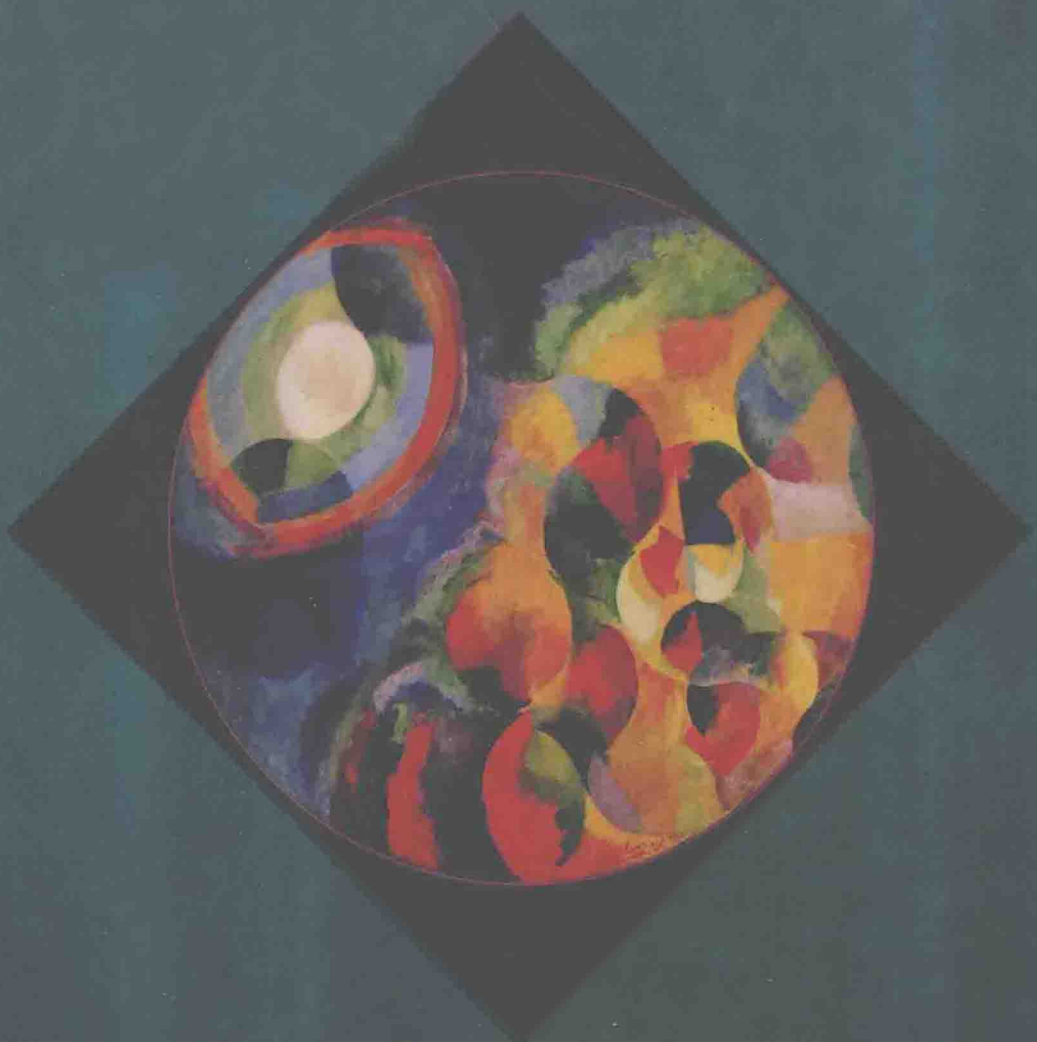
SALSA for the Desktop CD INCLUDED

DATABASE PROCESSING

FUNDAMENTALS, DESIGN, AND IMPLEMENTATION

SIXTH EDITION



DAVID M. KROENKE

6th EDITION

DATABASE PROCESSING

Fundamentals, Design, and Implementation

David M. Kroenke

Prentice Hall, Upper Saddle River, N.J. 07458

Library of Congress Cataloging-in-Publication Data
Kroenke, David M.
Database processing : fundamentals, design, and implementation / David M. Kroenke. — 6th ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-13-737842-4
1. Database management. I. Title.
QA76.9.D3K737 1997
005.74—dc21 97-62
CIP

Acquisitions Editor: JO-ANN DELUCA
Editorial Assistant: MARC OLIVER
Executive Marketing Manager: NANCY EVANS
Editor-in-Chief: PJ BOARDMAN
Director of Production and Manufacturing: JOANNE JAY
Managing Editor: KATHERINE EVANCIE
Production Editor: SUSAN RIFKIN
Production Coordinator: CINDY SPREDER
Manufacturing Manager: VINCENT SCELTA
Senior Manufacturing Supervisor: PAUL SMOLENSKI
Manufacturing Buyer: ALANA ZDINAK
Electronic Art Supervisor: WARREN FISCHBACH
Artist: ANNIE BARTELL
Senior Designer: SUZANNE BEHNKE
Interior Design: ROSEMARIE VOTTA
Cover Design: PAUL GOURHAM
Composition: CARLISLE COMMUNICATIONS, LTD.
Cover Art/Photo: DELAUNAY, ROBERT. Simultaneous Contrasts: Sun and Moon. (Dated on painting 1912.) Oil on canvas, 53" in diameter. The Museum of Modern Art, New York. Purchase.

Printed in the United States of America
10 9 8 7 6 5 4 3 2

ISBN 0-13-737842-4

Prentice-Hall International (UK) Limited, London
Prentice-Hall of Australia Pty. Limited, Sydney
Prentice-Hall Canada, Inc., Toronto
Prentice-Hall Hispanoamericana, S.A., Mexico
Prentice-Hall of India Private Limited, New Delhi
Prentice-Hall of Japan, Inc., Tokyo
Simon & Schuster Asia Pte. Ltd., Singapore
Editora Prentice-Hall do Brasil, Ltda., Rio de Janeiro

PREFACE

This text is intended for use in a one-term course on database processing. The purpose of this text is, first and foremost, to teach the fundamental skills students need to be successful developing personal databases on their own and working as members of a team when building workgroup and enterprise databases. Given this foundation, the second purpose is to discuss and illustrate emerging technology and applications.

■ Course Foundations

Data Modeling
Relational Database Design
Knowledge of SQL
Application Design Fundamentals
Practice with Different Examples
Multi-User Processing

For a database to be effective, its structure must logically reflect the users' view of their world. Such consistency does not arise automatically. No wizard will provide it. Instead, it must be developed through the hard work of understanding the users' model of their world. This means that any person or team who hopes to build effective databases and database applications must be able to create data models of the users' needs. Hence, I believe learning solid modeling skills is critically important in this course—in fact, the single most important skill to be learned.

Of course, once the users' requirements have been modeled, they must be translated into a database design that accurately represents this model. Therefore, the second foundation skill of this course is database design. Today, the graduates of this course will encounter relational databases almost exclusively; hence, this text presents the skills necessary to transform data models into relational database designs.

In fact, with the exceptions of Chapter 17, which presents object-oriented database processing, and Chapter 18, which surveys the DL/I and the CODASYL DBTG models, all of this text assumes the use of relational database technology. This leads to the third foundational skill: knowledge of SQL.

SQL is the lingua franca of database processing. While one could argue that SQL has failed as an end-user query tool, supplanted today by Query by Example and Query by Form, there is no question that SQL is the standard means by which application programs access and update relational databases. SQL statements are important for creating views for forms and reports as well.

No organization wants a database for the sheer joy of having a database. People want to keep track of things and they want to pose queries and generate reports to gain information about those things. Database applications are the means by which this is done, and therefore, a knowledge of the relationship of application design to database design and the users' data model is another foundation skill. This topic overlaps with the systems development course. To avoid conflict and repetition with that course, this text focuses on application design considerations that pertain directly to databases and data models. See, for example, Chapter 8 and the ODBC section in Chapter 14. Hence, whether or not the student has already taken the systems class, the material presented here should be understandable and useful.

Finally, there is no substitute for experience. To ensure that every student can have hands-on experience with a software product for creating data models and database designs, this text is accompanied by the commercial version of SALSA for the Desktop. Of course, if your university has licenses for other products such as Microsoft Access, you will also want to use those products. The goal in providing

SALSA is to ensure that students have at least one commercial product at their disposal. More information on the use of Access and SALSA is presented later in the Preface.

Years ago, database products were so difficult to use that the students had time to develop, at best, a single database. Often even that was too much work. Today, however, with the productivity inherent in modern products, students should be able to develop several data models, if not complete databases. If the projects are suitably chosen, the students will learn firsthand how different data models impact the design and characteristics of the databases and related applications.

Finally, many databases are used by more than one person at a time. This means students need to learn the special requirements, techniques, and technologies of multi-user processing including concurrency control, reliability, security, and database administration.

Important New Topics in This Edition

ODBC
Data Warehousing
Intranet and Internet Databases
Multimedia DBs
Object Databases
SQL3 and ODMB

Databases are everywhere. Whereas at one time databases were the sole province of the MIS department that applied the technology only to enterprise databases on mainframes behind locked doors, today we find databases on personal computers, in client-server workgroup settings, and as the backbone of data warehouses, multimedia applications, and network databases on corporate intranets and the Internet.

This text discusses a number of new technologies that support these applications. The Open Database Connectivity standard has become important in client-server applications and will increasingly be important for network applications as well. Many organizations are currently attempting to increase their return on their data resource by developing data warehouses. Both intranets and the Internet have become new delivery mechanisms for database applications. Such networks are likely to dramatically increase the need for skilled database designers. In addition, databases are being used to support multimedia applications. In the future, it is likely that most database applications will include at least some multimedia elements.

Object-oriented databases are also becoming important, although, at present, they are far less common than relational databases. As object-oriented programming continues to increase in popularity, the demand for such databases will grow. It is difficult to assess whether such databases will result as an outgrowth of relational databases (along the lines of the SQL3 standard), or whether organizations will find a way to migrate to DBMS with completely new architectures (along the lines of the ODMG standard). Today, I believe the wise course of action is to follow both, while focusing the bulk of this course on the current form of the relational model.

Organization of This Book

This text is organized into six parts. Part I provides an introduction. A number of different types of databases and applications are presented in Chapter 1, along with a brief history of database processing. Chapter 2 illustrates the components of a DBMS and provides an overview of the process of building a database and its applications.

The focus of Part II is on data modeling. Chapter 3 explains the concepts and constructs of the entity-relationship model and illustrates its use. Chapter 4 presents the semantic object in a similar way. Either of these models can be used to express the users' data requirements.

Part III discusses the transformation of data models into relational database designs. Chapter 5 sets the stage by discussing the relational model and normalization. Normalization techniques are then used in Chapter 6 to explain the transformation of entity-relationship data models into relational designs and in Chapter 7 to explain the transformation of semantic object data models into relational designs. Chapter 8 concludes Part III by discussing database application design. The presentation in this chapter is limited to those aspects of application design that pertain directly to database structures.

The implementation of relational databases is presented next in Part IV. Chapter 9 discusses the foundations of relational implementation and briefly surveys relational algebra. Chapter 10 is devoted to a thorough discussion of SQL and Chapter 11 illustrates the implementation of a personal database using the features and functions of Microsoft Access.

Part V discusses concepts, techniques, technology, and products that are important for multi-user databases. Chapter 12 introduces multi-user processing alternatives and describes reliability, security, and database administration. It contains a brief survey of distributed database alternatives, as well. Chapter 13 illustrates the implementation of multi-user data using DB2. This chapter is important not only because it describes a mainframe DBMS, but also because it illustrates the means by which SQL can be embedded in application programs. Client-server database processing and ODBC are the subjects of Chapter 14, and the sharing of enterprise data and data warehousing are discussed in Chapter 15. Part V concludes with Chapter 16 and its discussion of the use of database technology on intranets and the Internet. This chapter also discusses the design of two different multimedia databases.

Nonrelational database processing is addressed in Part VI. Chapter 17 discusses object-oriented databases and standards. It includes a brief summary of object-oriented programming, along with a description of the need for object persistence and a discussion of persistence via traditional file structures, relational databases, and ODBMS. Both the SQL3 and ODMG standards are summarized. Finally, Chapter 18 discusses the major elements of DL/I and the CODASYL DBTG models. Little new development is done with the models, but students may encounter them when working on maintenance or database conversion projects.

I have endeavored to limit the discussions in Chapters 14 through 18 so that all of this material could be presented in one course. The needs of employers vary from one locality to another, and it may be that some of this material is unimportant to your students. If so, Chapters 14 through 18 can be read in any order and any of the chapters can be omitted without loss of continuity.

■ Use of Microsoft Access and SALSA for the Desktop

I have used Microsoft Access to illustrate discussion points throughout this text. I understand that many institutions do not have a license for students to use this product, and further, not every professor would choose this product for illustration. Using Access is an attempt to balance, on the one hand, the need for the student to see how abstract concepts apply to a real DBMS, and, on the other, the desire not to indicate that Access is the only choice. In fact, Access is an excellent personal DBMS product, and it also happens to be, by a wide margin, the world's most popular DBMS product. Hence, students are most likely to encounter it. IBM (Lotus's) Approach and Borland's Paradox for Windows are two other excellent products.

This text includes a copy of SALSA for the Desktop. Except for a section in Chapter 17, I have not illustrated the use of this product in this text. I have, however, presented the semantic object model, as a model, using product-independent illustrations. I have done this because I believe that students will find the semantic object model easy to use and powerful. Dan Rogers, writing in the *VB Tech Journal,* said, "The arrangement (of semantic objects) feels quite natural and it leads me to believe that (semantic object modeling) is a break-through modeling technique that can bridge the gap between the business and technical worlds."[1]

Using SALSA is entirely optional; no section in the text depends on it. As stated earlier, however, I believe that the students' experiences in data modeling will be significantly enhanced if they do perform exercises and accomplish some of the end-of-chapter projects. Be aware that SALSA, however, is designed to hide the technology that this course teaches. Thus, after building a semantic object model, the user clicks a button and all of the relations, foreign keys, intersection tables, SQL, and so forth are generated behind the scenes. Thus, while SALSA can be used to create models and to observe the impact of changes in those models on forms and reports, the students will need to cause SALSA to print supplementary reports to determine the relational structure that is generated from their models. Or, if you have Microsoft Access, the students can have SALSA generate an Access database and then examine that database in Access. These issues are discussed in the SALSA projects at the end of Chapters 4 and 7.

My hope is that using SALSA can raise the level of discourse in this class. Instead of spending many hours discussing the placement of foreign keys, we can spend some time on foreign keys, but invest the bulk of the time discussing how well a particular data model represents the needs of the business that it is intended to support. I believe that this hope is consistent with the general trend in technology. Over time, all DBMS vendors will find ways to hide more and more of the underlying data structures, and data modeling will become the primary focus of the database development process.

Acknowledgments

I would like to thank the following reviewers of this book who provided many helpful and insightful comments:

Linda Behrens, *University of Central Oklahoma*
Donna Davis, *University of Southern Mississippi*
Gordon Everest, *University of Minnesota*
William R. Friend, *Edmonds Community College*
Richard D. Heath, *St. Cloud State University*
Andrew Kreutzer, *Richard Stockton College*
Julian M. Scher, *New Jersey Institute of Technology*

Also, at Prentice Hall, I would like to thank my editor, Jo-Ann DeLuca; her assistant, Marc Oliver; and Susan Rifkin, who managed the production of this book.

Thanks also to Jim Simpson, John Wall, Kevin Vitale, and John Nicol of Wall Data, who have supported the development of SALSA and who generously agreed to include SALSA for the Desktop in this text. I am indebted to all of the SALSA developers—thanks especially to Chris Olds and Kawai Kenji with whom I have had many enjoyable conversations. Finally, a special thanks to Marvin Mall without whom SALSA might not ever have been finished.

David Kroenke
Seattle, Washington

[1]Roger, Dan "Manage Data with Modeling Tools" *VB Tech Journal,* December 1996, p. 30.

BRIEF CONTENTS

CONTENTS

PART I

INTRODUCTION

Part I introduces the topic of database processing. Chapter 1 surveys the topic and describes four applications that typify the wide range of databases in use today, as well as the benefits and evolution of database technology. Chapter 2 introduces the development of databases and database applications and describes the elements of a database and the functions of a DBMS as well.

The goal of this part is to set the stage for the discussion of data modeling that begins in Chapter 3.

CHAPTER 1

INTRODUCTION TO DATABASE PROCESSING

This book presents the fundamental skills you will need to be able to design and develop a database and its related applications. Many students find this subject enjoyable and interesting, even though it can be challenging. Learning this material will certainly be worth the effort you put forth as these skills are in great demand, and many fulfilling and well-paying database development positions are available. The reason for this demand is that almost every business person needs the benefits of database technology, but very few know how to use that technology and few want to learn. Hence, there is a great need for people who can serve as liaisons between business users and database technology. The goal of this book is to help you learn the knowledge and skills you will need to provide that service.

FOUR DATABASE EXAMPLES

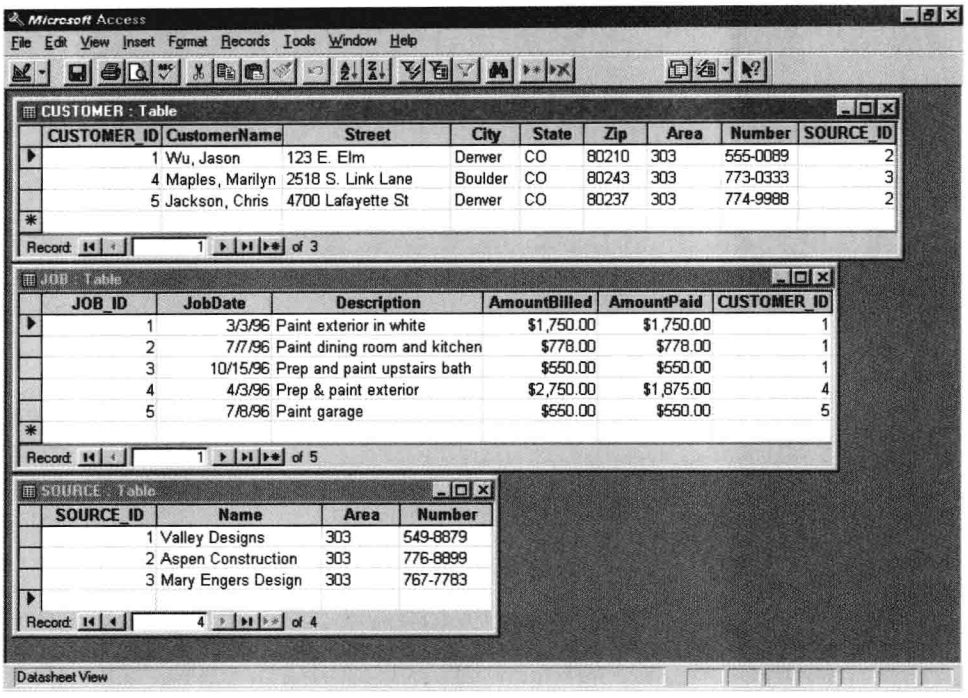
The purpose of a database is to help people keep track of things. The classical database applications concern the tracking of items like orders, customers, jobs, employees, phone calls, or other items of interest to a business person. Recently, as databases have become more readily available, database technology has been applied to new areas such as databases for the Internet or for organizational intranets. Also, databases are increasingly used to create and maintain multimedia applications. Consider several typical examples.

MARY RICHARDS HOUSEPAINTING

Mary Richards is a professional housepainter who owns and operates a small company consisting of herself, another professional painter, and, when needed, part-time painters. Mary has been in business for ten years and has earned a reputation as a high-quality painter who works for a reasonable rate. Mary gets most of her work through repeat business from customers who hire her to paint their houses and also from their word-of-mouth referrals. In addition, Mary gets some work from building contractors and professional interior designers.

Customers remember Mary far better than she remembers them. Indeed, sometimes she is embarrassed when a customer calls and says something like, "Hi Mary, this is John Maples. You painted my house three years ago." Mary knows she

FIGURE 1-1 *Tables of Data for Mary Richards Housepainting*



is supposed to remember the caller and the work she did for him, but since she paints more than fifty houses a year, it usually is difficult for her to do so. This situation becomes worse when the customer says something like, "My neighbor liked the job you did on our house and would like something similar done to her house."

In order to help her memory and to keep better track of her business records, Mary had a consultant develop a database and database application that she uses on her personal computer. The database stores records regarding customers, jobs, and sources in the form of tables, as shown in the example in Figure 1-1.

It is the job of a program called a database management system (DBMS) to store and retrieve the data in these tables. Unfortunately, when such data are in the form of tables, they are not very useful to Mary. Rather, she would like to know how customers and jobs and referrals relate to one another, for example, what jobs she has done for a particular customer or what customers have been referred by a particular person.

To provide this capability, Mary's consultant created a database application that processes data entry forms and produces reports. Consider the example form in Figure 1-2. Here, Mary keys in data about customers such as name, phone, and address. She also links the customer to a particular referral source and keys in data about jobs performed for the customer. This data can then be displayed in reports like the one shown in Figure 1-3.

It is the function of the database application and the DBMS to process the form and store the data that are entered into tables like those in Figure 1-1. Similarly, the application and DBMS extract data from tables like those in Figure 1-1 to create a report like the one in Figure 1-3.

Consider again the data in Figure 1-1, and notice that the rows in the tables cross-reference and are linked to one another. Each JOB contains the Customer_ID of the CUSTOMER who purchased that JOB, and each CUSTOMER contains the Source_ID of the person who referred that customer. These references are used to