

CONCEPTS OF
DATABASE
MANAGEMENT

EIGHTH EDITION

Phil Pratt • Mary Last

5 REASONS

to buy your textbooks
and course materials at

CENGAGE **brain**.com

- 1 SAVINGS:**
Prices up to 75% off, daily coupons, and free shipping on orders over \$25
- 2 CHOICE:**
Multiple format options including textbook, eBook and eChapter rentals
- 3 CONVENIENCE:**
Anytime, anywhere access of eBooks or eChapters via mobile devices
- 4 SERVICE:**
Free eBook access while your text ships, and instant access to online homework products
- 5 STUDY TOOLS:**
Study tools* for your text, plus writing, research, career and job search resources
**availability varies*



Find your course materials and start saving at:
www.cengagebrain.com

Source Code: 14M-AA0107

Engaged with you.
www.cengage.com

 **CENGAGE**
Learning®

Copyright 2015 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

CONCEPTS OF DATABASE MANAGEMENT

CONCEPTS OF DATABASE MANAGEMENT

Eighth Edition

Philip J. Pratt
Grand Valley State University
Mary Z. Last



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

This is an electronic version of the print textbook. Due to electronic rights restrictions, some third party content may be suppressed. Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. The publisher reserves the right to remove content from this title at any time if subsequent rights restrictions require it. For valuable information on pricing, previous editions, changes to current editions, and alternate formats, please visit www.cengage.com/highered to search by ISBN#, author, title, or keyword for materials in your areas of interest.

**Concepts of Database Management,
Eighth Edition**
Philip J. Pratt and Mary Z. Last

Product Manager: Clara Goosman
Content Developer: Wendy Langerud, S4 Carlisle
Managing Developer: Jeremy Judson
Product Assistant: Brad Sullender
Senior Marketing Manager: Eric La Scola
IP Analyst: Sara Crane
Senior IP Project Manager: Kathryn Kucharek
Manufacturing Planner: Ron Montgomery
Art and Design Direction, Production
Management, and Composition:
Lumina Datamatics, Inc.
Cover Image: © majcot/Shutterstock.com

© 2015, 2012 Cengage Learning®

WCN: 02-200-203

ALL RIGHTS RESERVED. No part of this work covered by the copyright herein may be reproduced, transmitted, stored, or used in any form or by any means graphic, electronic, or mechanical, including but not limited to photocopying, recording, scanning, digitizing, taping, Web distribution, information networks, or information storage and retrieval systems, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without the prior written permission of the publisher.

For product information and technology assistance, contact us at
Cengage Learning Customer & Sales Support, 1-800-354-9706.

For permission to use material from this text or product,
submit all requests online at www.cengage.com/permissions.

Further permissions questions can be emailed to
permissionrequest@cengage.com.

Some of the product names and company names used in this book have been used for identification purposes only and may be trademarks or registered trademarks of their respective manufacturers and sellers.

Library of Congress Control Number: 2014944683

ISBN: 978-1-285-42710-2

Cengage Learning

20 Channel Center Street
Boston, MA 02210
USA

Screenshots for this book were created using Microsoft Access®, and were used with permission from Microsoft.

Microsoft and the Office logo are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries. Cengage Learning is an independent entity from the Microsoft Corporation, and not affiliated with Microsoft in any manner.

Oracle is a registered trademark, and Oracle11g is a trademark of Oracle Corporation.

The programs in this book are for instructional purposes only. They have been tested with care, but are not guaranteed for any particular intent beyond educational purposes. The author and the publisher do not offer any warranties or representations, nor do they accept any liabilities with respect to the programs.

Cengage Learning, reserves the right to revise this publication and make changes from time to time in its content without notice.

Cengage Learning is a leading provider of customized learning solutions with office locations around the globe, including Singapore, the United Kingdom, Australia, Mexico, Brazil, and Japan. Locate your local office at: www.cengage.com/global.

Cengage Learning products are represented in Canada by Nelson Education, Ltd.

To learn more about Cengage Learning Solutions, visit www.cengage.com.

Purchase any of our products at your local college store or at our preferred online store www.cengagebrain.com.

TABLE OF CONTENTS

Preface	xi
Chapter 1 <i>Introduction to Database Management</i>	1
TAL Distributors Background	1
Database Background	4
Database Management Systems	10
Advantages of Database Processing	13
Disadvantages of Database Processing	15
Introduction to the Colonial Adventure Tours Database Case	16
Introduction to the Solmaris Condominium Group Database Case	21
Summary	25
Key Terms	25
Review Questions	25
TAL Distributors Exercises	26
Colonial Adventure Tours Case	26
Solmaris Condominium Group Case	27
Chapter 2 <i>The Relational Model 1: Introduction, QBE, and Relational Algebra</i>	29
Relational Databases	29
Query-By-Example (QBE)	32
Simple Queries	33
Simple Criteria	36
Compound Criteria	37
Computed Fields	41
Functions	42
Grouping	44
Sorting	45
Sorting on Multiple Keys	46
Joining Tables	48
Joining Multiple Tables	50
Using an Update Query	52
Using a Delete Query	52
Using a Make-Table Query	53
Relational Algebra	56
SELECT	56
PROJECT	57
JOIN	57
Normal Set Operations	59
Product	61
Division	61
Summary	63
Key Terms	64
Review Questions	64
TAL Distributors Exercises: QBE	65
TAL Distributors Exercises: Relational Algebra	66
Colonial Adventure Tours Case	67
Solmaris Condominium Group Case	68
Chapter 3 <i>The Relational Model 2: SQL</i>	71
Getting Started with SQL	71
Getting Started with Access 2010 or 2013	71

Table Creation	72
Simple Retrieval	74
Compound Conditions	80
Computed Fields	83
Using Special Operators (LIKE and IN)	85
Sorting	87
Built-In Functions	89
Subqueries	92
Grouping	93
Joining Tables	95
Union	99
Updating Tables	100
Creating a Table from a Query	101
Summary of SQL Commands	103
Summary	111
Key Terms	111
Review Questions	112
TAL Distributors Exercises	112
Colonial Adventure Tours Case	113
Solmaris Condominium Group Case	114
Chapter 4 <i>The Relational Model 3: Advanced Topics</i>	115
Views	115
Indexes	122
Security	126
Integrity Rules	127
Entity Integrity	127
Referential Integrity	128
Legal-Values Integrity	131
Structure Changes	132
Making Complex Changes	135
System Catalog	136
Stored Procedures	138
Triggers	138
Triggers in Access 2013	138
Summary	142
Key Terms	142
Review Questions	143
TAL Distributors Exercises	144
Colonial Adventure Tours Case	145
Solmaris Condominium Group Case	146
Chapter 5 <i>Database Design 1: Normalization</i>	149
Functional Dependence	151
Keys	153
First Normal Form	155
Second Normal Form	156
Third Normal Form	160
Incorrect Decompositions	163
Multivalued Dependencies and Fourth Normal Form	167
Avoiding the Problem with Multivalued Dependencies	170
Application to Database Design	171
Summary	173
Key Terms	173
Review Questions	173
TAL Distributors Exercises	174
Colonial Adventure Tours Case	175
Solmaris Condominium Group Case	176

Chapter 6 <i>Database Design 2: Design Method</i>	177
User Views	178
Information-Level Design Method	178
Represent the User View as a Collection of Tables	178
Normalize the Tables	180
Identify All Keys	180
Database Design Language (DBDL)	181
Entity-Relationship (E-R) Diagrams	182
Merge the Result into the Design	183
Database Design Examples	184
Physical-Level Design	195
Top-Down Versus Bottom-Up Design	196
Survey Form	197
Obtaining Information from Existing Documents	198
One-to-One Relationship Considerations	202
Many-to-Many Relationship Considerations	205
Nulls and Entity Subtypes	207
Avoiding Problems with Third Normal Form when Merging Tables	211
The Entity-Relationship Model	211
Summary	217
Key Terms	218
Review Questions	218
TAL Distributors Exercises	220
Colonial Adventure Tours Case	220
Solmaris Condominium Group Case	221
Chapter 7 <i>DBMS Functions</i>	223
Update and Retrieve Data	224
Provide Catalog Services	225
Support Concurrent Update	226
The Concurrent Update Problem	226
Avoiding the Lost Update Problem	230
Two-Phase Locking	231
Deadlock	234
Locking on PC-Based DBMSs	235
Timestamping	236
Recover Data	236
Journaling	236
Forward Recovery	238
Backward Recovery	239
Recovery on PC-Based DBMSs	239
Provide Security Services	240
Encryption	240
Authentication	240
Authorizations	241
Views	241
Privacy	241
Provide Data Integrity Features	242
Support Data Independence	244
Adding a Field	244
Changing the Length of a Field	244
Creating an Index	244
Adding or Changing a Relationship	244
Support Data Replication	245
Provide Utility Services	246

Summary	247
Key Terms	247
Review Questions	248
TAL Distributors Exercises	249
Colonial Adventure Tours Case	249
Solmaris Condominium Group Case	250
Chapter 8 Database Administration	253
Database Policy Formulation and Enforcement	254
Access Privileges	254
Security	256
Disaster Planning	257
Archiving	258
Other Database Administrative Functions	259
DBMS Evaluation and Selection	260
DBMS Maintenance	264
Data Dictionary Management	264
Training	265
Technical Functions	265
Database Design	265
Testing	265
Performance Tuning	266
Summary	270
Key Terms	270
Review Questions	271
TAL Distributors Exercises	271
Colonial Adventure Tours Case	272
Solmaris Condominium Group Case	272
Chapter 9 Database Management Approaches	275
Distributed Databases	275
Characteristics of Distributed DBMSs	277
Location Transparency	277
Replication Transparency	277
Fragmentation Transparency	278
Advantages of Distributed Databases	279
Disadvantages of Distributed Databases	280
Rules for Distributed Databases	283
Client/Server Systems	283
Advantages of Client/Server Systems	286
Web Access to Databases	287
XML	290
Data Warehouses	294
Data Warehouse Structure and Access	295
Rules for OLAP Systems	298
Object-Oriented DBMSs	299
What Is an Object-Oriented DBMS?	299
Objects and Classes	300
Methods and Messages	301
Inheritance	303
Unified Modeling Language (UML)	303
Rules for OODBMSs	306
Summary	307
Key Terms	308
Review Questions	309

TAL Distributors Exercises	311
Colonial Adventure Tours Case	311
Solmaris Condominium Group Case	312
Appendix A <i>Comprehensive Design Example: Marvel College</i>	313
Marvel College Requirements	313
General Description	313
Report Requirements	314
Update (Transaction) Requirements	317
Marvel College Information-Level Design	318
Final Information-Level Design	336
Exercises	337
Appendix B <i>SQL Reference</i>	345
ALTER TABLE	345
Column or Expression List (SELECT Clause)	345
Computed Fields	346
Functions	346
Conditions	346
Simple Conditions	346
Compound Conditions	346
BETWEEN Conditions	347
LIKE Conditions	347
IN Conditions	347
CREATE INDEX	347
CREATE TABLE	348
CREATE VIEW	349
Data Types	349
DELETE Rows	349
DROP INDEX	350
DROP TABLE	350
GRANT	350
INSERT	351
Integrity	351
Join	351
REVOKE	352
SELECT	352
SELECT INTO	353
Subqueries	353
UNION	354
UPDATE	354
Appendix C <i>“How Do I” Reference</i>	355
Appendix D <i>Answers to Odd-Numbered Review Questions</i>	357
Chapter 1—Introduction to Database Management	357
Chapter 2—The Relational Model 1: Introduction, QBE, and Relational Algebra	358
Chapter 3—The Relational Model 2: SQL	358
Chapter 4—The Relational Model 3: Advanced Topics	359
Chapter 5—Database Design 1: Normalization	360
Chapter 6—Database Design 2: Design Method	361
Chapter 7—DBMS Functions	362
Chapter 8—Database Administration	363
Chapter 9—Database Management Approaches	364

Appendix E	<i>Using Access to Create a Web App</i>	367
	Examining a Web App	367
	Table Restrictions for Web Apps	369
	Creating a Web App	372
	Creating a Table for a Web App	374
	Using Views in a Web App	375
	Viewing Data in a Web App	377
	Creating a New View for a Web App	379
	Creating a New Object in a Web App	379
	Running a Web App	381
	Showing a Relationship in List View	384
	Running a Web App From a Browser	384
	Creating a Summary View in a Web App	384
	Key Terms	389
	TAL Distributors Exercises	389
Appendix F	<i>A Systems Analysis Approach to Information-Level Requirements</i>	391
	Information Systems	391
	System Requirement Categories	392
	Output Requirements	392
	Input Requirements	393
	Processing Requirements	393
	Technical and Constraining Requirements	393
	Determining System Requirements	394
	Interviews	394
	Questionnaires	394
	Document Collection	394
	Observation	394
	Research	394
	Transitioning from Systems Analysis to Systems Design	394
	Key Terms	396
	Exercises	396
Glossary		397
Index		407

PREFACE

The advent of database management systems for personal computers in the 1980s moved database management beyond the realm of database professionals and into the hands of everyday users from all segments of the population. A field once limited to highly trained users of large, mainframe, database-oriented application systems became an essential productivity tool for such diverse groups as home computer users, small business owners, and end-users in large organizations.

The major PC-based database software systems have continually added features to increase their ease of use, allowing users to enjoy the benefits of database tools relatively quickly. Truly effective use of such a product, however, requires more than just knowledge of the product itself, although that knowledge is obviously important. It requires a general knowledge of the database environment, including topics such as database design, database administration, and application development using these systems. While the depth of understanding required is certainly not as great for the majority of users as it is for the information technology professional, a lack of any understanding in these areas precludes effective use of the product in all but the most limited applications.

ABOUT THIS BOOK

This book is intended for anyone who is interested in gaining some familiarity with database management. It is appropriate for students in introductory database classes in computer science or information systems programs. It is appropriate for students in database courses in related disciplines, such as business, at either the undergraduate or graduate level. Such students require a general understanding of the database environment. In addition, courses introducing students of any discipline to database management have become increasingly popular over the past few years, and this book is ideal for such courses. It is also appropriate for individuals considering purchasing a PC-based database package and who want to make effective use of such a package.

This book assumes that students have some familiarity with computers; a single introductory course is all the background that is required. While students need not have any background in programming to use this book effectively, there are certain areas where some programming experience will allow them to explore topics in more depth.

CHANGES TO THE EIGHTH EDITION

The Eighth Edition includes the following new features and content:

- Hands-on steps for creating and using Microsoft Access data macros to accomplish the same functionality as SQL triggers.
- General information about creating Web apps to allow data to be shared easily using the Web.
- A discussion of the systems analysis approach for determining the requirements needed as the starting point for database design, including descriptions of the requirements you need to gather and how to gather these requirements.
- A new case for TAL Distributors is used to illustrate the concepts in each chapter of the book, and is also used in the end-of-chapter exercises.
- Two new cases for Colonial Adventure Tours and Solmaris Condominium Group are used in the end-of-chapter cases.
- Critical-thinking questions and exercises that reinforce problem-solving and analytical skills are included in each chapter.

SPECIAL FEATURES

As in the Seventh Edition, the SQL material is covered using Access. Also included are generic forms of all examples that students can use on a variety of platforms, including Oracle. The Eighth Edition continues the two appendices that provide a useful reference for anyone wanting to use SQL effectively. Appendix B includes a command reference of all the SQL commands and operators that are taught in the chapters. Students can use this appendix as a quick resource when constructing commands. Each command includes a short description, a table that shows the required and optional clauses and operators, and an example and its results. Appendix C provides students with an opportunity to ask a question, such as “How do I delete rows?,” and to identify the appropriate section in Appendix B to use to find the answer. Appendix C is extremely valuable when students know what they want to accomplish, but can’t remember the exact SQL command they need.

In addition to the section of Review Questions, the end of each chapter includes three sets of exercises—one featuring the TAL Distributors database and the others featuring the Colonial Adventure Tours database and the Solmaris Condominium Group database—that give students “hands-on” experiences with the concepts found in the chapter.

As in the previous edition, the Eighth Edition covers entity-relationship diagrams. The database design material includes a discussion of the entity-relationship model as a database model. It also includes a discussion of a characterization of various types of primary keys.

The TAL Distributors, Colonial Adventure Tours, and Solmaris Condominium Group databases will be available at www.cengage.com and are usable with Access 2010 and Access 2013. For those students using database management systems that run scripts (such as Oracle), the data files also include the script files that create the tables and add the data to the tables in the databases used in the book.

For instructors who want to use an Access or SQL text as a companion to the Eighth Edition, the Instructor’s Manual for this book includes detailed tips on integrating the Eighth Edition with other books from Cengage Learning that cover Access 2010 and Access 2013 and SQL (for more information, see the “Teaching Tools” section in this preface).

Detailed Coverage of the Relational Model, including Query-By-Example (QBE) and SQL

The book features detailed coverage of the important aspects of the relational model, including comprehensive coverage of SQL. It also covers QBE and relational algebra as well as advanced aspects of the model, such as views, the use of indexes, the catalog, and relational integrity rules.

Normalization Coverage

The Eighth Edition covers first normal form, second normal form, third normal form (Boyce-Codd normal form), and fourth normal form. The book describes in detail the update anomalies associated with lower normal forms as part of the motivation for the need for higher normal forms. Finally, the book examines correct and incorrect ways to normalize tables. This book specifically addresses this by showing students some of the mistakes people can make in the normalization process, explaining why the approach is incorrect, demonstrating the problems that would result from incorrect normalizations, and, most importantly, identifying how to avoid these mistakes.

Views Coverage

This text covers the important topic of views. It discusses the creation and use of views as well as the advantages of using views.

Database Design

The important process of database design is given detailed treatment. A highly useful method for designing databases is presented and illustrated through a variety of examples. In addition to the method, this text includes important design topics such as the use of survey forms, obtaining information by reviewing existing documents, special relationship considerations, and entity subtypes. Appendix A contains a comprehensive design example that illustrates how to apply the complete design process to a large and complex set of requirements. After mastering the design method presented in this text,

students should be able to produce correct database designs for future database requirements they encounter.

Functions Provided by a Database Management System

With such a wide range of features included in current database management systems, it is important for students to know the functions that such systems should provide. These functions are presented and discussed in detail.

Database Administration

While database administration (DBA) is absolutely essential in the mainframe environment, it is also important in a personal computer environment, especially when the database is shared among several users. Thus, this text includes a detailed discussion of the database administration function.

Database Management System Selection

The process of selecting a database management system is important, considering the number of available systems from which to choose. Unfortunately, selecting the correct database management system is not an easy task. To prepare students to be able to do an effective job in this area, the text includes a detailed discussion of the selection process together with a comprehensive checklist that greatly assists in making such a selection.

Advanced Topics

The text also covers distributed database management systems, client/server systems, data warehouses, object-oriented database management systems, Web access to databases, and XML. Each of these topics encompasses an enormous amount of complex information, but the goal is to introduce students to these important topics. The text also includes detailed coverage of data macros in Access. The book includes detailed coverage of Web apps. In addition, the book discusses the systems analysis approach to determining the requirements needed as the starting point for database design. After describing information systems, we describe the requirements you need to gather and how to gather these requirements.

Numerous Realistic Examples

The book contains numerous examples illustrating each of the concepts. A running “case” example—TAL Distributors—is used throughout the book to illustrate concepts. The examples are realistic and represent the kinds of problems students will encounter in the design, manipulation, and administration of databases. Exercises that use the TAL Distributors case are included at the end of each chapter. In addition, there is another complete set of exercises at the end of each chapter that feature a second and third case—Colonial Adventure Tours and Solmaris Condominium Group—giving students a chance to apply what they have learned to a database that they have not seen in the chapter material.

Review Material

This text contains a wide variety of questions. At key points within the chapters, students are asked questions to reinforce their understanding of the material before proceeding. The answers to these questions follow the questions. A summary and a list of key terms appear at the end of each chapter, followed by review questions that test the students’ knowledge of the important points in the chapter and that occasionally test their ability to apply what they have learned. The answers to the odd-numbered review questions are provided in Appendix D. Each chapter also contains hands-on exercises related to the TAL Distributors, Colonial Adventure Tours, and Solmaris Condominium Group case examples. Critical-thinking questions that reinforce problem-solving and analytical skills are included for review questions and hands-on exercises.

Teaching Tools

When this book is used in an academic setting, instructors may obtain the following teaching tools from Cengage Learning through their sales representative or by visiting www.cengage.com:

- **Instructor's Manual** The Instructor's Manual has been carefully prepared and tested to ensure its accuracy and dependability. The Instructor's Manual includes suggestions and strategies for using this text, including the incorporation of companion texts on Access or SQL for those instructors who desire to do so. For instructors who want to use an Access or SQL text as a companion to the Eighth Edition, the Instructor's Manual for this book includes detailed tips on integrating the Eighth Edition with the following books, also published by Cengage Learning: *Microsoft Access 2013: Introductory Concepts and Techniques*, *Microsoft Access 2013: Complete Concepts and Techniques*, and *Microsoft Access 2013: Comprehensive Concepts and Techniques*, by Pratt and Last; and *A Guide to SQL, Eighth Edition*, by Pratt and Last.
- **Data and Solution Files** Data and solution files are available at www.cengage.com. Data files consist of copies of the TAL Distributors, Colonial Adventure Tours, and Solmaris Condominium Group databases that are usable in Access 2010 and Access 2013, and script files to create the tables and data in these databases in other systems, such as Oracle.

Cengage Learning Testing Powered by Cognero is a flexible, online system that allows you to:

- author, edit, and manage test bank content from multiple Cengage Learning solutions
- create multiple test versions in an instant
- deliver tests from your LMS, your classroom or wherever you want
- **PowerPoint Presentations** Microsoft PowerPoint slides are included for each chapter as a teaching aid for classroom presentations, to make available to students on a network for chapter review, or to be printed for classroom distribution. Instructors can add their own slides for additional topics they introduce to the class. The presentations are available at www.cengage.com.
- **Figure Files** Figure files are included so that instructors can create their own presentations using figures appearing in the text.

ORGANIZATION OF THE TEXTBOOK

This text includes nine chapters covering general database topics that are relevant to any database management system. A brief description of the organization of topics in the chapters and an overview of each chapter's contents follows.

Introduction

Chapter 1 provides a general introduction to the field of database management.

The Relational Model

The relational model is covered in detail in Chapters 2, 3, and 4. Chapter 2 covers the data definition and manipulation aspects of the model using QBE and relational algebra. The text uses Access 2013 to illustrate the QBE material. The relational algebra section includes the entire relational algebra. (*Note:* The extra material on relational algebra is optional and can be omitted if desired.)

Chapter 3 is devoted exclusively to SQL. The SQL material is illustrated using Access, but the chapter also includes generic versions of all examples that can be used with a variety of platforms, including Oracle.

Chapter 4 covers some advanced aspects of the relational model such as views, the use of indexes, the catalog, relational integrity rules, stored procedures, triggers, and data macros.

Database Design

Chapters 5 and 6 are devoted to database design. Chapter 5 covers the normalization process, which enables students to identify and correct bad designs. This chapter discusses and illustrates the use of

first, second, third, and fourth normal forms. (*Note:* The material on fourth normal form is optional and can be omitted if desired.)

Chapter 6 presents a method for database design using many examples. The material includes entity-relationship diagrams and their role in database design. It also includes discussions of several special design issues as well as the use of survey forms, obtaining information by reviewing existing documents, special relationship considerations, and entity subtypes. After completing Chapter 6, students can further challenge themselves by completing Appendix A, which includes a comprehensive design example that illustrates the application of the complete design process to a large and complex set of requirements. (*Note:* Chapters 5 and 6 can be covered immediately after Chapter 2 if desired.)

Database Management System Functions

Chapter 7 discusses the features that should be provided by a full-functioned PC-based database management system. This chapter includes coverage of journaling, forward recovery, backward recovery, authentication, and authorizations.

Database Administration

Chapter 8 is devoted to the role of database administration. Also included in this chapter is a discussion of the process of selecting a database management system.

Database Management Approaches

Chapter 9 provides an overview of several advanced topics: distributed databases, client/server systems, Web access to databases, XML and related document specification standards, data warehouses, and object-oriented databases.

GENERAL NOTES TO THE STUDENT

There are many places in the text where special questions have been embedded. Sometimes the purpose of these questions is to ensure that you understand some crucial material before you proceed. In other cases, the questions are designed to give you the chance to consider some special concept in advance of its actual presentation. In all cases, the answers to these questions follow each question. You could simply read the question and its answer. You will receive maximum benefit from the text, however, if you take the time to work out the answers to the questions and then check your answer against the one provided before continuing.

The end-of-chapter material consists of a summary, a list of key terms, review questions, and exercises for the TAL Distributors, Colonial Adventure Tours, and Solmaris Condominium Group databases. The summary briefly describes the material covered in the chapter. The review questions require you to recall and apply the important material in the chapter. (The answers to the odd-numbered review questions appear in Appendix D.) The TAL Distributors, Colonial Adventure Tours, and Solmaris Condominium Group exercises test your knowledge of the chapter material; your instructor will assign one or more of these exercises for you to complete. Review questions and exercises include critical-thinking questions to challenge your problem-solving and analytical skills.

ACKNOWLEDGMENTS

We would like to acknowledge the following individuals who all made contributions during the preparation of this book during its multiple editions. We also appreciate the efforts of the following individuals, who have been invaluable during this book's development: Clara Goosman, Product Manager; Jennifer Feltri, Content Project Manager; Nicole Ashton, Quality Assurance tester; Arul Joseph Raj, Senior Project Manager; and Wendy Langerud, Content Developer.

We have again been privileged to work with Jessica Evans as our Development Editor. Thank you, Jess, for your knowledge, skill, guidance, and energy, and for your hard work and positive influence on our work. We appreciate your dedication and many contributions to this book, value you as a friend, and wish the very best to you and your family. We have said it many times before, but we need to say it again. You are the best!!

CHAPTER 1

INTRODUCTION TO DATABASE MANAGEMENT

LEARNING OBJECTIVES

- Introduce TAL Distributors, the company that is used as the basis for many of the examples throughout the text
- Introduce basic database terminology
- Describe database management systems (DBMSs)
- Explain the advantages and disadvantages of database processing
- Introduce Colonial Adventure Tours, the company that is used in a case that appears throughout the text
- Introduce Solmaris Condominium Group, the company that is used in another case that appears throughout the text

INTRODUCTION

In this chapter, you will examine the requirements of TAL Distributors, a company that will be used in many examples in this chapter and in the rest of the text. You will learn how TAL Distributors initially stored its data, what problems employees encountered with the storage method, and why management decided to use a database management system. Then you will study the basic terminology and concepts of databases and database management systems, and learn the advantages and disadvantages of database processing. Finally, you will examine the database requirements for Colonial Adventure Tours and Solmaris Condominium Group, the companies featured in the cases that appear at the end of each chapter.

TAL DISTRIBUTORS BACKGROUND

TAL Distributors is a wholesaler of finely crafted wooden toys, games, and puzzles. Since its inception, the company has used spreadsheet software to maintain data about customers, orders, inventory, and sales representatives. Management has determined that the company's recent growth means it is no longer feasible to use spreadsheets to maintain its data.

What has led the managers at TAL Distributors to this decision? One of the company's spreadsheets, shown in Figure 1-1, displays sample order data and illustrates the company's problems with the spreadsheet approach. For each order, the spreadsheet displays the number and name of the customer placing the order; the number and date of the order; the number, description, number ordered, quoted price, and storehouse number of the item ordered; and the number of the sales rep assigned to the customer. Note that The Everything Shop order (order number 51610) and the first Almondton General Store order (order number 51617) appear in two rows because these customers purchased two different items in their orders. In the second Almondton General Store order (order number 51623), the customer purchased three different items.

Orders requiring more than one spreadsheet row

Customer Number	Customer Name	Order Number	Order Date	Item Number	Description	Number Ordered	Quoted Price	Store house	Rep Number
126	Toys Galore	51608	10/12/2015	CD33	Wood Block Set (48 piece)	5	\$86.99	1	15
126	Toys Galore	51619	10/15/2015	FD11	Rocking Horse	2	\$121.95	3	15
260	Brookings Direct	51614	10/13/2015	FD11	Rocking Horse	1	\$124.95	3	30
334	The Everything Shop	51610	10/12/2015	KL78	Pick Up Sticks	25	\$10.95	1	45
334	The Everything Shop	51610	10/12/2015	TR40	Tic Tac Toe	10	\$13.99	2	45
386	Johnson's Department Store	51613	10/13/2015	DL51	Classic Railway Set	5	\$104.95	3	30
586	Almondton General Store	51617	10/15/2015	NL89	Wood Block Set (62 piece)	4	\$115.99	3	45
586	Almondton General Store	51617	10/15/2015	TW35	Fire Engine	3	\$116.95	2	45
586	Almondton General Store	51623	10/15/2015	DR67	Giant Star Brain Teaser	5	\$29.95	2	45
586	Almondton General Store	51623	10/15/2015	FH24	Puzzle Gift Set	12	\$36.95	1	45
586	Almondton General Store	51623	10/15/2015	KD34	Pentominos Brain Teaser	10	\$13.10	2	45
796	Unique Gifts	51625	10/16/2015	MT03	Zauberkasten Brain Teaser	8	\$45.79	1	45

FIGURE 1-1 Sample orders spreadsheet

Redundancy is one problem that employees have with the orders spreadsheet. **Redundancy** is the duplication of data or the storing of the same data in more than one place. In the orders spreadsheet, redundancy occurs in the Customer Name column because the name of a customer is stored in more than one place. Both rows for customer number 334, for example, store “The Everything Shop” as the customer name. In the orders spreadsheet, redundancy also occurs in other columns, such as the Order Date and Description columns.

Q & A

Question: What problems does redundancy cause?

Answer: Redundancy wastes space because you're storing the same data in multiple places. This extra space results in larger spreadsheets that require more space in memory and on disk and that take longer to save and open.

When you need to change data, redundancy also makes your changes more cumbersome and time-consuming. For example, if you incorrectly enter “The Everything Shop” in the Customer Name column, you would need to correct it in two places. Even if you use the global find-and-replace feature, multiple changes require more computer time than does a single change.

Finally, redundancy can lead to inconsistencies. For example, you might enter “The Every Thing Shop” and “The Everything Shop” on separate rows in the Customer Name column, and then be unsure about the correct version of this customer's name. Further, if this customer's name is spelled two different ways and you use the search feature with one of the two values, you'd find a single match instead of two matches.

Difficulty accessing related data is another problem that employees at TAL Distributors encounter with their spreadsheets. For example, if you want to see a customer's address and an item's standard price, you must open and search other spreadsheets that contain this data.

Spreadsheets also have limited security features to protect data from being accessed by unauthorized users. A spreadsheet's data-sharing features also prevent multiple employees from updating data in one spreadsheet at the same time. Finally, if the increase in sales and growth at TAL Distributors continues at its planned rate, spreadsheets have inherent size limitations that will eventually force the company to split its order data into multiple spreadsheets. Splitting the spreadsheets would create further redundancy.

Having decided to replace its spreadsheet software, management has determined that TAL Distributors must maintain the following information about its sales reps, customers, and items inventory:

- The sales rep number, last name, first name, address, total commission, and commission rate for each sales rep.
- The customer number, name, address, current balance, and credit limit for each customer, as well as the number of the sales rep who represents the customer.
- The item number, description, number of units on hand, item category, number of the store-house where the item is stored, and unit price for each item in inventory.

TAL Distributors must also store information about orders. Figure 1-2 shows a sample order.

TAL Distributors		INVOICE		
555 Main Street San Rita, TX 78364 Phone (512) 555-0190 Fax (512) 555-0191		ORDER: 51610 DATE: October 12, 2015		
TO: CUSTOMER: 334 The Everything Shop 342 Magee St. Congaree, CA 90097		SALES REP: 45 Hui Tian		
ITEM NUMBER	DESCRIPTION	NUMBER ORDERED	PRICE	TOTAL
KL78	Pick Up Sticks	25	10.95	273.75
TR40	Tic Tac Toe	10	13.99	139.90
			Extensions	
			TOTAL	413.65

Make all checks payable to TAL Distributors
Total due in 15 days. Overdue accounts subject to a service charge of 1% per month.

Thank you for your business!

FIGURE 1-2 Sample order

- The heading (top) of the order contains the TAL Distributors' name and address, the word "Invoice," the order number and date; the customer's number, name, and address; and the sales rep's number and name.
- The body of the order contains one or more order lines, sometimes called line items. Each order line contains an item number, a description, the number of units of the item ordered, and the quoted price for the item. Each order line also contains a total, usually called an extension, which is the result of multiplying the number ordered by the quoted price.
- The footing (bottom) of the order contains the order total.

TAL Distributors must also store the following items for each customer's order:

- For each order, the company must store the order number, the date the order was placed, and the number of the customer that placed the order. The customer's name and address and the number of the sales rep who represents the customer are stored with the customer information. The name of the sales rep is stored with the sales rep information.
- For each order line, the company must store the order number, the item number, the number of units ordered, and the quoted price. Remember that the item description is stored with the information about items. The result of multiplying the number of units ordered by the quoted price is not stored because the computer can calculate it when necessary.
- The overall order total is not stored. Instead, the computer calculates the total whenever an order is printed or displayed on the screen.

The problem facing TAL Distributors is common to many businesses and individuals that need to store and retrieve data in an efficient and organized way. Furthermore, most organizations are interested in more than one category of information. For example, TAL Distributors is interested in categories such as sales reps, customers, orders, and items. A school is interested in students, faculty, and classes; a real estate agency is interested in clients, houses, and agents; and a car dealership is interested in customers, vehicles, and manufacturers.

Besides wanting to store data that pertains to more than one category, TAL Distributors is also interested in the relationships between the categories. For example, TAL Distributors employees want to be able to associate orders with the customers that ordered them, the sales reps who coordinated the orders, and the items that the customers requested. Likewise, a real estate agency wants to know not only about clients, houses, and agents but also about the relationships between clients and houses (which clients have listed which houses and which clients have expressed interest in which houses). A real estate agency also wants to know about the relationships between agents and houses (which agent sold which house, which agent is listing which house, and which agents are receiving commissions for which houses).

DATABASE BACKGROUND

After studying the alternatives to using spreadsheet software, TAL Distributors decided to switch to a database system. A database is a structure that contains information about many different categories of information and about the relationships between those categories. The TAL Distributors database, for example, will contain information about sales reps, customers, orders, and items. It will also provide facts that relate sales reps to the customers they represent and customers to the orders they currently have placed.

With a database, employees can enter the number of a particular order and identify which customer placed the order, as well as which items the customer ordered. Alternately, employees can start with a customer and find all orders the customer placed, together with which items the customer ordered and the amount of the commission earned by the customer's sales rep. Using a database, not only can TAL Distributors maintain its data better, but it also can use the data in the database to produce a variety of reports and to answer different types of questions.

There are some terms and concepts in the database environment that are important for you to know. For instance, the terms *entity*, *attribute*, and *relationship* are fundamental when discussing databases. An **entity** is a person, place, object, event, or idea for which you want to store and process data. The entities of interest to TAL Distributors, for example, are sales reps, customers, orders, and items.

An **attribute** is a characteristic or property of an entity. The term is used in this text exactly as it is used in everyday English. For the entity *person*, for example, the list of attributes might include such things as eye color and height. For TAL Distributors, the attributes of interest for the entity *customer* are such things as customer name, street, city, and so on. An attribute is also called a **field** or **column** in many database systems.

Figure 1-3 shows two entities, Rep (short for Sales Rep) and Customer, along with the attributes for each entity. The Rep entity has nine attributes: RepNum, LastName, FirstName, Street, City, State, PostalCode, Commission, and Rate. The attributes are the same as the columns in a spreadsheet. The Customer entity has nine attributes: CustomerNum, CustomerName, Street, City, State, PostalCode, Balance, CreditLimit, and RepNum.

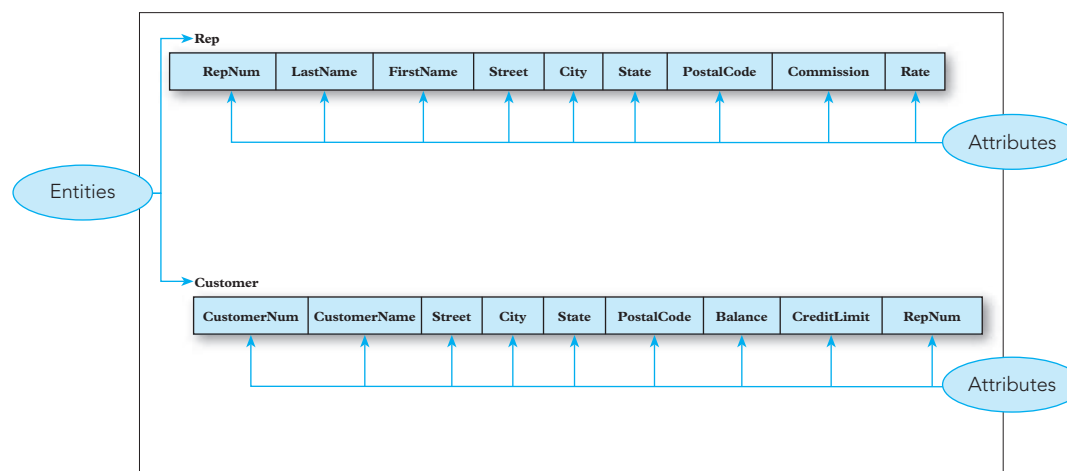


FIGURE 1-3 Entities and attributes

The final key database term is relationship. A **relationship** is an association between entities. There is an association between reps and customers, for example, at TAL Distributors. A rep is associated with all of his or her customers, and a customer is associated with its rep. Technically, you say that a rep is *related to* all of his or her customers, and a customer is *related to* its rep.

This particular relationship is called a **one-to-many relationship** because each rep is associated with *many* customers, but each customer is associated with only *one* rep. In this type of relationship, the word *many* is used differently than in everyday English; it might not always indicate a large number. In this context, for example, the term *many* means that a rep can be associated with *any* number of customers. That is, a given rep can be associated with zero, one, or more customers.

A one-to-many relationship often is represented visually in the manner shown in Figure 1-4. In such a diagram, entities and attributes are represented in precisely the same way as they are shown in Figure 1-3. A line connecting the entities represents the relationship. The *one* part of the relationship (in this case, Rep)

does not have an arrow on its end of the line, and the *many* part of the relationship (in this case, Customer) is indicated by a single-headed arrow.

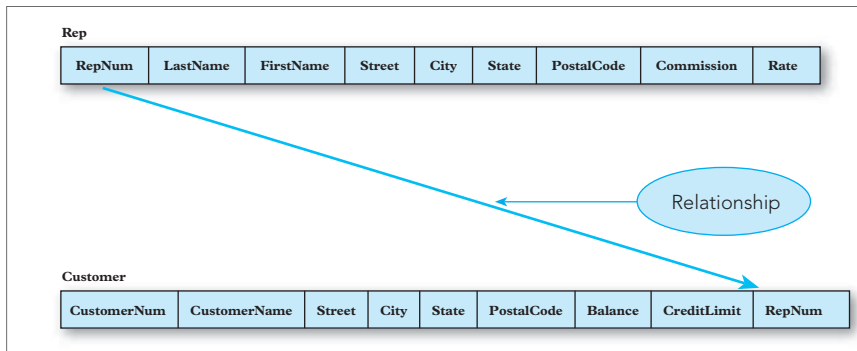


FIGURE 1-4 One-to-many relationship

Spreadsheets, word-processed documents, Web pages, and other computer information sources are stored in files. Basically, a file that is used to store data, which is often called a **data file**, is the computer counterpart to an ordinary paper file you might keep in a file cabinet, an accounting ledger, or other place. A database, however, is more than a file. Unlike a typical data file, a database can store information about multiple entities.

A database also holds information about the relationships among the various entities. Not only will the TAL Distributors database have information about both reps and customers, it will also hold information relating reps to the customers they service, customers to orders, items to orders, and so on. Formally, a **database** is a structure that can store information about multiple types of entities, the attributes of those entities, and the relationships between the entities.

How does a database handle entities, attributes of entities, and relationships between entities? Entities and attributes are fairly simple. Each entity has its own table. In the TAL Distributors database, for example, there will be one table for reps, one table for customers, and so on. The attributes of an entity become the columns in the table. In the table for reps, for example, there will be a column for the rep number, a column for the rep last name, and so on.

What about relationships between entities? At TAL Distributors, there is a one-to-many relationship between reps and customers. (Each rep is related to the many customers that he or she represents, and each customer is related to the one rep who represents the customer.) How is this relationship handled in a database system? It is handled by using common columns in the two tables. Consider Figure 1-4 again. The RepNum column in the Rep table and the RepNum column in the Customer table are used to implement the relationship between reps and customers. Given a rep, you can use these columns to determine all the customers that he or she represents; given a customer, you can use these columns to find the rep who represents the customer.

How will TAL Distributors store its data in a database? Figure 1-5 shows sample data for TAL Distributors.

Rep

RepNum	LastName	FirstName	Street	City	State	PostalCode	Commission	Rate
15	Campos	Rafael	724 Vinca Dr.	Grove	CA	90092	\$23,457.50	0.06
30	Gradey	Megan	632 Liatris St.	Fullton	CA	90085	\$41,317.00	0.08
45	Tian	Hui	1785 Tyler Ave.	Northfield	CA	90098	\$27,789.25	0.06
60	Sefton	Janet	267 Oakley St.	Congaree	CA	90097	\$0.00	0.06

Customer

CustomerNum	CustomerName	Street	City	State	PostalCode	Balance	CreditLimit	RepNum
126	Toys Galore	28 Laketon St.	Fullton	CA	90085	\$1,210.25	\$7,500.00	15
260	Brookings Direct	452 Columbus Dr.	Grove	CA	90092	\$575.00	\$10,000.00	30
334	The Everything Shop	342 Magee St.	Congaree	CA	90097	\$2,345.75	\$7,500.00	45
386	Johnson's Department Store	124 Main St.	Northfield	CA	90098	\$879.25	\$7,500.00	30
440	Grove Historical Museum Store	3456 Central Ave.	Fullton	CA	90085	\$345.00	\$5,000.00	45
502	Cards and More	167 Hale St.	Mesa	CA	90104	\$5,025.75	\$5,000.00	15
586	Almondton General Store	3345 Devon Ave.	Almondton	CA	90125	\$3,456.75	\$15,000.00	45
665	Cricket Gift Shop	372 Oxford St.	Grove	CA	90092	\$678.90	\$7,500.00	30
713	Cress Store	12 Rising Sun Ave.	Congaree	CA	90097	\$4,234.60	\$10,000.00	15
796	Unique Gifts	786 Passmore St.	Northfield	CA	90098	\$124.75	\$7,500.00	45
824	Kline's	945 Gilham St.	Mesa	CA	90104	\$2,475.99	\$15,000.00	30
893	All Season Gifts	382 Wildwood Ave.	Fullton	CA	90085	\$935.75	\$7,500.00	15

Orders

OrderNum	OrderDate	CustomerNum
51608	10/12/2015	126
51610	10/12/2015	334
51613	10/13/2015	386
51614	10/13/2015	260
51617	10/15/2015	586
51619	10/15/2015	126
51623	10/15/2015	586
51625	10/16/2015	796

OrderLine

OrderNum	ItemNum	NumOrdered	QuotedPrice
51608	CD33	5	\$86.99
51610	KL78	25	\$10.95
51610	TR40	10	\$13.99
51613	DL51	5	\$104.95
51614	FD11	1	\$124.95
51617	NL89	4	\$115.99
51617	TW35	3	\$116.95
51619	FD11	2	\$121.95
51623	DR67	5	\$29.95
51623	FH24	12	\$36.95
51623	KD34	10	\$13.10
51625	MT03	8	\$45.79

Item

ItemNum	Description	OnHand	Category	Storehouse	Price
AH74	Patience	9	GME	3	\$22.99
BR23	Skittles	21	GME	2	\$29.99
CD33	Wood Block Set (48 piece)	36	TOY	1	\$89.49
DL51	Classic Railway Set	12	TOY	3	\$107.95
DR67	Giant Star Brain Teaser	24	PZL	2	\$31.95
DW23	Mancala	40	GME	3	\$50.00
FD11	Rocking Horse	8	TOY	3	\$124.95
FH24	Puzzle Gift Set	65	PZL	1	\$38.95
KA12	Cribbage Set	56	GME	3	\$75.00
KD34	Pentominoes Brain Teaser	60	PZL	2	\$14.95
KL78	Pick Up Sticks	110	GME	1	\$10.95
MT03	Zauberkasten Brain Teaser	45	PZL	1	\$45.79
NL89	Wood Block Set (62 piece)	32	TOY	3	\$119.75
TR40	Tic Tac Toe	75	GME	2	\$13.99
TW35	Fire Engine	30	TOY	2	\$118.95

FIGURE 1-5 Sample data for TAL Distributors

In the Rep table, you see that there are four reps whose numbers are 15, 30, 45, and 60. The name of sales rep 15 is Rafael Campos. His street address is 724 Vinca Dr. He lives in Grove, CA, and his postal code is 90092. His total commission is \$23,457.50, and his commission rate is 0.06 (6%).

TAL Distributors has 12 customers, which are identified with the numbers 126, 260, 334, 386, 440, 502, 586, 665, 713, 796, 824, and 893. The name of customer number 126 is Toys Galore. This customer's address is 28 Laketon St. in Fullton, CA, with a postal code of 90085. The customer's current balance is \$1,210.25, and its credit limit is \$7,500.00. The number 15 in the RepNum column indicates that Toys Galore is represented by sales rep 15 (Rafael Campos).

Skipping to the table named Item, you see that there are 15 items, whose item numbers are AH74, BR23, CD33, DL51, DR67, DW23, FD11, FH24, KA12, KD34, KL78, MT03, NL89, TR40, and TW35. Item AH74 is Patience, and TAL Distributors has nine units of this item on hand. The Patience item is in the GME (games) category and is located in storehouse 3. The price of the Patience game is \$22.99. Other categories are PZL (puzzles) and TOY (toys).

Moving back to the table named Orders, you see that there are eight orders, which are identified with the numbers 51608, 51610, 51613, 51614, 51617, 51619, 51623, and 51625. Order number 51608 was placed on October 12, 2015, by customer 126 (Toys Galore).

NOTE

In some database systems, the word "Order" has a special purpose. Having a table named Order could cause problems in such systems. For this reason, TAL Distributors uses the table name Orders instead of Order.

The table named OrderLine might seem strange at first glance. Why do you need a separate table for the order lines? Couldn't the order lines be included in the Orders table? The answer is yes. The Orders table could be structured as shown in Figure 1-6. Notice that this table contains the same orders as those shown in Figure 1-5, with the same dates and customers. In addition, each table row in Figure 1-6 contains all the order lines for a given order. Examining the second row, for example, you see that order 51610 has two order lines. One of the order lines is for 25 units of item KL78 at \$10.95 each, and the other order line is for 10 units of item TR40 at \$13.99 each.

Orders

OrderNum	OrderDate	CustomerNum	ItemNum	NumOrdered	QuotedPrice
51608	10/12/2015	126	CD33	5	\$86.99
51610	10/12/2015	334	KL78	25	\$10.95
			TR40	10	\$13.99
51613	10/13/2015	386	DL51	5	\$104.95
51614	10/13/2015	260	FD11	1	\$124.95
51617	10/15/2015	586	NL89	4	\$115.99
			TW35	3	\$116.95
51619	10/15/2015	126	FD11	2	\$121.95
51623	10/15/2015	586	DR67	5	\$29.95
			FH24	12	\$36.95
			KD34	10	\$13.10
51625	10/16/2015	796	MT03	8	\$45.79

FIGURE 1-6 Alternative Orders table structure

Q & A

Question: How is the information in Figure 1-5 represented in Figure 1-6?

Answer: Examine the OrderLine table shown in Figure 1-5 and note the second and third rows. The second row indicates that there is an order line in order 51610 for 25 units of item KL78 at \$10.95 each. The third row indicates that there is an order line in order 51610 for ten units of item TR40 at \$13.99 each. Thus, the information in Figure 1-6 is represented in Figure 1-5 in two separate rows rather than in one row.

It might seem inefficient to use two rows to store information that can be represented in one row. There is a problem, however, with the arrangement shown in Figure 1-6—the table is more complicated. In Figure 1-5, there is a single entry at each position in the table. In Figure 1-6, some of the individual positions within the table contain multiple entries, thus making it difficult to track the information between columns. In the row for order number 51610, for example, it is crucial to know that the item number KL78 corresponds to the number 25 in the NumOrdered column (not to the 10) and that it corresponds to \$10.95 in the QuotedPrice column (not to \$13.99). In addition, having a more complex table means that there are practical issues to worry about, such as:

- How much room do you allow for these multiple entries?
- What happens when an order requires more order lines than you have allowed room for?
- Given an item, how do you determine which orders contain order lines for that item?

Certainly, none of these problems is unsolvable. These problems do add a level of complexity, however, that is not present in the arrangement shown in Figure 1-5. In Figure 1-5, there are no multiple entries to worry about, it doesn't matter how many order lines exist for any order, and it is easy to find every order that contains an order line for a given item (just look for all order lines with the given item number in the ItemNum column). In general, this simpler structure is preferable, which is why order lines appear in a separate table.

To test your understanding of the TAL Distributors data, use the data shown in Figure 1-5 to answer the following questions.

Q & A

Question: What are the numbers of the customers represented by Rafael Campos?

Answer: 126, 502, 713, and 893. (Look up the RepNum value for Rafael Campos in the Rep table and obtain the number 15. Then find all customers in the Customer table that have the number 15 in the RepNum column.)

Q & A

Question: What is the name of the customer that placed order 51613, and what is the name of the rep who represents this customer?

Answer: Johnson's Department Store is the customer, and Megan Gradey is the rep. (Look up the CustomerNum value in the Orders table for order number 51613 and obtain the number 386. Then, find the customer in the Customer table with a CustomerNum value of 386. Using this customer's RepNum value, which is 30, find the name of the rep in the Rep table.)

Q & A

Question: List all the items that appear in order 51617. For each item, give the description, number ordered, and quoted price.

Answer: Item number: NL89, description: Wood Block Set (62 piece), number ordered: 4, and quoted price: \$115.99. Also, item number: TW35, description: Fire Engine, number ordered: 3, and quoted price: \$116.95. (Look up each OrderLine table row in which the order number is 51617. Each row contains an item number, the number ordered, and the quoted price. Use the item number to look up the corresponding description in the Item table.)