schemas 717
 - 15-3b XML Presentation 719
 - 15-3c XML Applications 721

- 15-4 Cloud Computing Services 722**
 - 15-4a Cloud Implementation Types 725
 - 15-4b Characteristics of Cloud Services 725
 - 15-4c Types of Cloud Services 726
 - 15-4d Cloud Services: Advantages and Disadvantages 727
 - 15-4e SQL Data Services 729
- Summary 730 • Key Terms 731 • Review Questions 731 • Problems 732

Part 6: Database Administration 733

Chapter 16: Database Administration and Security 734

- 16-1 Data as a Corporate Asset 735**
- 16-2 The Need for a Database and Its Role in an Organization 736**
- 16-3 Introduction of a Database: Special Considerations 738**
- 16-4 The Evolution of Database Administration 739**
- 16-5 The Database Environment's Human Component 743**
 - 16-5a The DBA's Managerial Role 745
 - 16-5b The DBA's Technical Role 750
- 16-6 Security 757**
 - 16-6a Security Policies 758
 - 16-6b Security Vulnerabilities 758
 - 16-6c Database Security 760
- 16-7 Database Administration Tools 761**
 - 16-7a The Data Dictionary 762
 - 16-7b Case Tools 764
- 16-8 Developing a Data Administration Strategy 767**
- 16-9 The DBA's Role in the Cloud 768**
- 16-10 The DBA at Work: Using Oracle for Database Administration 769**
 - 16-10a Oracle Database Administration Tools 770
 - 16-10b Ensuring That the RDBMS Starts Automatically 770
 - 16-10c Creating Tablespaces and Datafiles 772
 - 16-10d Managing Users and Establishing Security 774
 - 16-10e Customizing the Database Initialization Parameters 776
- Summary 777 • Key Terms 779 • Review Questions 779
- Glossary 782
- Index 793

The following appendices are included on the Instructor and Student Companion Sites at www.cengagebrain.com.

Appendix A1: Designing Databases with Visio Professional 2010: A Tutorial

Appendix A2: Designing Databases with Visio 2013: A Tutorial

Appendix B: The University Lab: Conceptual Design

Appendix C: The University Lab: Conceptual Design Verification, Logical Design, and Implementation

Appendix D: Converting an ER Model into a Database Structure

Appendix E: Comparison of ER Modeling Notations

Appendix F: Client/Server Systems

Appendix G: Object-Oriented Databases

Appendix H: Unified Modeling Language (UML)

Appendix I: Databases in Electronic Commerce

Appendix J: Web Database Development with ColdFusion

Appendix K: The Hierarchical Database Model

Appendix L: The Network Database Model

Appendix M: MS Access Tutorial

Appendix N: Creating a New Database Using Oracle 12c

Appendix O: Data Warehouse Implementation Factors

Appendix P: Working with MongoDB

Appendix Q: Working with Neo4j