9TH EDITION

# CONCEPTS OF DATABASE MANAGEMENT

STARKS • PRATT • LAST

# CONCEPTS OF DATABASE MANAGEMENT

# CONCEPTS OF DATABASE MANAGEMENT

Ninth Edition

**Joy L. Starks**
*Indiana University—Purdue University Indianapolis*
**Philip J. Pratt**
*Grand Valley State University*
**Mary Z. Last**

CENGAGE

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

**Concepts of Database Management, Ninth Edition**
**Joy L. Starks, Philip J. Pratt, and Mary Z. Last**

SVP, GM Skills & Global Product Management: Jonathan Lau

Product Team Manager: Kristin McNary

Associate Product Manager: Kate Mason

Senior Content Development Manager: Leigh Hefferon

Content Developer: Maria Gargulio and Tyler Sally

Marketing Director: Michele McTighe

Marketing Manager: Stephanie Albracht

Production Director: Patty Stephan

Content Project Manager: Michele Stulga

Art Director: Diana Graham

Cover Designer: Roycroft Design (roycroftdesign.com)

Production Service/Composition: Lumina Datamatics, Inc.



Library of Congress Control Number: 2017963668

ISBN: 978-1-337-09342-2

**Cengage**
20 Channel Center Street
Boston, MA 02210
USA



Printed in the United States of America
Print Number: 01 Print Year: 2018

# TABLE OF CONTENTS

# 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 also is 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 NINTH EDITION

The Ninth Edition includes the following new features and content:

- New "Your Turn" exercises to fully engage students in critical thinking about what they have just learned.
- Full color screen shots using Access 2016.
- 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 BITS Corporation is used to illustrate the concepts in each chapter of the book, and is also used in the end-of-chapter exercises.
- A new case for Sports Physical Therapy, along with a case for Colonial Adventure Tours, 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.
- Concepts of big data are presented across many chapter topics.
- A new appendix covering the use of MySQL with the database cases.

## SPECIAL FEATURES

As in the Eighth 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 Ninth 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 cannot remember the exact SQL command they need.

A new Appendix D introduces MySQL with instructions for downloading and installing both the server and the MySQL Workbench user interface. Students learn how to connect to the server, open and manipulate an SQL file, enter and save SQL scripts, and use the command line.

In addition to the section of Review Questions, the end of each chapter includes three sets of exercises—one featuring the BITS Corporation database and the others featuring the Colonial Adventure Tours database and the Sports Physical Therapy database—that give students "hands-on" experiences with the concepts found in the chapter.

As in the previous edition, the Ninth 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 BITS Corporation, Colonial Adventure Tours, and Sports Physical Therapy databases will be available at *www.cengagebrain.com* and are usable with Access 2010, Access 2013, and Access 2016. 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 Ninth Edition, the Instructor's Manual for this book includes detailed tips on integrating the Ninth Edition with other books from Cengage Learning that cover Access 2010, Access 2013, Access 2016, 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 Ninth 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 describes the process of beginning from a user perspective and then 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, with examples both in Access and SQL.

## Database Administration

While database administration (DBA) is absolutely essential in the mainframe environment, it also is 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 coverage of data macros in Access. In addition, the book presents the systems analysis approach to determining the requirements needed as the starting point for database design. After describing information systems, the book describes 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—BITS Corporation—is used throughout the book to demonstrate concepts. The examples are realistic and represent the kinds of real-world problems students will encounter in the design, manipulation, and administration of databases. Exercises that use the BITS Corporation 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 features a second and third case—Colonial Adventure Tours and Sports Physical Therapy—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. Each chapter also contains hands-on exercises related to the BITS Corporation, Colonial Adventure Tours, and Sports Physical Therapy 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 Ninth Edition, the Instructor's Manual for this book includes detailed tips on integrating the Ninth Edition with the following books, also published by Cengage Learning: *Microsoft Access 2013: Introductory Concepts and Techniques, Microsoft Access 2016: Complete Concepts and Techniques,* and *Microsoft Access 2016: Comprehensive Concepts and Techniques,* 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 BITS Corporation, Colonial Adventure Tours, and Sports Physical Therapy databases that are usable in Access 2010, Access 2013, and Access 2016, and script files to create the tables and data in these databases in other systems, such as Oracle and MySQL.

**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.cengagebrain.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 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 2016 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 and MySQL.

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, and Appendix E, A Systems Analysis Approach to Information-level 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.

You also will find *Your Turn* exercises, which allow you to stop, and try to apply the concept. These critical thinking exercises help you solidify the process and well as solve the problem. The text then follows through with a sample.

The end-of-chapter material consists of a summary, a list of key terms, review questions, and exercises for the BITS Corporation, Colonial Adventure Tours, and Sports Physical Therapy 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 BITS Corporation, Colonial Adventure Tours, and Sports Physical Therapy 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: Kate Mason, Associate Product Manager; Michele Stulga, Content Project Manager, Maria Garguilo and Tyler Sally, Content Developers; Diana Graham, Art Director; and Sumathy Kumaran, Associate Product Manager at Lumina Datamatics, Inc.

CHAPTER 1

# INTRODUCTION TO DATABASE MANAGEMENT

## LEARNING OBJECTIVES

- Introduce Burk IT Solutions (BITS), 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, a company that is used in a case that appears at the end of each chapter
- Introduce Sports Physical Therapy, a company that is used in another case that appears at the end of each chapter

## INTRODUCTION

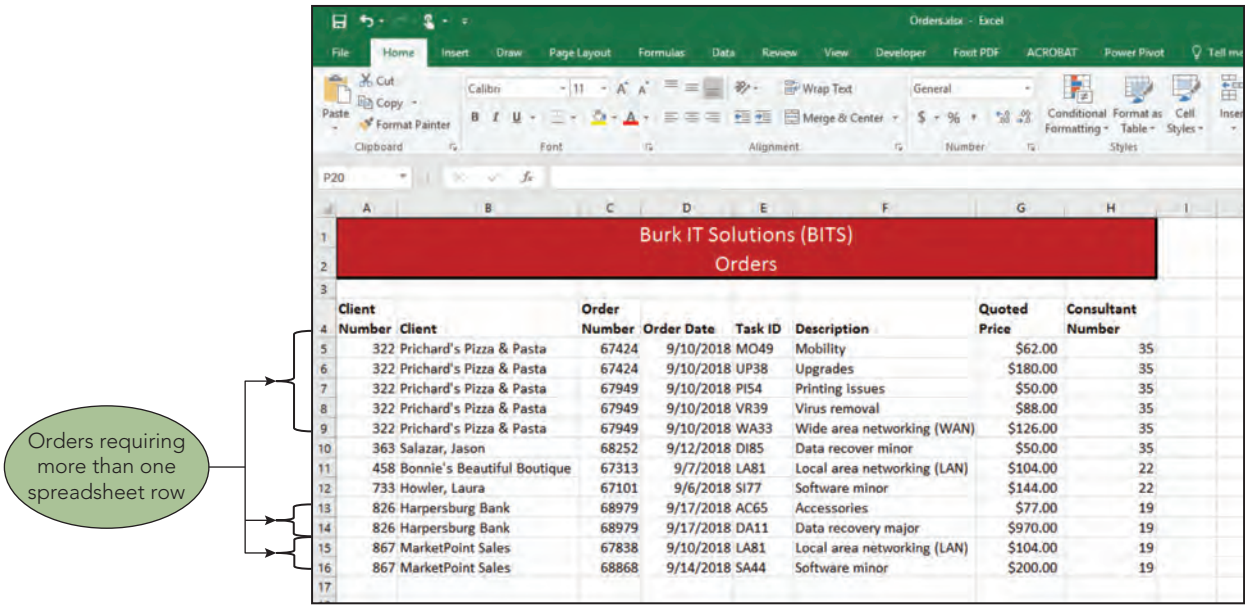
In this chapter, you will examine the requirements of Burk IT Solutions (BITS), a company that will be used in many examples in this chapter and in the rest of the text. You will learn how BITS initially stored its data, what problems employees encountered with the storage method, and why management decided to employ a database management system (DBMS). Then you will study the basic terminology and concepts of databases, database management systems, and big data. You will learn the advantages and disadvantages of database processing. Finally, you will examine the database requirements for Colonial Adventure Tours and Sports Physical Therapy, the companies featured in the cases that appear at the end of each chapter.

## BITS COMPANY BACKGROUND

Burk IT Solutions (BITS) is a local computer hardware and software consulting company whose IT consultants perform functions such as hardware repair, software installation, networking solutions, and system security—for both individuals and small businesses. As the company was getting started, they kept track of their clients in a spreadsheet; they used a homegrown job order/inventory program to keep track of work orders. Management has now determined that the company's recent growth means it is no longer feasible to use those programs to maintain its data.

What has led the managers at BITS to this decision? One of the company's spreadsheets, shown in Figure 1-1 on the next page, displays sample work order data, and illustrates the company's problems with the spreadsheet approach. For each work order, the spreadsheet displays the number and name of the client, the work order number and date, the task ID, a description, the quoted price or estimate, and the number of the consultant assigned to the client. Note that Harpersburg Bank (order number 68979) appears in two rows because this client needed two different jobs performed in its order. In the case of Prichard's Pizza & Pasta, the company placed two different orders (order numbers 67424 and 67949). In the first order, the client needed help with mobility (connectivity), which would also require an upgrade. In the second order,

the client had printer issues along with a possible virus. The client also was experiencing difficulty with the network between two stores (wide area networking). The result was five lines in the spreadsheet, two work order numbers, and various job task IDs.




| Client Number | Client | Order Number | Order Date | Task ID | Description | Quoted Price | Consultant Number |
|---|---|---|---|---|---|---|---|
| 322 | Prichard's Pizza & Pasta | 67424 | 9/10/2018 | MO49 | Mobility | $62.00 | 35 |
| 322 | Prichard's Pizza & Pasta | 67424 | 9/10/2018 | UP38 | Upgrades | $180.00 | 35 |
| 322 | Prichard's Pizza & Pasta | 67949 | 9/10/2018 | PI54 | Printing issues | $50.00 | 35 |
| 322 | Prichard's Pizza & Pasta | 67949 | 9/10/2018 | VR39 | Virus removal | $88.00 | 35 |
| 322 | Prichard's Pizza & Pasta | 67949 | 9/10/2018 | WA33 | Wide area networking (WAN) | $126.00 | 35 |
| 363 | Salazar, Jason | 68252 | 9/12/2018 | DI85 | Data recover minor | $50.00 | 35 |
| 458 | Bonnie's Beautiful Boutique | 67313 | 9/7/2018 | LA81 | Local area networking (LAN) | $104.00 | 22 |
| 733 | Howler, Laura | 67101 | 9/6/2018 | SI77 | Software minor | $144.00 | 22 |
| 826 | Harpersburg Bank | 68979 | 9/17/2018 | AC65 | Accessories | $77.00 | 19 |
| 826 | Harpersburg Bank | 68979 | 9/17/2018 | DA11 | Data recovery major | $970.00 | 19 |
| 867 | MarketPoint Sales | 67838 | 9/10/2018 | LA81 | Local area networking (LAN) | $104.00 | 19 |
| 867 | MarketPoint Sales | 68868 | 9/14/2018 | SA44 | Software minor | $200.00 | 19 |

**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 Client column because the name of a client is stored in more than one place. Both rows for client number 867, for example, store "MarketPoint Sales" as the client name. In the Orders spreadsheet, redundancy also occurs in other columns, such as the Client Number and Order Number columns.

## Q & A 1-1

**Question:** What problems does redundancy cause?
**Answer:** Redundancy can cause inconsistencies in the data, leading to missing information and poor decision making from the data. The accuracy of the data is the most important factor. For example, you might enter "MarketPoint Sales" and "Market Point Sales" on separate rows in the Client column, and then be unsure about the correct version of this client's name. Further, if this client's name is spelled in two different ways and you use the search feature with one of the two values, you would find a single match instead of two matches.

When you need to change data, redundancy also makes your changes more cumbersome and time-consuming. For example, if you incorrectly enter "Harpersberg Bank" in the Client column, you would need to correct it in two places. Even if you use the global find-and-replace feature, multiple changes require more editing time than does a single change.

Finally, while storage space is relatively inexpensive, 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. The files also take longer to save and open.

Difficulty accessing related data is another problem that employees at BITS encounter with their spreadsheets. For example, if you want to see a client's address and the scheduled date and time, 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. In addition, 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 work orders at BITS 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 BITS must maintain the following information about its consultants, clients, categories of IT tasks, and work orders:

- The consultant number, last name, first name, address, normal weekly hours, and rate of pay for each consultant.
- The client number, name, address, current balance, and credit limit for each client, as well as the number of the consultant who typically works with the client.
- The order number, task, description, scheduled date, and quoted estimate.

BITS must store information about orders for invoicing purposes. Figure 1-2 shows a sample invoice.

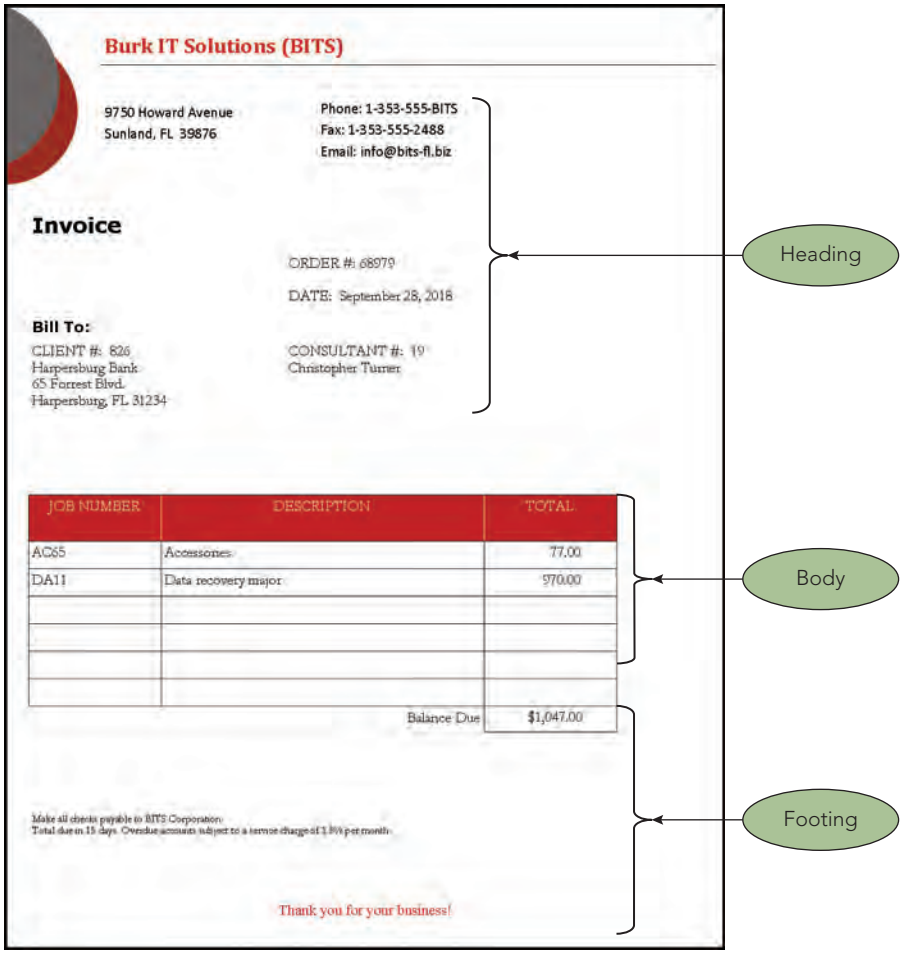



**FIGURE 1-2** Sample invoice

- The heading (top) of the order contains the BITS Corporation's name, address, phone, fax, and email; the word "Invoice"; the order number and date; the client's number, name, and address; and the consultant's number and name.

- The body of the order contains one or more order lines, sometimes called line items. Each order line contains a job number, a description, and the total for the item.
- The footing (bottom) of the order contains the balance due.

BITS also must store the following items for each client's order:

- For each work order, the company must store the order number, the date the order was placed, and the number of the client that placed the order. The client's name and address as well as the number of the consultant who represents the client are stored with the client information. The name of the consultant is stored with the consultant information.
- For each order line, the company must store the order number, the task ID, the scheduled date of the repair, and the quoted estimate or price. If the job may result in taking more time or resources, the client is called and the quoted price is adjusted. Remember that the description and task category are stored with the information about the IT task.
- 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 BITS 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, BITS is interested in categories such as consultants, clients, orders, and tasks. A school is interested in students, faculty, and classes; a real estate agency is interested in clients, houses, and agents; a distributor is interested customers, orders, and inventory; and a car dealership is interested in clients, vehicles, and manufacturers.

Besides wanting to store data that pertains to more than one task, BITS is interested in the relationships between the clients, and consultants. For example, BITS may want to assign consultants that specialize in one area of IT. They need to be able to associate orders with the clients that ordered them, the consultants who coordinated the work, and the jobs that the client 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 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 SOLUTION

After studying the alternatives to using spreadsheet software, BITS decided to switch to a database system. A database is a structure that contains data about many different categories of information and about the relationships between those categories. The BITS database, for example, will contain information about consultants, clients, orders, and tasks. It also will provide facts that relate consultants to the clients they service, and clients to the work orders they currently have placed.

With a database, employees can enter the number of a particular work order and identify which client placed the order. Alternately, employees can start with a client and find all work orders the client placed, together with descriptions of the task. Using a database, BITS not only can maintain its data better but also can use the data in the database to produce a variety of reports and to answer different types of questions.

### Database Terminology

There are some terms and concepts in the database environment that are important 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 BITS, for example, are consultants, clients, orders, and tasks. Entities sometimes are represented by a **table** of data in database systems.

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 BITS, the attributes of interest for the entity *client* are such things as client 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, Consultant and Client, along with the attributes for each entity. The Consultant entity has nine attributes: ConsltNum, LastName, FirstName, Street, City, State, ZipCode, Hours, and Rate. The attributes are the same as the columns in a spreadsheet. The Client entity has nine attributes: ClientNum, ClientName, Street, City, State, ZipCode, Balance, CreditLimit, and ConsltNum. NOTE: Entity (table) names and attribute (field) names should be easy to understand, concise, indicative of their content, and contain no spaces.

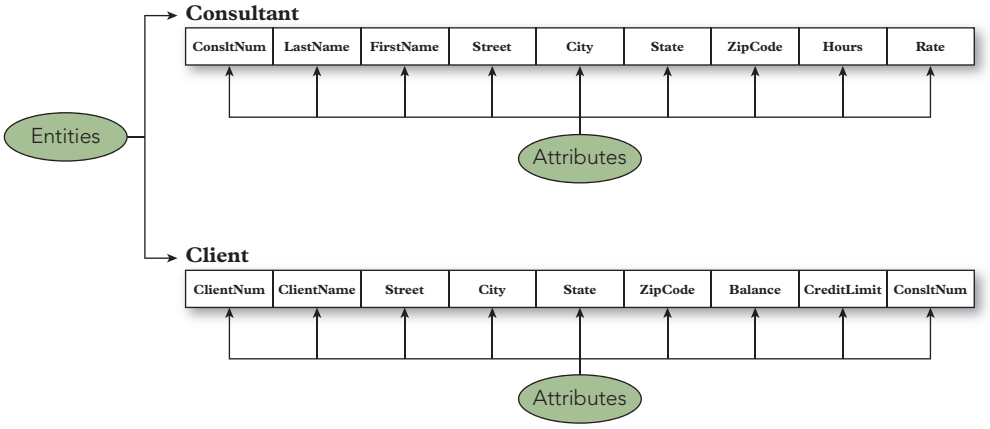



**FIGURE 1-3** Entities and attributes

The final key database term is relationship. A **relationship** is an association between entities. There is an association between consultants and clients; for example, at BITS, a consultant is associated with all of his or her clients, and a client is associated with its consultant. Technically speaking, a consultant is *related to* all of his or her clients, and a client is *related to* its consultant.

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

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

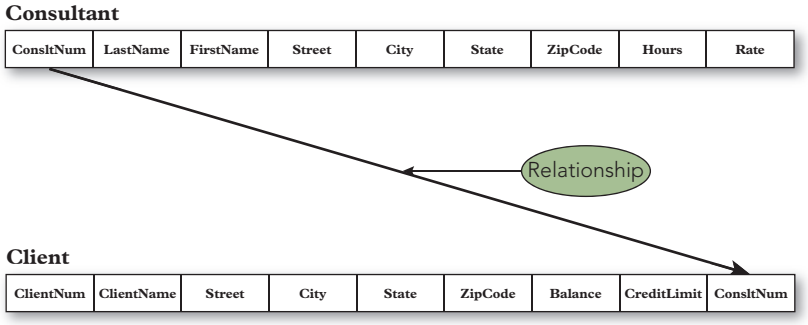



**FIGURE 1-4** One-to-many relationship

## Storing Data

Spreadsheets, word-processed documents, webpages, and other computer information sources are stored in files. A file that is used to store data, 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.

Additionally, a database holds information about the relationships among the various entities. Not only will the BITS database have information about both consultants and clients, it also will hold information relating consultants to the clients they service, clients to work orders, tasks to work 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 among the entities.

How does a database handle these entities, attributes of entities, and relationships among entities? Entities and attributes are fairly simple. Each entity has its own table. In the BITS database, for example, there will be one table for consultants, one table for clients, and so on. The attributes of an entity become the columns in the table. In the table for consultants, for example, there will be a column for the consultant number, a column for the consultant last name, and so on. Within each table, a **row** of data corresponds to one record. A **record** is a group of fields related to one item in a table.

What about relationships between entities? At BITS, there is a one-to-many relationship between consultants and clients. (Each consultant is related to the many clients that he or she represents, and each client is related to the one consultant who represents the client.) How is this relationship handled in a database system? It is handled by using common columns in the two tables. Consider Figure 1-4 on the previous page again. The ConsltNum column in the Consultant table and the ConsltNum column in the Client table are used to implement the relationship between consultants and clients. (It is not unusual to abbreviate column names in a database.) Given a consultant, you can use these columns to determine all the clients that he or she represents; given a client, you can use these columns to find the consultant who represents the client.

How will BITS store its data via tables in a database? Figure 1-5 shows sample data for BITS.

In the Consultant table, you see that there are four consultants whose numbers are 19, 22, 35, and 51. The name of consultant 19 is Christopher Turner. His street address is 554 Brown Dr. He lives in Tri City, FL, and his zip code is 32889. He typically works 40 hours a week with a pay rate of $22.50 per hour.

BITS has 12 clients at this time, which are identified with the numbers 143, 175, 299, 322, 363, 405, 449, 458, 677, 733, 826, 867. The name of client number 143 is Jarrod Hershey. (The last name is listed first for alphabetical/sorting reasons. Not all clients have a first and last name.) This client's address is 135 E. Mill Street in Easton, FL, with a zip code of 33998. The client's current balance is $1,904.55, and its credit limit is $2,500.00. The number 19 in the ConsltNum column indicates that Jarrod Hershey is represented by consultant 19 (Christopher Turner—see Consultant table).

In the table named Tasks, you see that BITS currently has 16 tasks, whose task ID numbers are AC65, DA11, DI85, HA63, HI31, LA81, MO49, OT99, PI54, SA44, SI77, SI91, UP38, VR39, WA33, and WC19. TaskID AC65 is Accessories, and BITS normal pricing is $80.00 for installing and troubleshooting accessories such as storage devices and monitors. The Accessories item is in the ACC category. Other categories include DRM (data recovery), HAM (hardware issues), and SOM (software issues), among others. The company has a $50 minimum charge on all service calls.

In the table named WorkOrders, you see that there are eight orders, which are identified with the numbers 67101, 67313, 67424, 67838, 67949, 68252, 68868, and 68979. Order number 67101 was placed on September 6, 2018, by client 733 (Laura Howler—see Client table).

### Consultant

| ConsltNum | LastName | FirstName | Street | City | State | ZipCode | Hours | Rate |
|---|---|---|---|---|---|---|---|---|
| 19 | Turner | Christopher | 554 Brown Dr. | Tri City | FL | 32889 | 40 | $22.50 |
| 22 | Jordan | Patrick | 2287 Port Rd. | Easton | FL | 33998 | 40 | $22.50 |
| 35 | Allen | Sarah | 82 Elliott St. | Lizton | FL | 34344 | 35 | $20.00 |
| 51 | Shields | Tom | 373 Lincoln Ln. | Sunland | FL | 39876 | 10 | $15.00 |

### Client

| ClientNum | ClientName | Street | City | State | ZipCode | Balance | CreditLimit | ConsltNum |
|---|---|---|---|---|---|---|---|---|
| 143 | Hershey, Jarrod | 135 E. Mill Street | Easton | FL | 33998 | $1,904.55 | $2,500.00 | 19 |
| 175 | Goduto, Sean | 12 Saratoga Parkway | Tri City | FL | 32889 | $2,814.55 | $5,000.00 | 19 |
| 299 | Two Crafty Cousins | 9787 NCR 350 West | Sunland | FL | 39876 | $8,354.00 | $10,000.00 | 22 |
| 322 | Prichard's Pizza & Pasta | 501 Air Parkway | Lizton | FL | 34344 | $7,335.55 | $10,000.00 | 35 |
| 363 | Salazar, Jason | 56473 Cherry Tree Dr. | Easton | FL | 33998 | $900.75 | $2,500.00 | 35 |
| 405 | Fisherman's Spot Shop | 49 Elwood Ave. | Harpersburg | FL | 31234 | $4,113.40 | $7,500.00 | 19 |
| 449 | Seymour, Lindsey | 4091 Brentwood Ln | Amo | FL | 34466 | $557.70 | $5,000.00 | 22 |
| 458 | Bonnie's Beautiful Boutique | 9565 Ridge Rd. | Tri City | FL | 32889 | $4,053.80 | $7,500.00 | 22 |
| 677 | Yates, Nick | 231 Day Rd. | Sunland | FL | 39876 | $2,523.80 | $2,500.00 | 35 |
| 733 | Howler, Laura | 1368 E. 1000 S. | Lizton | FL | 34344 | $3,658.05 | $5,000.00 | 22 |
| 826 | Harpersburg Bank | 65 Forrest Blvd. | Harpersburg | FL | 31234 | $6,824.55 | $10,000.00 | 19 |
| 867 | MarketPoint Sales | 826 Host St. | Easton | FL | 33998 | $3,089.00 | $5,000.00 | 19 |

### Tasks

| TaskID | Description | Category | Price |
|---|---|---|---|
| AC65 | Accessories | ACC | $80.00 |
| DA11 | Data recovery major | DRM | $175.00 |
| DI85 | Data recovery minor | DRM | $50.00 |
| HA63 | Hardware major | HAM | $225.00 |
| HI31 | Hardware minor | HAM | $165.70 |
| LA81 | Local area networking (LAN) | LAN | $104 00 |
| MO49 | Mobility | MOB | $65.00 |
| OT99 | Other work | OTH | $99.99 |
| PI54 | Printing issues | PRI | $50.00 |
| SA44 | Software major | SOM | $200.00 |
| SI77 | Software minor | SOM | $144.00 |
| SI91 | Security install/repair | SIR | $126.00 |
| UP38 | Upgrades | UPG | $185.00 |
| VR39 | Virus removal | VIR | $90.00 |
| WA33 | Wide area networking (WAN) | WAN | $130.00 |
| WC19 | Web connectivity | WEC | $75.00 |

### OrderLine

| OrderNum | TaskID | ScheduledDate | QuotedPrice |
|---|---|---|---|
| 67101 | SI77 | 9/10/2018 | $144.00 |
| 67313 | LA81 | 9/12/2018 | $104.00 |
| 67424 | MO49 | 9/14/2018 | $65.00 |
| 67424 | UP38 | 9/14/2018 | $185.00 |
| 67838 | LA81 | 9/20/2018 | $104.00 |
| 67949 | PI54 | 9/21/2018 | $50.00 |
| 67949 | VR39 | 9/21/2018 | $88.00 |
| 67949 | WA33 | 9/21/2018 | $126.00 |
| 68252 | DI85 | 9/24/2018 | $50.00 |
| 68868 | SA44 | 9/24/2018 | $200.00 |
| 68979 | AC65 | 9/27/2018 | $77.00 |
| 68979 | DA11 | 9/27/2018 | $970.00 |

### WorkOrders

| OrderNum | OrderDate | ClientNum |
|---|---|---|
| 67101 | 9/6/2018 | 733 |
| 67313 | 9/7/2018 | 458 |
| 67424 | 9/10/2018 | 322 |
| 67838 | 9/10/2018 | 867 |
| 67949 | 9/10/2018 | 322 |
| 68252 | 9/12/2018 | 363 |
| 68868 | 9/14/2018 | 867 |
| 68979 | 9/17/2018 | 826 |

**FIGURE 1-5** Sample data for BITS

The table named OrderLine on the previous page 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 WorkOrders table? The answer is yes. The WorkOrders 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 on the previous page, with the same dates and clients. In addition, each table row in Figure 1-6 contains all the order lines for a given order. Examining the third row, for example, you see that order 67424 has two order lines. One of the order lines is for MO49 (mobility issues), and the quoted price is $65.00. The other order line is for UP38 (upgrades), and the quoted price is $185.00.

**WorkOrders**

| OrderNum | OrderDate | ClientNum | TaskID | QuotedPrice |
|---|---|---|---|---|
| 67101 | 9/6/2018 | 733 | SI77 | $144.00 |
| 67313 | 9/7/2018 | 458 | LA81 | $104.00 |
| 67424 | 9/10/2018 | 322 | MO49 | $65.00 |
| | | | UP38 | $185.00 |
| 67838 | 9/10/2018 | 867 | LA81 | $104.00 |
| 67949 | 9/10/2018 | 322 | PI54 | $50.00 |
| | | | VR39 | $88.00 |
| | | | WA33 | $126.00 |
| 68252 | 9/12/2018 | 363 | DI85 | $50.00 |
| 68868 | 9/14/2018 | 867 | SA44 | $200.00 |
| 68979 | 9/17/2018 | 826 | AC65 | $77.00 |
| | | | DA11 | $970.00 |

**FIGURE 1-6** Alternative WorkOrders table structure

## Q & A 1-2

**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 third and fourth rows. The third row indicates that there is an order line in order number 67424 for task MO49 with a quoted price of $65.00. The fourth row indicates that there is an order line in order 67424 for upgrades with a quoted price of $185.00. Thus, the information in Figure 1-6 is represented in Figure 1-5 with two separate rows rather than in one row.

## Q & A 1-3

**Question:** Why is the quoted price in the OrderLine table different from the price listed in the Tasks table?
**Answer:** The estimator at BITS Corporation talks to each client or customer as he or she calls in to request services, and then enters the work order and order line. The estimator evaluates the need and may adjust the price up or down depending on the situation and how much time may be involved. In the Tasks table, the prices are listed for a typical hour related to the task at hand. The actual service or repair may take more time. For example, Task DA11 is listed at $175.00. However, in the last order line, the estimator, after talking with the client, quoted a price of $970.00 for the large amount of work involved.

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 OrderLine 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 67424, for example, it is crucial to know that TaskID UP38 corresponds to the dollar figure $185.00 in the QuotedPrice column, not to the $65.00.

In addition, having a more complex table means that there are practical issues to worry about, such as the following:

- 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 a task ID, how do you determine which orders contain order lines for that task?

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 on page 7. In Figure 1-5, there are no multiple entries to worry about, it does not matter how many order lines exist for any work order, and it is easy to find every order that contains an order line for a given task (just look for all order lines with the given TaskID). In general, this simpler structure is preferable, which is why the order lines appear in a separate table.

To test your understanding of the BITS data, use the data shown in Figure 1-5 on page 7 to answer the following questions.

## Q & A 1-4

**Question:** What are the numbers of the clients represented by Christopher Turner?
**Answer:** 143, 175, 405, and 867. (Look up the ConsltNum value for Christopher Turner in the Consultant table and obtain the number 19. Then find all clients in the Client table that have the number 19 in the ConsltNum column.)

## Q & A 1-5

**Question:** What is the name of the client that placed order 67424, and what is the name of the consultant who represents this client?
**Answer:** Prichard's Pizza & Pasta is the client, and Sarah Allen is the consultant. (Look up the ClientNum value in the Orders table for order number 67424 and obtain the number 322. Then find the client in the Client table with a ClientNum value of 322. Using this client's ConsltNum value, which is 35, find the name of the consultant in the Consultant table.)

## Q & A 1-6

**Question:** List all the items that appear in order 67949. For each item, give the description, number ordered, and quoted price.
**Answer:** TaskID: PI54; description: Printing issues; category: PRI; and quoted price: $50.00. Also, TaskID: VR39; description: Virus removal; category: VIR; and quoted price $88.00. Finally, TaskID: WA33; description: Wide area networking (WAN); category: WAN; and quoted price: $126.00. The scheduled date is 9/21/2018. (Look up each OrderLine table row in which the order number is 67949. Each row contains a TaskID, the ScheduledDate, and the QuotedPrice. Use the TaskID to look up the corresponding description in the Tasks table.)

## Q & A 1-7

**Question:** Why is the QuotedPrice column in the OrderLine table? Couldn't you just use the task ID to look up the price in the Tasks table?
**Answer:** If the QuotedPrice column did not appear in the OrderLine table, you would need to obtain the price for a service on an order line by looking up the price in the Tasks table. Although this might not be a bad practice, it prevents BITS from charging different prices to different clients for the same item. Because BITS wants the flexibility to quote and charge different prices to different clients, the QuotedPrice column is included in the OrderLine table. If you examine the OrderLine table, you will see cases in which the estimated price matches the actual price in the Tasks table and cases in which the estimated price differs. For example, in order number 67949, the scheduler at BITS quoted a price to Prichard's Pizza & Pasta of 126.00 (for TaskID WA33) rather than the regular price of 130.00 (shown in the Tasks table). The reduction might lead you to think the client received a slight discount for its multiple task order.