Business Intelligence and Analytics

Systems for Decision Support

TENTH EDITION

Ramesh Sharda • Dursun Delen • Efraim Turban

ALWAYS LEARNING

PEARSON

TENTH EDITION

BUSINESS INTELLIGENCE AND ANALYTICS:

SYSTEMS FOR DECISION SUPPORT

Global Edition

Ramesh Sharda
Oklahoma State University

Dursun Delen
Oklahoma State University

Efraim Turban
University of Hawaii

With contributions by

J. E. Aronson
The University of Georgia

Ting-Peng Liang
National Sun Yat-sen University

David King
JDA Software Group, Inc.

PEARSON

Boston Columbus Indianapolis New York San Francisco Upper Saddle River
Amsterdam Cape Town Dubai London Madrid Milan Munich Paris Montréal Toronto
Delhi Mexico City São Paulo Sydney Hong Kong Seoul Singapore Taipei Tokyo

Editor in Chief: Stephanie Wall
Executive Editor: Bob Horan
Publisher, Global Edition: Laura Dent
Senior Acquisitions Editor, Global Edition: Steven Jackson
Program Manager Team Lead: Ashley Santora
Program Manager: Denise Vaughn
Marketing Manager, International: Kristin Schneider
Project Manager Team Lead: Judy Leale
Project Manager: Tom Benfatti
Assistant Project Editor, Global Edition: Paromita Banerjee
Operations Specialist: Michelle Klein
Senior Manufacturing Controller, Production: Trudy Kimber
Creative Director: Jayne Conte
Cover Image Credit: © Robert Adrian Hillman
Cover Printer: Courier Kendallville
Cover Designer: Jodi Notowitz at Wicked Design
Full-Service Project Management: George Jacob, Integra Software Solutions.
Text Font: ITC Garamond Std

Pearson Education Limited
Edinburgh Gate
Harlow
Essex CM20 2JE
England

and Associated Companies throughout the world

Visit us on the World Wide Web at: www.pearsonglobaleditions.com

Authorized adaptation from the United States edition, entitled Business Intelligence and Analytics: Systems for Decision Support, 10th edition, ISBN 978-0-133-05090-5, by Ramesh Sharda, Dursun Delen, and Efraim Turban, published by Pearson Education © 2014.

ISBN 10: 1-292-00920-9
ISBN 13: 978-1-292-00920-9

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

9 8 7 6 5 4 3 2 1
14 13 12 11 10

Typeset in ITC Garamond Std. Integra Software Solutions

Printed and bound by Courier Kendalville in The United States of America

BRIEF CONTENTS

CONTENTS

PREFACE

Analytics has become the technology driver of this decade. Companies such as IBM, Oracle, Microsoft, and others are creating new organizational units focused on analytics that help businesses become more effective and efficient in their operations. Decision makers are using more computerized tools to support their work. Even consumers are using analytics tools directly or indirectly to make decisions on routine activities such as shopping, healthcare, and entertainment. The field of decision support systems (DSS)/business intelligence (BI) is evolving rapidly to become more focused on innovative applications of data streams that were not even captured some time back, much less analyzed in any significant way. New applications turn up daily in healthcare, sports, entertainment, supply chain management, utilities, and virtually every industry imaginable.

The theme of this revised edition is BI and analytics for enterprise decision support. In addition to traditional decision support applications, this edition expands the reader's understanding of the various types of analytics by providing examples, products, services, and exercises by discussing Web-related issues throughout the text. We highlight Web intelligence/Web analytics, which parallel BI/business analytics (BA) for e-commerce and other Web applications. The book is supported by a Web site (**pearsonglobaleditions.com/sharda**) and also by an independent site at **dssbibook.com**. We will also provide links to software tutorials through a special section of the Web site.

The purpose of this book is to introduce the reader to these technologies that are generally called *analytics* but have been known by other names. The core technology consists of DSS, BI, and various decision-making techniques. We use these terms interchangeably. This book presents the fundamentals of the techniques and the manner in which these systems are constructed and used. We follow an EEE approach to introducing these topics: **Exposure, Experience,** and **Explore**. The book primarily provides **exposure** to various analytics techniques and their applications. The idea is that a student will be inspired to learn from how other organizations have employed analytics to make decisions or to gain a competitive edge. We believe that such **exposure** to what is being done with analytics and how it can be achieved is the key component of learning about analytics. In describing the techniques, we also introduce specific software tools that can be used for developing such applications. The book is not limited to any one software tool, so the students can **experience** these techniques using any number of available software tools. Specific suggestions are given in each chapter, but the student and the professor are able to use this book with many different software tools. Our book's companion Web site will include specific software guides, but students can gain **experience** with these techniques in many different ways. Finally, we hope that this **exposure** and **experience** enable and motivate readers to **explore** the potential of these techniques in their own domain. To facilitate such **exploration,** we include exercises that direct them to Teradata University Network and other sites as well that include team-oriented exercises where appropriate. We will also highlight new and innovative applications that we learn about on the book's companion Web sites.

Most of the specific improvements made in this tenth edition concentrate on three areas: reorganization, content update, and a sharper focus. Despite the many changes, we have preserved the comprehensiveness and user friendliness that have made the text a market leader. We have also reduced the book's size by eliminating older and redundant material and by combining material that was not used by a majority of professors. At the same time, we have kept several of the classical references intact. Finally, we present accurate and updated material that is not available in any other text. We next describe the changes in the tenth edition.

WHAT'S NEW IN THE TENTH EDITION?

With the goal of improving the text, this edition marks a major reorganization of the text to reflect the focus on analytics. The last two editions transformed the book from the traditional DSS to BI and fostered a tight linkage with the Teradata University Network (TUN). This edition is now organized around three major types of analytics. The new edition has many timely additions, and the dated content has been deleted. The following major specific changes have been made:

- ***New organization.*** The book is now organized around three types of analytics: descriptive, predictive, and prescriptive, a classification promoted by INFORMS. After introducing the topics of DSS/BI and analytics in Chapter 1 and covering the foundations of decision making and decision support in Chapter 2, the book begins with an overview of data warehousing and data foundations in Chapter 3. This part then covers descriptive or reporting analytics, specifically, visualization and business performance measurement. Chapters 5–8 cover predictive analytics. Chapters 9–12 cover prescriptive and decision analytics as well as other decision support systems topics. Some of the coverage from Chapter 3–4 in previous editions will now be found in the new Chapters 9 and 10. Chapter 11 covers expert systems as well as the new rule-based systems that are commonly built for implementing analytics. Chapter 12 combines two topics that were key chapters in earlier editions—knowledge management and collaborative systems. Chapter 13 is a new chapter that introduces big data and analytics. Chapter 14 concludes the book with discussion of emerging trends and topics in business analytics, including location intelligence, mobile computing, cloud-based analytics, and privacy/ethical considerations in analytics. This chapter also includes an overview of the analytics ecosystem to help the user explore all of the different ways one can participate and grow in the analytics environment. Thus, the book marks a significant departure from the earlier editions in organization. Of course, it is still possible to teach a course with a traditional DSS focus with this book by covering Chapters 1–4, Chapters 9–12, and possibly Chapter 14.
- ***New chapters.*** The following chapters have been added:

  Chapter 8, *"Web Analytics, Web Mining, and Social Analytics."* This chapter covers the popular topics of Web analytics and social media analytics. It is an almost entirely new chapter (95% new material).
  Chapter 13, *"Big Data and Analytics."* This chapter introduces the hot topics of Big Data and analytics. It covers the basics of major components of Big Data techniques and charcteristics. It is also a new chapter (99% new material).
  Chapter 14, *"Business Analytics: Emerging Trends and Future Impacts."* This chapter examines several new phenomena that are already changing or are likely to change analytics. It includes coverage of geospatial in analytics, location-based analytics applications, consumer-oriented analytical applications, mobile platforms, and cloud-based analytics. It also updates some coverage from the previous edition on ethical and privacy considerations. It concludes with a major discussion of the analytics ecosystem (90% new material).
- ***Streamlined coverage.*** We have made the book shorter by keeping the most commonly used content. We also mostly eliminated the preformatted online content. Instead, we will use a Web site to provide updated content and links on a regular basis. We also reduced the number of references in each chapter.
- ***Revamped author team.*** Building upon the excellent content that has been prepared by the authors of the previous editions (Turban, Aronson, Liang, King, Sharda, and Delen), this edition was revised by Ramesh Sharda and Dursun Delen.

Both Ramesh and Dursun have worked extensively in DSS and analytics and have industry as well as research experience.

- ***A live-update Web site.*** Adopters of the textbook will have access to a Web site that will include links to news stories, software, tutorials, and even YouTube videos related to topics covered in the book. This site will be accessible at **http://dssbibook.com**.
- ***Revised and updated content.*** Almost all of the chapters have new opening vignettes and closing cases that are based on recent stories and events. In addition, application cases throughout the book have been updated to include recent examples of applications of a specific technique/model. These application case stories now include suggested questions for discussion to encourage class discussion as well as further exploration of the specific case and related materials. New Web site links have been added throughout the book. We also deleted many older product links and references. Finally, most chapters have new exercises, Internet assignments, and discussion questions throughout.

Specific changes made in chapters that have been retained from the previous editions are summarized next:

Chapter 1, "An Overview of Business Intelligence, Analytics, and Decision Support," introduces the three types of analytics as proposed by INFORMS: descriptive, predictive, and prescriptive analytics. A noted earlier, this classification is used in guiding the complete reorganization of the book itself. It includes about 50 percent new material. All of the case stories are new.

Chapter 2, "Foundations and Technologies for Decision Making," combines material from earlier Chapters 1, 2, and 3 to provide a basic foundation for decision making in general and computer-supported decision making in particular. It eliminates some duplication that was present in Chapters 1–3 of the previous editions. It includes 35 percent new material. Most of the cases are new.

Chapter 3, "Data Warehousing"
- 30 percent new material, including the cases
- New opening case
- Mostly new cases throughout
- NEW: A historic perspective to data warehousing—how did we get here?
- Better coverage of multidimensional modeling (star schema and snowflake schema)
- An updated coverage on the future of data warehousing

Chapter 4, "Business Reporting, Visual Analytics, and Business Performance Management"
- 60 percent of the material is new—especially in visual analytics and reporting
- Most of the cases are new

Chapter 5, "Data Mining"
- 25 percent of the material is new
- Most of the cases are new

Chapter 6, "Techniques for Predictive Modeling"
- 55 percent of the material is new
- Most of the cases are new
- New sections on SVM and *k*NN

Chapter 7, "Text Analytics, Text Mining, and Sentiment Analysis"
- 50 percent of the material is new
- Most of the cases are new
- New section (1/3 of the chapter) on sentiment analysis

Chapter 8, "Web Analytics, Web Mining, and Social Analytics" (New Chapter)

- 95 percent of the material is new

Chapter 9, "Model-Based Decision Making: Optimization and Multi-Criteria Systems"

- All new cases
- Expanded coverage of analytic hierarchy process
- New examples of mixed-integer programming applications and exercises
- About 50 percent new material

In addition, all the Microsoft Excel–related coverage has been updated to work with Microsoft Excel 2010.

Chapter 10, "Modeling and Analysis: Heuristic Search Methods and Simulation"

- This chapter now introduces genetic algorithms and various types of simulation models
- It includes new coverage of other types of simulation modeling such as agent-based modeling and system dynamics modeling
- New cases throughout
- About 60 percent new material

Chapter 11, "Automated Decision Systems and Expert Systems"

- Expanded coverage of automated decision systems including examples from the airline industry
- New examples of expert systems
- New cases
- About 50 percent new material

Chapter 12, "Knowledge Management and Collaborative Systems"

- Significantly condensed coverage of these two topics combined into one chapter
- New examples of KM applications
- About 25 percent new material

Chapters 13 and 14 are mostly new chapters, as described earlier.

We have retained many of the enhancements made in the last editions and updated the content. These are summarized next:

- ***Links to Teradata University Network (TUN).*** Most chapters include new links to TUN (**teradatauniversitynetwork.com**). We encourage the instructors to register and join **teradatauniversitynetwork.com** and explore various content available through the site. The cases, white papers, and software exercises available through TUN will keep your class fresh and timely.
- ***Book title.*** As is already evident, the book's title and focus have changed substantially.
- ***Software support.*** The TUN Web site provides software support at no charge. It also provides links to free data mining and other software. In addition, the site provides exercises in the use of such software.

THE SUPPLEMENT PACKAGE: PEARSONGLOBALEDITIONS.COM/SHARDA

A comprehensive and flexible technology-support package is available to enhance the teaching and learning experience. The following instructor and student supplements are available on the book's Web site, **pearsonglobaleditions.com/sharda**:

- ***Instructor's Manual.*** The Instructor's Manual includes learning objectives for the entire course and for each chapter, answers to the questions and exercises at the end of each chapter, and teaching suggestions (including instructions for projects). The Instructor's Manual is available on the secure faculty section of **pearsonglobaleditions.com/sharda**.

- ***Test Item File and TestGen Software.*** The Test Item File is a comprehensive collection of true/false, multiple-choice, fill-in-the-blank, and essay questions. The questions are rated by difficulty level, and the answers are referenced by book page number. The Test Item File is available in Microsoft Word and in TestGen. Pearson Education's test-generating software is available from **www.pearsonglobaleditions.com/irc**. The software is PC/MAC compatible and preloaded with all of the Test Item File questions. You can manually or randomly view test questions and drag-and-drop to create a test. You can add or modify test-bank questions as needed. Our TestGens are converted for use in BlackBoard, WebCT, Moodle, D2L, and Angel. These conversions can be found on **pearsonglobaleditions.com/sharda**. The TestGen is also available in Respondus and can be found on **www.respondus.com**.
- ***PowerPoint slides.*** PowerPoint slides are available that illuminate and build on key concepts in the text. Faculty can download the PowerPoint slides from **pearsonglobaleditions.com/sharda**.

ACKNOWLEDGMENTS

Many individuals have provided suggestions and criticisms since the publication of the first edition of this book. Dozens of students participated in class testing of various chapters, software, and problems and assisted in collecting material. It is not possible to name everyone who participated in this project, but our thanks go to all of them. Certain individuals made significant contributions, and they deserve special recognition.

First, we appreciate the efforts of those individuals who provided formal reviews of the first through tenth editions (school affiliations as of the date of review):

Robert Blanning, Vanderbilt University
Ranjit Bose, University of New Mexico
Warren Briggs, Suffolk University
Lee Roy Bronner, Morgan State University
Charles Butler, Colorado State University
Sohail S. Chaudry, University of Wisconsin–La Crosse
Kathy Chudoba, Florida State University
Wingyan Chung, University of Texas
Woo Young Chung, University of Memphis
Paul "Buddy" Clark, South Carolina State University
Pi'Sheng Deng, California State University–Stanislaus
Joyce Elam, Florida International University
Kurt Engemann, Iona College
Gary Farrar, Jacksonville University
George Federman, Santa Clara City College
Jerry Fjermestad, New Jersey Institute of Technology
Joey George, Florida State University
Paul Gray, Claremont Graduate School
Orv Greynholds, Capital College (Laurel, Maryland)
Martin Grossman, Bridgewater State College
Ray Jacobs, Ashland University
Leonard Jessup, Indiana University
Jeffrey Johnson, Utah State University
Jahangir Karimi, University of Colorado Denver
Saul Kassicieh, University of New Mexico
Anand S. Kunnathur, University of Toledo

Shao-ju Lee, California State University at Northridge
Yair Levy, Nova Southeastern University
Hank Lucas, New York University
Jane Mackay, Texas Christian University
George M. Marakas, University of Maryland
Dick Mason, Southern Methodist University
Nick McGaughey, San Jose State University
Ido Millet, Pennsylvania State University–Erie
Benjamin Mittman, Northwestern University
Larry Moore, Virginia Polytechnic Institute and State University
Simitra Mukherjee, Nova Southeastern University
Marianne Murphy, Northeastern University
Peter Mykytyn, Southern Illinois University
Natalie Nazarenko, SUNY College at Fredonia
Souren Paul, Southern Illinois University
Joshua Pauli, Dakota State University
Roger Alan Pick, University of Missouri–St. Louis
W. "RP" Raghupaphi, California State University–Chico
Loren Rees, Virginia Polytechnic Institute and State University
David Russell, Western New England College
Steve Ruth, George Mason University
Vartan Safarian, Winona State University
Glenn Shephard, San Jose State University
Jung P. Shim, Mississippi State University
Meenu Singh, Murray State University
Randy Smith, University of Virginia
James T.C. Teng, University of South Carolina
John VanGigch, California State University at Sacramento
David Van Over, University of Idaho
Paul J.A. van Vliet, University of Nebraska at Omaha
B. S. Vijayaraman, University of Akron
Howard Charles Walton, Gettysburg College
Diane B. Walz, University of Texas at San Antonio
Paul R. Watkins, University of Southern California
Randy S. Weinberg, Saint Cloud State University
Jennifer Williams, University of Southern Indiana
Steve Zanakis, Florida International University
Fan Zhao, Florida Gulf Coast University

Pearson would like to thank and acknowledge the following people for their work on the Global Edition:

Contributors
M. Reza Abdi, Bradford University School of Management
Stefania Paladini, Coventry University
Xavier Pierron, Coventry University
Krish Saha, Coventry University
Mark Sallos, Coventry University
Reviewers
Chun Kit Lok, University of Hong Kong
Liu Qizhang, National University of Singapore
May-Lin Yap, Universiti Teknoligi MARA

Several individuals contributed material to the text or the supporting material. Susan Baxley and Dr. David Schrader of Teradata provided special help in identifying new TUN content for the book and arranging permissions for the same. Peter Horner, editor of *OR/MS Today,* allowed us to summarize new application stories from *OR/MS Today* and *Analytics Magazine*. We also thank INFORMS for their permission to highlight content from *Interfaces*. Prof. Rick Wilson contributed some examples and exercise questions for Chapter 9. Assistance from Natraj Ponna, Daniel Asamoah, Amir Hassan-Zadeh, Kartik Dasika, Clara Gregory, and Amy Wallace (all of Oklahoma State University) is gratefully acknowledged for this edition. We also acknowledge Narges Kasiri (Ithaca College) for the write-up on system dynamics modeling and Jongswas Chongwatpol (NIDA, Thailand) for the material on SIMIO software. For the previous edition, we acknowledge the contributions of Dave King (JDA Software Group, Inc.) and Jerry Wagner (University of Nebraska–Omaha). Major contributors for earlier editions include Mike Goul (Arizona State University) and Leila A. Halawi (Bethune-Cookman College), who provided material for the chapter on data warehousing; Christy Cheung (Hong Kong Baptist University), who contributed to the chapter on knowledge management; Linda Lai (Macau Polytechnic University of China); Dave King (JDA Software Group, Inc.); Lou Frenzel, an independent consultant whose books *Crash Course in Artificial Intelligence and Expert Systems* and *Understanding of Expert Systems* (both published by Howard W. Sams, New York, 1987) provided material for the early editions; Larry Medsker (American University), who contributed substantial material on neural networks; and Richard V. McCarthy (Quinnipiac University), who performed major revisions in the seventh edition.

Previous editions of the book have also benefited greatly from the efforts of many individuals who contributed advice and interesting material (such as problems), gave feedback on material, or helped with class testing. These individuals are Warren Briggs (Suffolk University), Frank DeBalough (University of Southern California), Mei-Ting Cheung (University of Hong Kong), Alan Dennis (Indiana University), George Easton (San Diego State University), Janet Fisher (California State University, Los Angeles), David Friend (Pilot Software, Inc.), the late Paul Gray (Claremont Graduate School), Mike Henry (OSU), Dustin Huntington (Exsys, Inc.), Subramanian Rama Iyer (Oklahoma State University), Angie Jungermann (Oklahoma State University), Elena Karahanna (The University of Georgia), Mike McAulliffe (The University of Georgia), Chad Peterson (The University of Georgia), Neil Rabjohn (York University), Jim Ragusa (University of Central Florida), Alan Rowe (University of Southern California), Steve Ruth (George Mason University), Linus Schrage (University of Chicago), Antonie Stam (University of Missouri), Ron Swift (NCR Corp.), Merril Warkentin (then at Northeastern University), Paul Watkins (The University of Southern California), Ben Mortagy (Claremont Graduate School of Management), Dan Walsh (Bellcore), Richard Watson (The University of Georgia), and the many other instructors and students who have provided feedback.

Several vendors cooperated by providing development and/or demonstration software: Expert Choice, Inc. (Pittsburgh, Pennsylvania), Nancy Clark of Exsys, Inc. (Albuquerque, New Mexico), Jim Godsey of GroupSystems, Inc. (Broomfield, Colorado), Raimo Hämäläinen of Helsinki University of Technology, Gregory Piatetsky-Shapiro of KDNuggets.com, Logic Programming Associates (UK), Gary Lynn of NeuroDimension Inc. (Gainesville, Florida), Palisade Software (Newfield, New York), Jerry Wagner of Planners Lab (Omaha, Nebraska), Promised Land Technologies (New Haven, Connecticut), Salford Systems (La Jolla, California), Sense Networks (New York, New York), Gary Miner of StatSoft, Inc. (Tulsa, Oklahoma), Ward Systems Group, Inc. (Frederick, Maryland), Idea Fisher Systems, Inc. (Irving, California), and Wordtech Systems (Orinda, California).

Special thanks to the Teradata University Network and especially to Hugh Watson, Michael Goul, and Susan Baxley, Program Director, for their encouragement to tie this book with TUN and for providing useful material for the book.

Many individuals helped us with administrative matters and editing, proofreading, and preparation. The project began with Jack Repcheck (a former Macmillan editor), who initiated this project with the support of Hank Lucas (New York University). Judy Lang collaborated with all of us, provided editing, and guided us during the entire project through the eighth edition.

Finally, the Pearson team is to be commended: Executive Editor Bob Horan, who orchestrated this project; Kitty Jarrett, who copyedited the manuscript; and the production team, Tom Benfatti at Pearson, George and staff at Integra Software Services, who transformed the manuscript into a book.

We would like to thank all these individuals and corporations. Without their help, the creation of this book would not have been possible. Ramesh and Dursun want to specifically acknowledge the contributions of previous coauthors Janine Aronson, David King, and T. P. Liang, whose original contributions constitute significant components of the book.

R.S.

D.D.

E.T.

Note that Web site URLs are dynamic. As this book went to press, we verified that all the cited Web sites were active and valid. Web sites to which we refer in the text sometimes change or are discontinued because companies change names, are bought or sold, merge, or fail. Sometimes Web sites are down for maintenance, repair, or redesign. Most organizations have dropped the initial "www" designation for their sites, but some still use it. If you have a problem connecting to a Web site that we mention, please be patient and simply run a Web search to try to identify the new site. Most times, the new site can be found quickly. Some sites also require a free registration before allowing you to see the content. We apologize in advance for this inconvenience.

ABOUT THE AUTHORS

Ramesh Sharda (M.B.A., Ph.D., University of Wisconsin–Madison) is director of the Ph.D. in Business for Executives Program and Institute for Research in Information Systems (IRIS), ConocoPhillips Chair of Management of Technology, and a Regents Professor of Management Science and Information Systems in the Spears School of Business at Oklahoma State University (OSU). About 200 papers describing his research have been published in major journals, including *Operations Research*, *Management Science*, *Information Systems Research*, *Decision Support Systems*, and *Journal of MIS*. He cofounded the AIS SIG on Decision Support Systems and Knowledge Management (SIGDSS). Dr. Sharda serves on several editorial boards, including those of *INFORMS Journal on Computing*, *Decision Support Systems*, and *ACM Transactions on Management Information Systems*. He has authored and edited several textbooks and research books and serves as the co-editor of several book series (Integrated Series in Information Systems, Operations Research/Computer Science Interfaces, and Annals of Information Systems) with Springer. He is also currently serving as the executive director of the Teradata University Network. His current research interests are in decision support systems, business analytics, and technologies for managing information overload.

Dursun Delen (Ph.D., Oklahoma State University) is the Spears and Patterson Chairs in Business Analytics, Director of Research for the Center for Health Systems Innovation, and Professor of Management Science and Information Systems in the Spears School of Business at Oklahoma State University (OSU). Prior to his academic career, he worked for a privately owned research and consultancy company, Knowledge Based Systems Inc., in College Station, Texas, as a research scientist for five years, during which he led a number of decision support and other information systems–related research projects funded by federal agencies such as DoD, NASA, NIST, and DOE. Dr. Delen's research has appeared in major journals including *Decision Support Systems, Communications of the ACM, Computers and Operations Research, Computers in Industry, Journal of Production Operations Management, Artificial Intelligence in Medicine,* and *Expert Systems with Applications*, among others. He recently published four textbooks: *Advanced Data Mining Techniques* with Springer, 2008; *Decision Support and Business Intelligence Systems* with Prentice Hall, 2010; *Business Intelligence: A Managerial Approach*, with Prentice Hall, 2010; and *Practical Text Mining*, with Elsevier, 2012. He is often invited to national and international conferences for keynote addresses on topics related to data/text mining, business intelligence, decision support systems, and knowledge management. He served as the general co-chair for the 4th International Conference on Network Computing and Advanced Information Management (September 2–4, 2008, in Seoul, South Korea) and regularly chairs tracks and mini-tracks at various information systems conferences. He is the associate editor-in-chief for *International Journal of Experimental Algorithms*, associate editor for *International Journal of RF Technologies* and *Journal of Decision Analytics*, and is on the editorial boards of five other technical journals. His research and teaching interests are in data and text mining, decision support systems, knowledge management, business intelligence, and enterprise modeling.

Efraim Turban (M.B.A., Ph.D., University of California, Berkeley) is a visiting scholar at the Pacific Institute for Information System Management, University of Hawaii. Prior to this, he was on the staff of several universities, including City University of Hong Kong; Lehigh University; Florida International University; California State University, Long

Beach; Eastern Illinois University; and the University of Southern California. Dr. Turban is the author of more than 100 refereed papers published in leading journals, such as *Management Science*, *MIS Quarterly*, and *Decision Support Systems*. He is also the author of 20 books, including *Electronic Commerce: A Managerial Perspective* and *Information Technology for Management*. He is also a consultant to major corporations worldwide. Dr. Turban's current areas of interest are Web-based decision support systems, social commerce, and collaborative decision making.

PART

I

Decision Making and Analytics

An Overview

LEARNING OBJECTIVES FOR PART I

- Understand the need for business analytics
- Understand the foundations and key issues of managerial decision making
- Understand the major categories and applications of business analytics
- Learn the major frameworks of computerized decision support: analytics, decision support systems (DSS), and business intelligence (BI)

This book deals with a collection of computer technologies that support managerial work—essentially, decision making. These technologies have had a profound impact on corporate strategy, performance, and competitiveness. These techniques broadly encompass analytics, business intelligence, and decision support systems, as shown throughout the book. In Part I, we first provide an overview of the whole book in one chapter. We cover several topics in this chapter. The first topic is managerial decision making and its computerized support; the second is frameworks for decision support. We then introduce business analytics and business intelligence. We also provide examples of applications of these analytical techniques, as well as a preview of the entire book. The second chapter within Part I introduces the foundational methods for decision making and relates these to computerized decision support. It also covers the components and technologies of decision support systems.

CHAPTER

1

An Overview of Business Intelligence, Analytics, and Decision Support

LEARNING OBJECTIVES

- Understand today's turbulent business environment and describe how organizations survive and even excel in such an environment (solving problems and exploiting opportunities)
- Understand the need for computerized support of managerial decision making
- Understand an early framework for managerial decision making
- Learn the conceptual foundations of the decision support systems (DSS[1]) methodology
- Describe the business intelligence (BI) methodology and concepts and relate them to DSS
- Understand the various types of analytics
- List the major tools of computerized decision support

The business environment (climate) is constantly changing, and it is becoming more and more complex. Organizations, private and public, are under pressures that force them to respond quickly to changing conditions and to be innovative in the way they operate. Such activities require organizations to be agile and to make frequent and quick strategic, tactical, and operational decisions, some of which are very complex. Making such decisions may require considerable amounts of relevant data, information, and knowledge. Processing these, in the framework of the needed decisions, must be done quickly, frequently in real time, and usually requires some computerized support.

This book is about using business analytics as computerized support for managerial decision making. It concentrates on both the theoretical and conceptual foundations of decision support, as well as on the commercial tools and techniques that are available. This introductory chapter provides more details of these topics as well as an overview of the book. This chapter has the following sections:

[1]The acronym *DSS* is treated as both singular and plural throughout this book. Similarly, other acronyms, such as *MIS* and *GSS*, designate both plural and singular forms. This is also true of the word *analytics*.

1.1 OPENING VIGNETTE: Magpie Sensing Employs Analytics to Manage a Vaccine Supply Chain Effectively and Safely

Cold chain in healthcare is defined as the temperature-controlled supply chain involving a system of transporting and storing vaccines and pharmaceutical drugs. It consists of three major components—transport and storage equipment, trained personnel, and efficient management procedures. The majority of the vaccines in the cold chain are typically maintained at a temperature of 35–46 degrees Fahrenheit [2–8 degrees Centigrade]. Maintaining cold chain integrity is extremely important for healthcare product manufacturers.

Especially for the vaccines, improper storage and handling practices that compromise vaccine viability prove a costly, time-consuming affair. Vaccines must be stored properly from manufacture until they are available for use. Any extreme temperatures of heat or cold will reduce vaccine potency; such vaccines, if administered, might not yield effective results or could cause adverse effects.

Effectively maintaining the temperatures of storage units throughout the healthcare supply chain in real time—i.e., beginning from the gathering of the resources, manufacturing, distribution, and dispensing of the products—is the most effective solution desired in the cold chain. Also, the location-tagged real-time environmental data about the storage units helps in monitoring the cold chain for spoiled products. The chain of custody can be easily identified to assign product liability.

A study conducted by the Centers for Disease Control and Prevention (CDC) looked at the handling of cold chain vaccines by 45 healthcare providers around United States and reported that three-quarters of the providers experienced serious cold chain violations.

A WAY TOWARD A POSSIBLE SOLUTION

Magpie Sensing, a start-up project under Ebers Smith and Douglas Associated LLC, provides a suite of cold chain monitoring and analysis technologies for the healthcare industry. It is a shippable, wireless temperature and humidity monitor that provides real-time, location-aware tracking of cold chain products during shipment. Magpie Sensing's solutions rely on rich analytics algorithms that leverage the data gathered from the monitoring devices to improve the efficiency of cold chain processes and predict cold storage problems before they occur.

Magpie sensing applies all three types of analytical techniques—descriptive, predictive, and prescriptive analytics—to turn the raw data returned from the monitoring devices into actionable recommendations and warnings.

The properties of the cold storage system, which include the set point of the storage system's thermostat, the typical range of temperature values in the storage system, and the duty cycle of the system's compressor, are monitored and reported in real time. This information helps trained personnel to ensure that the storage unit is properly configured to store a particular product. All the temperature information is displayed on a Web dashboard that shows a graph of the temperature inside the specific storage unit.

Based on information derived from the monitoring devices, Magpie's predictive analytic algorithms can determine the set point of the storage unit's thermostat and alert the system's users if the system is incorrectly configured, depending upon the various types of products stored. This offers a solution to the users of consumer refrigerators where the thermostat is not temperature graded. Magpie's system also sends alerts about possible temperature violations based on the storage unit's average temperature and subsequent compressor cycle runs, which may drop the temperature below the freezing point. Magpie's predictive analytics further report possible human errors, such as failure to shut the storage unit doors or the presence of an incomplete seal, by analyzing the temperature trend and alerting users via Web interface, text message, or audible alert before the temperature bounds are actually violated. In a similar way, a compressor or a power failure can be detected; the estimated time before the storage unit reaches an unsafe temperature also is reported, which prepares the users to look for backup solutions such as using dry ice to restore power.

In addition to predictive analytics, Magpie Sensing's analytics systems can provide prescriptive recommendations for improving the cold storage processes and business decision making. Prescriptive analytics help users dial in the optimal temperature setting, which helps to achieve the right balance between freezing and spoilage risk; this, in turn, provides a cushion-time to react to the situation before the products spoil. Its prescriptive analytics also gather useful meta-information on cold storage units, including the times of day that are busiest and periods where the system's doors are opened, which can be used to provide additional design plans and institutional policies that ensure that the system is being properly maintained and not overused.

Furthermore, prescriptive analytics can be used to guide equipment purchase decisions by constantly analyzing the performance of current storage units. Based on the storage system's efficiency, decisions on distributing the products across available storage units can be made based on the product's sensitivity.

Using Magpie Sensing's cold chain analytics, additional manufacturing time and expenditure can be eliminated by ensuring that product safety can be secured throughout the supply chain and effective products can be administered to the patients. Compliance with state and federal safety regulations can be better achieved through automatic data gathering and reporting about the products involved in the cold chain.

QUESTIONS FOR THE OPENING VIGNETTE

1. What information is provided by the descriptive analytics employed at Magpie Sensing?
2. What type of support is provided by the predictive analytics employed at Magpie Sensing?
3. How does prescriptive analytics help in business decision making?
4. In what ways can actionable information be reported in real time to concerned users of the system?
5. In what other situations might real-time monitoring applications be needed?

WHAT WE CAN LEARN FROM THIS VIGNETTE

This vignette illustrates how data from a business process can be used to generate insights at various levels. First, the graphical analysis of the data (termed *reporting analytics*) allows users to get a good feel for the situation. Then, additional analysis using **data mining** techniques can be used to estimate what future behavior would be like. This is the domain of predictive analytics. Such analysis can then be taken to create specific recommendations for operators. This is an example of what we call prescriptive analytics. Finally, this opening vignette also suggests that innovative applications of analytics can create new business ventures. Identifying opportunities for applications of analytics and assisting with decision making in specific domains is an emerging entrepreneurial opportunity.

Sources: Magpiesensing.com, "Magpie Sensing Cold Chain Analytics and Monitoring," **magpiesensing.com/wp-content/uploads/2013/01/ColdChainAnalyticsMagpieSensing-Whitepaper.pdf** (accessed July 2013); Centers for Disease Control and Prevention, Vaccine Storage and Handling, **http://www.cdc.gov/vaccines/pubs/pinkbook/vac-storage.html#storage** (accessed July 2013); A. Zaleski, "Magpie Analytics System Tracks Cold-Chain Products to Keep Vaccines, Reagents Fresh" (2012), **technicallybaltimore.com/profiles/startups/magpie-analytics-system-tracks-cold-chain-products-to-keep-vaccines-reagents-fresh** (accessed February 2013).

1.2 CHANGING BUSINESS ENVIRONMENTS AND COMPUTERIZED DECISION SUPPORT

The opening vignette illustrates how a company can employ technologies to make sense of data and make better decisions. Companies are moving aggressively to computerized support of their operations. To understand why companies are embracing computerized support, including business intelligence, we developed a model called the *Business Pressures–Responses–Support Model*, which is shown in Figure 1.1.

The Business Pressures–Responses–Support Model

The Business Pressures–Responses–Support Model, as its name indicates, has three components: business pressures that result from today's business climate, responses (actions taken) by companies to counter the pressures (or to take advantage of the opportunities available in the environment), and computerized support that facilitates the monitoring of the environment and enhances the response actions taken by organizations.

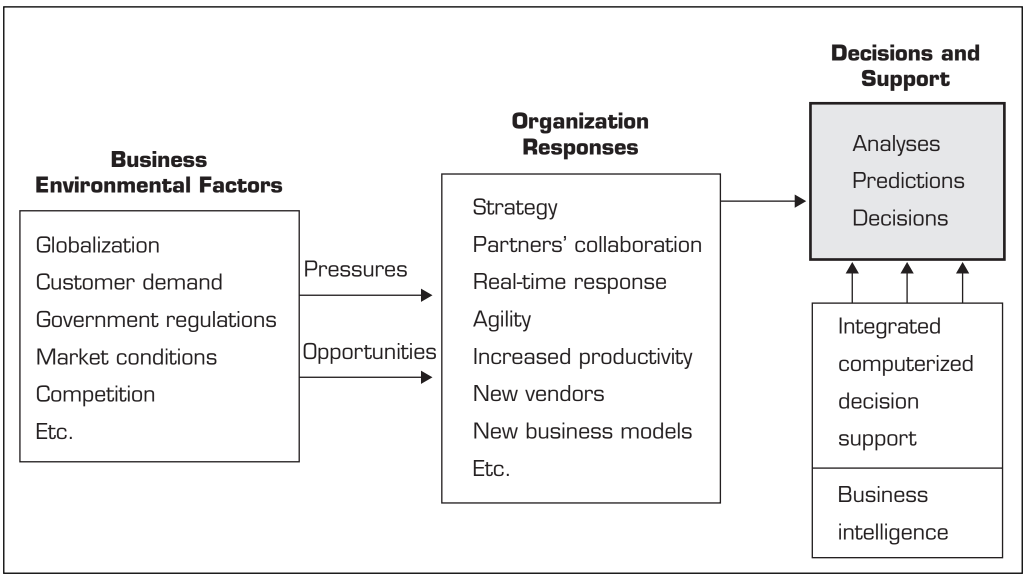

FIGURE 1.1 The Business Pressures–Responses–Support Model.

THE BUSINESS ENVIRONMENT The environment in which organizations operate today is becoming more and more complex. This complexity creates opportunities on the one hand and problems on the other. Take globalization as an example. Today, you can easily find suppliers and customers in many countries, which means you can buy cheaper materials and sell more of your products and services; great opportunities exist. However, globalization also means more and stronger competitors. Business environment factors can be divided into four major categories: *markets, consumer demands, technology*, and *societal*. These categories are summarized in Table 1.1.

Note that the *intensity* of most of these factors increases with time, leading to more pressures, more competition, and so on. In addition, organizations and departments within organizations face decreased budgets and amplified pressures from top managers to increase performance and profit. In this kind of environment, managers must respond quickly, innovate, and be agile. Let's see how they do it.

ORGANIZATIONAL RESPONSES: BE REACTIVE, ANTICIPATIVE, ADAPTIVE, AND PROACTIVE Both private and public organizations are aware of today's business environment and pressures. They use different actions to counter the pressures. Vodafone New Zealand Ltd (Krivda, 2008), for example, turned to BI to improve communication and to support executives in its effort to retain existing customers and increase revenue from these customers. Managers may take other actions, including the following:

- Employ strategic planning.
- Use new and innovative business models.
- Restructure business processes.
- Participate in business alliances.
- Improve corporate information systems.
- Improve partnership relationships.

TABLE 1.1 Business Environment Factors That Create Pressures on Organizations

Factor	Description
Markets	Strong competition
	Expanding global markets
	Booming electronic markets on the Internet
	Innovative marketing methods
	Opportunities for outsourcing with IT support
	Need for real-time, on-demand transactions
Consumer demands	Desire for customization
	Desire for quality, diversity of products, and speed of delivery
	Customers getting powerful and less loyal
Technology	More innovations, new products, and new services
	Increasing obsolescence rate
	Increasing information overload
	Social networking, Web 2.0 and beyond
Societal	Growing government regulations and deregulation
	Workforce more diversified, older, and composed of more women
	Prime concerns of homeland security and terrorist attacks
	Necessity of Sarbanes-Oxley Act and other reporting-related legislation
	Increasing social responsibility of companies
	Greater emphasis on sustainability

- Encourage innovation and creativity.
- Improve customer service and relationships.
- Employ social media and mobile platforms for e-commerce and beyond.
- Move to make-to-order production and on-demand manufacturing and services.
- Use new IT to improve communication, data access (discovery of information), and collaboration.
- Respond quickly to competitors' actions (e.g., in pricing, promotions, new products and services).
- Automate many tasks of white-collar employees.
- Automate certain decision processes, especially those dealing with customers.
- Improve decision making by employing analytics.

Many, if not all, of these actions require some computerized support. These and other response actions are frequently facilitated by computerized decision support (DSS).

CLOSING THE STRATEGY GAP One of the major objectives of computerized decision support is to facilitate closing the gap between the current performance of an organization and its desired performance, as expressed in its mission, objectives, and goals, and the strategy to achieve them. In order to understand why computerized support is needed and how it is provided, especially for decision-making support, let's look at managerial decision making.

SECTION 1.2 REVIEW QUESTIONS

1. List the components of and explain the Business Pressures–Responses–Support Model.
2. What are some of the major factors in today's business environment?
3. What are some of the major response activities that organizations take?

1.3 MANAGERIAL DECISION MAKING

Management is a process by which organizational goals are achieved by using resources. The resources are considered inputs, and attainment of goals is viewed as the output of the process. The degree of success of the organization and the manager is often measured by the ratio of outputs to inputs. This ratio is an indication of the organization's *productivity*, which is a reflection of the *organizational and managerial performance*.

The level of productivity or the success of management depends on the performance of managerial functions, such as planning, organizing, directing, and controlling. To perform their functions, managers engage in a continuous process of making decisions. Making a decision means selecting the best alternative from two or more solutions.

The Nature of Managers' Work

Mintzberg's (2008) classic study of top managers and several replicated studies suggest that managers perform 10 major roles that can be classified into three major categories: *interpersonal, informational*, and *decisional* (see Table 1.2).

To perform these roles, managers need information that is delivered efficiently and in a timely manner to personal computers (PCs) on their desktops and to mobile devices. This information is delivered by networks, generally via Web technologies.

In addition to obtaining information necessary to better perform their roles, managers use computers directly to support and improve decision making, which is a key task

TABLE 1.2 Mintzberg's 10 Managerial Roles

Role	Description
Interpersonal	
Figurehead	Is symbolic head; obliged to perform a number of routine duties of a legal or social nature
Leader	Is responsible for the motivation and activation of subordinates; responsible for staffing, training, and associated duties
Liaison	Maintains self-developed network of outside contacts and informers who provide favors and information
Informational	
Monitor	Seeks and receives a wide variety of special information (much of it current) to develop a thorough understanding of the organization and environment; emerges as the nerve center of the organization's internal and external information
Disseminator	Transmits information received from outsiders or from subordinates to members of the organization; some of this information is factual, and some involves interpretation and integration
Spokesperson	Transmits information to outsiders about the organization's plans, policies, actions, results, and so forth; serves as an expert on the organization's industry
Decisional	
Entrepreneur	Searches the organization and its environment for opportunities and initiates improvement projects to bring about change; supervises design of certain projects
Disturbance handler	Is responsible for corrective action when the organization faces important, unexpected disturbances
Resource allocator	Is responsible for the allocation of organizational resources of all kinds; in effect, is responsible for the making or approval of all significant organizational decisions
Negotiator	Is responsible for representing the organization at major negotiations

Sources: Compiled from H. A. Mintzberg, *The Nature of Managerial Work*. Prentice Hall, Englewood Cliffs, NJ, 1980; and H. A. Mintzberg, *The Rise and Fall of Strategic Planning*. The Free Press, New York, 1993.

that is part of most of these roles. Many managerial activities in all roles revolve around decision making. *Managers, especially those at high managerial levels, are primarily decision makers*. We review the decision-making process next but will study it in more detail in the next chapter.

The Decision-Making Process

For years, managers considered decision making purely an art—a talent acquired over a long period through experience (i.e., learning by trial-and-error) and by using intuition. Management was considered an art because a variety of individual styles could be used in approaching and successfully solving the same types of managerial problems. These styles were often based on creativity, judgment, intuition, and experience rather than on systematic quantitative methods grounded in a scientific approach. However, recent research suggests that companies with top managers who are more focused on persistent work (almost dullness) tend to outperform those with leaders whose main strengths are interpersonal communication skills (Kaplan et al., 2008; Brooks, 2009). It is more important to emphasize methodical, thoughtful, analytical decision making rather than flashiness and interpersonal communication skills.

Managers usually make decisions by following a four-step process (we learn more about these in Chapter 2):

1. Define the problem (i.e., a decision situation that may deal with some difficulty or with an opportunity).
2. Construct a model that describes the real-world problem.
3. Identify possible solutions to the modeled problem and evaluate the solutions.
4. Compare, choose, and recommend a potential solution to the problem.

To follow this process, one must make sure that sufficient alternative solutions are being considered, that the consequences of using these alternatives can be reasonably predicted, and that comparisons are done properly. However, the environmental factors listed in Table 1.1 make such an evaluation process difficult for the following reasons:

- Technology, information systems, advanced search engines, and globalization result in more and more alternatives from which to choose.
- Government regulations and the need for compliance, political instability and terrorism, competition, and changing consumer demands produce more uncertainty, making it more difficult to predict consequences and the future.
- Other factors are the need to make rapid decisions, the frequent and unpredictable changes that make trial-and-error learning difficult, and the potential costs of making mistakes.
- These environments are growing more complex every day. Therefore, making decisions today is indeed a complex task.

Because of these trends and changes, it is nearly impossible to rely on a trial-and-error approach to management, especially for decisions for which the factors shown in Table 1.1 are strong influences. Managers must be more sophisticated; they must use the new tools and techniques of their fields. Most of those tools and techniques are discussed in this book. Using them to support decision making can be extremely rewarding in making effective decisions. In the following section, we look at why we need computer support and how it is provided.

SECTION 1.3 REVIEW QUESTIONS

1. Describe the three major managerial roles, and list some of the specific activities in each.
2. Why have some argued that management is the same as decision making?
3. Describe the four steps managers take in making a decision.

1.4 INFORMATION SYSTEMS SUPPORT FOR DECISION MAKING

From traditional uses in payroll and bookkeeping functions, computerized systems have penetrated complex managerial areas ranging from the design and management of automated factories to the application of analytical methods for the evaluation of proposed mergers and acquisitions. Nearly all executives know that information technology is vital to their business and extensively use information technologies.

Computer applications have moved from transaction processing and monitoring activities to problem analysis and solution applications, and much of the activity is done with Web-based technologies, in many cases accessed through mobile devices. Analytics and BI tools such as data warehousing, data mining, online analytical processing (OLAP), **dashboards**, and the use of the Web for decision support are the cornerstones of today's modern management. Managers must have high-speed, networked information systems (wireline or wireless) to assist them with their most important task: making decisions. Besides the obvious growth in hardware, software, and network capacities, some developments have clearly contributed to facilitating growth of decision support and analytics in a number of ways, including the following:

- ***Group communication and collaboration.*** Many decisions are made today by groups whose members may be in different locations. Groups can collaborate and communicate readily by using Web-based tools as well as the ubiquitous smartphones. Collaboration is especially important along the supply chain, where partners—all the way from vendors to customers—must share information. Assembling a group of decision makers, especially experts, in one place can be costly. Information systems can improve the collaboration process of a group and enable its members to be at different locations (saving travel costs). We will study some applications in Chapter 12.
- ***Improved data management.*** Many decisions involve complex computations. Data for these can be stored in different databases anywhere in the organization and even possibly at Web sites outside the organization. The data may include text, sound, graphics, and video, and they can be in different languages. It may be necessary to transmit data quickly from distant locations. Systems today can search, store, and transmit needed data quickly, economically, securely, and transparently.
- ***Managing giant data warehouses and Big Data.*** Large data warehouses, like the ones operated by Walmart, contain terabytes and even petabytes of data. Special methods, including parallel computing, are available to organize, search, and mine the data. The costs related to data warehousing are declining. Technologies that fall under the broad category of Big Data have enabled massive data coming from a variety of sources and in many different forms, which allows a very different view into organizational performance that was not possible in the past.
- ***Analytical support.*** With more data and analysis technologies, more alternatives can be evaluated, forecasts can be improved, risk analysis can be performed quickly, and the views of experts (some of whom may be in remote locations) can be collected quickly and at a reduced cost. Expertise can even be derived directly from analytical systems. With such tools, decision makers can perform complex simulations, check many possible scenarios, and assess diverse impacts quickly and economically. This, of course, is the focus of several chapters in the book.
- ***Overcoming cognitive limits in processing and storing information.*** According to Simon (1977), the human mind has only a limited ability to process and store information. People sometimes find it difficult to recall and use information in an error-free fashion due to their cognitive limits. The term *cognitive limits* indicates that an individual's problem-solving capability is limited when a wide range of diverse information and knowledge is required. Computerized systems enable people to overcome their cognitive limits by quickly accessing and processing vast amounts of stored information (see Chapter 2).
- ***Knowledge management.*** Organizations have gathered vast stores of information about their own operations, customers, internal procedures, employee interactions, and so forth through the unstructured and structured communications taking place among the various stakeholders. Knowledge management systems (KMS, Chapter 12) have become sources of formal and informal support for decision making to managers, although sometimes they may not even be called *KMS*.
- ***Anywhere, any time support.*** Using wireless technology, managers can access information anytime and from any place, analyze and interpret it, and communicate with those involved. This perhaps is the biggest change that has occurred in the last few years. The speed at which information needs to be processed and converted into decisions has truly changed expectations for both consumers and businesses.

These and other capabilities have been driving the use of computerized decision support since the late 1960s, but especially since the mid-1990s. The growth of mobile technologies, social media platforms, and analytical tools has enabled a much higher level of information systems support for managers. In the next sections we study a historical classification of decision support tasks. This leads us to be introduced to decision support systems. We will then study an overview of technologies that have been broadly referred to as business intelligence. From there we will broaden our horizons to introduce various types of analytics.

SECTION 1.4 REVIEW QUESTIONS

1. What are some of the key system-oriented trends that have fostered IS-supported decision making to a new level?
2. List some capabilities of information systems that can facilitate managerial decision making.
3. How can a computer help overcome the cognitive limits of humans?

1.5 AN EARLY FRAMEWORK FOR COMPUTERIZED DECISION SUPPORT

An early framework for computerized decision support includes several major concepts that are used in forthcoming sections and chapters of this book. Gorry and Scott-Morton created and used this framework in the early 1970s, and the framework then evolved into a new technology called *DSS*.

The Gorry and Scott-Morton Classical Framework

Gorry and Scott-Morton (1971) proposed a framework that is a 3-by-3 matrix, as shown in Figure 1.2. The two dimensions are the degree of structuredness and the types of control.

	Type of Control		
Type of Decision	**Operational Control**	**Managerial Control**	**Strategic Planning**
Structured	1 Accounts receivable Accounts payable Order entry	2 Budget analysis Short-term forecasting Personnel reports Make-or-buy	3 Financial management Investment portfolio Warehouse location Distribution systems
Semistructured	4 Production scheduling Inventory control	5 Credit evaluation Budget preparation Plant layout Project scheduling Reward system design Inventory categorization	6 Building a new plant Mergers & acquisitions New product planning Compensation planning Quality assurance HR policies Inventory planning
Unstructured	7 Buying software Approving loans Operating a help desk Selecting a cover for a magazine	8 Negotiating Recruiting an executive Buying hardware Lobbying	9 R & D planning New tech development Social responsibility planning

FIGURE 1.2 Decision Support Frameworks.

DEGREE OF STRUCTUREDNESS The left side of Figure 1.2 is based on Simon's (1977) idea that decision-making processes fall along a continuum that ranges from highly structured (sometimes called *programmed*) to highly unstructured (i.e., *nonprogrammed*) decisions. Structured processes are routine and typically repetitive problems for which standard solution methods exist. *Unstructured processes* are fuzzy, complex problems for which there are no cut-and-dried solution methods.

An **unstructured problem** is one where the articulation of the problem or the solution approach may be unstructured in itself. In a **structured problem**, the procedures for obtaining the best (or at least a good enough) solution are known. Whether the problem involves finding an appropriate inventory level or choosing an optimal investment strategy, the objectives are clearly defined. Common objectives are cost minimization and profit maximization.

Semistructured problems fall between structured and unstructured problems, having some structured elements and some unstructured elements. Keen and Scott-Morton (1978) mentioned trading bonds, setting marketing budgets for consumer products, and performing capital acquisition analysis as semistructured problems.

TYPES OF CONTROL The second half of the Gorry and Scott-Morton framework (refer to Figure 1.2) is based on Anthony's (1965) taxonomy, which defines three broad categories that encompass all managerial activities: *strategic planning*, which involves defining long-range goals and policies for resource allocation; *management control*, the acquisition and efficient use of resources in the accomplishment of organizational goals; and *operational control*, the efficient and effective execution of specific tasks.

THE DECISION SUPPORT MATRIX Anthony's and Simon's taxonomies are combined in the nine-cell decision support matrix shown in Figure 1.2. The initial purpose of this matrix was to suggest different types of computerized support to different cells in the matrix. Gorry and Scott-Morton suggested, for example, that for *semistructured decisions* and *unstructured decisions*, conventional management information systems (MIS) and management science (MS) tools are insufficient. Human intellect and a different approach to computer technologies are necessary. They proposed the use of a supportive information system, which they called a *DSS*.

Note that the more structured and operational control-oriented tasks (such as those in cells 1, 2, and 4) are usually performed by lower-level managers, whereas the tasks in cells 6, 8, and 9 are the responsibility of top executives or highly trained specialists.

Computer Support for Structured Decisions

Computers have historically supported structured and some semistructured decisions, especially those that involve operational and managerial control, since the 1960s. Operational and managerial control decisions are made in all functional areas, especially in finance and production (i.e., operations) management.

Structured problems, which are encountered repeatedly, have a high level of structure. It is therefore possible to abstract, analyze, and classify them into specific categories. For example, a make-or-buy decision is one category. Other examples of categories are capital budgeting, allocation of resources, distribution, procurement, planning, and inventory control decisions. For each category of decision, an easy-to-apply prescribed model and solution approach have been developed, generally as quantitative formulas. Therefore, it is possible to use a *scientific approach* for automating portions of managerial decision making.

Computer Support for Unstructured Decisions

Unstructured problems can be only partially supported by standard computerized quantitative methods. It is usually necessary to develop customized solutions. However, such solutions may benefit from data and information generated from corporate or external data sources. Intuition and judgment may play a large role in these types of decisions, as may computerized communication and collaboration technologies, as well as knowledge management (see Chapter 12).

Computer Support for Semistructured Problems

Solving semistructured problems may involve a combination of standard solution procedures and human judgment. Management science can provide models for the portion of a decision-making problem that is structured. For the unstructured portion, a DSS can improve the quality of the information on which the decision is based by providing, for example, not only a single solution but also a range of alternative solutions, along with their potential impacts. These capabilities help managers to better understand the nature of problems and, thus, to make better decisions.

SECTION 1.5 REVIEW QUESTIONS

1. What are structured, unstructured, and semistructured decisions? Provide two examples of each.
2. Define *operational control, managerial control*, and *strategic planning*. Provide two examples of each.
3. What are the nine cells of the decision framework? Explain what each is for.
4. How can computers provide support for making structured decisions?
5. How can computers provide support to semistructured and unstructured decisions?

1.6 THE CONCEPT OF DECISION SUPPORT SYSTEMS (DSS)

In the early 1970s, Scott-Morton first articulated the major concepts of DSS. He defined **decision support systems (DSS)** as “interactive computer-based systems, which help decision makers utilize *data* and *models* to solve unstructured problems” (Gorry and Scott-Morton, 1971). The following is another classic DSS definition, provided by Keen and Scott-Morton (1978):

> Decision support systems couple the intellectual resources of individuals with the capabilities of the computer to improve the quality of decisions. It is a computer-based support system for management decision makers who deal with semistructured problems.

Note that the term *decision support system*, like *management information system* (MIS) and other terms in the field of IT, is a content-free expression (i.e., it means different things to different people). Therefore, there is no universally accepted definition of DSS. (We present additional definitions in Chapter 2.) Actually, DSS can be viewed as a *conceptual methodology*—that is, a broad, umbrella term. However, some view DSS as a narrower, specific decision support application.

DSS as an Umbrella Term

The term *DSS* can be used as an umbrella term to describe any computerized system that supports decision making in an organization. An organization may have a knowledge management system to guide all its personnel in their problem solving. Another organization may have separate support systems for marketing, finance, and accounting; a supply chain management (SCM) system for production; and several rule-based systems for product repair diagnostics and help desks. DSS encompasses them all.

Evolution of DSS into Business Intelligence

In the early days of DSS, managers let their staff do some supportive analysis by using DSS tools. As PC technology advanced, a new generation of managers evolved—one that was comfortable with computing and knew that technology can directly help make intelligent business decisions faster. New tools such as OLAP, data warehousing, data mining, and intelligent systems, delivered via Web technology, added promised capabilities and easy access to tools, models, and data for computer-aided decision making. These tools started to appear under the names *BI* and *business analytics* in the mid-1990s. We introduce these concepts next, and relate the DSS and BI concepts in the following sections.

SECTION 1.6 REVIEW QUESTIONS

1. Provide two definitions of *DSS*.
2. Describe *DSS* as an umbrella term.

1.7 A FRAMEWORK FOR BUSINESS INTELLIGENCE (BI)

The decision support concepts presented in Sections 1.5 and 1.6 have been implemented incrementally, under different names, by many vendors that have created tools and methodologies for decision support. As the enterprise-wide systems grew, managers were able to access user-friendly reports that enabled them to make decisions quickly. These systems, which were generally called *executive information systems* (EIS), then began to offer additional visualization, alerts, and performance measurement capabilities. By 2006, the major *commercial* products and services appeared under the umbrella term *business intelligence* (BI).

Definitions of BI

Business intelligence (BI) is an umbrella term that combines architectures, tools, databases, analytical tools, applications, and methodologies. It is, like DSS, a content-free expression, so it means different things to different people. Part of the confusion about BI lies in the flurry of acronyms and buzzwords that are associated with it (e.g., business performance management [BPM]). BI's major objective is to enable interactive access (sometimes in real time) to data, to enable manipulation of data, and to give business managers and analysts the ability to conduct appropriate analyses. By analyzing historical and current data, situations, and performances, decision makers get valuable insights that enable them to make more informed and better decisions. The process of BI is based on the *transformation* of data to information, then to decisions, and finally to actions.

A Brief History of BI

The term *BI* was coined by the Gartner Group in the mid-1990s. However, the concept is much older; it has its roots in the MIS reporting systems of the 1970s. During that period, reporting systems were static, two dimensional, and had no analytical capabilities. In the early 1980s, the concept of *executive information systems* (EIS) emerged. This concept expanded the computerized support to top-level managers and executives. Some of the

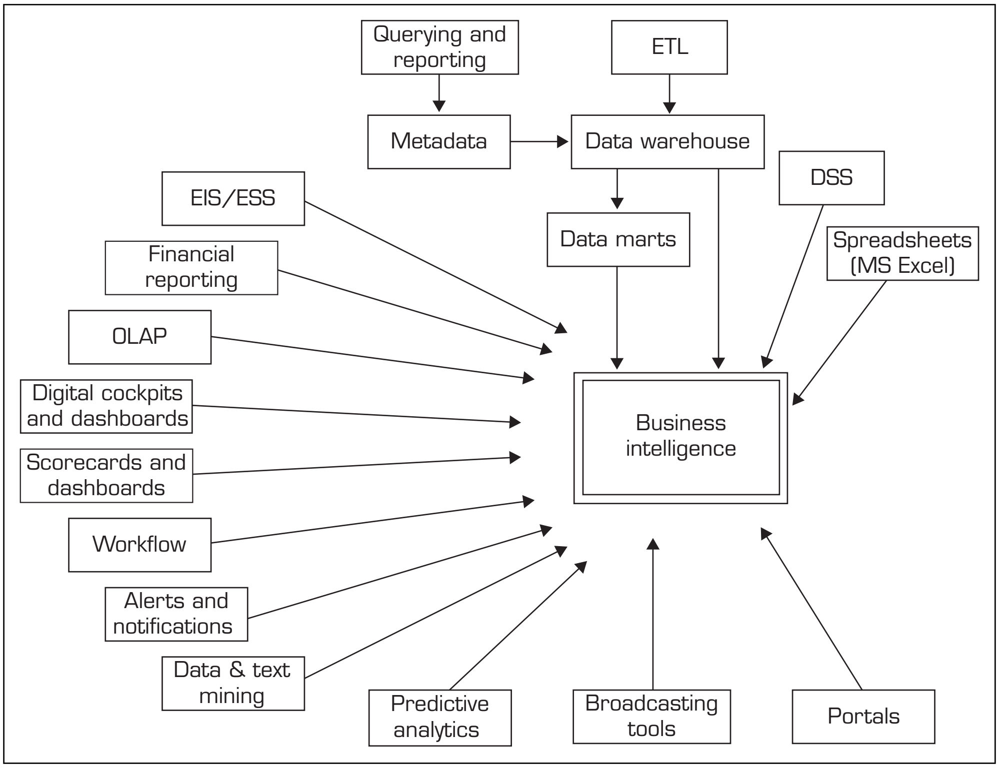

FIGURE 1.3 Evolution of Business Intelligence (BI).

capabilities introduced were dynamic multidimensional (ad hoc or on-demand) reporting, forecasting and prediction, trend analysis, drill-down to details, status access, and critical success factors. These features appeared in dozens of commercial products until the mid-1990s. Then the same capabilities and some new ones appeared under the name BI. Today, a good BI-based enterprise information system contains all the information executives need. So, the original concept of EIS was transformed into BI. By 2005, BI systems started to include *artificial intelligence* capabilities as well as powerful analytical capabilities. Figure 1.3 illustrates the various tools and techniques that may be included in a BI system. It illustrates the evolution of BI as well. The tools shown in Figure 1.3 provide the capabilities of BI. The most sophisticated BI products include most of these capabilities; others specialize in only some of them. We will study several of these capabilities in more detail in Chapters 5 through 9.

The Architecture of BI

A BI system has four major components: a *data warehouse*, with its source data; *business analytics*, a collection of tools for manipulating, mining, and analyzing the data in the data warehouse; *business performance management (BPM)* for monitoring and analyzing performance; and a *user interface* (e.g., a dashboard). The relationship among these components is illustrated in Figure 1.4. We will discuss these components in detail in Chapters 3 through 9.

Styles of BI

The architecture of BI depends on its applications. MicroStrategy Corp. distinguishes five styles of BI and offers special tools for each. The five styles are report delivery and alerting; enterprise reporting (using dashboards and scorecards); cube analysis (also known as slice-and-dice analysis); ad hoc queries; and statistics and data mining.

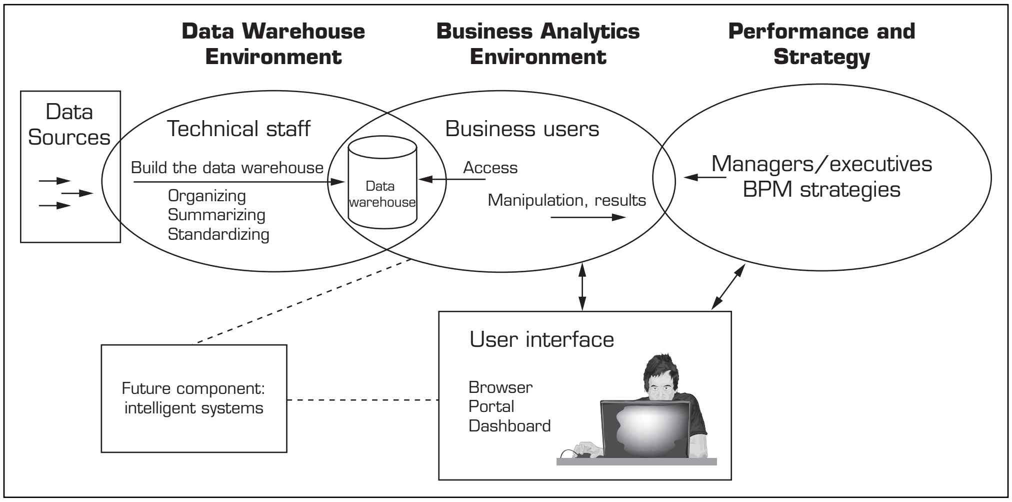

FIGURE 1.4 A High-Level Architecture of BI. *Source:* Based on W. Eckerson, *Smart Companies in the 21st Century: The Secrets of Creating Successful Business Intelligent Solutions.* The Data Warehousing Institute, Seattle, WA, 2003, p. 32, Illustration 5.

The Origins and Drivers of BI

Where did modern approaches to data warehousing (DW) and BI come from? What are their roots, and how do those roots affect the way organizations are managing these initiatives today? Today's investments in information technology are under increased scrutiny in terms of their bottom-line impact and potential. The same is true of DW and the BI applications that make these initiatives possible.

Organizations are being compelled to capture, understand, and harness their data to support decision making in order to improve business operations. Legislation and regulation (e.g., the Sarbanes-Oxley Act of 2002) now require business leaders to document their business processes and to sign off on the legitimacy of the information they rely on and report to stakeholders. Moreover, business cycle times are now extremely compressed; faster, more informed, and better decision making is therefore a competitive imperative. Managers need the *right information* at the *right time* and in the *right place*. This is the mantra for modern approaches to BI.

Organizations have to work smart. Paying careful attention to the management of BI initiatives is a necessary aspect of doing business. It is no surprise, then, that organizations are increasingly championing BI. You will hear about more BI successes and the fundamentals of those successes in Chapters 3 through 9. Examples of many applications of BI are provided in Table 1.3. Application Case 1.1 illustrates one such application of BI that has helped many airlines, as well as the companies offering such services to the airlines.

A Multimedia Exercise in Business Intelligence

Teradata University Network (TUN) includes some videos along the lines of the television show *CSI* to illustrate concepts of analytics in different industries. These are called "BSI Videos (Business Scenario Investigations)." Not only these are entertaining, but they also provide the class with some questions for discussion. For starters, please go to **teradatauniversitynetwork.com/teach-and-learn/library-item/?LibraryItemId=889**. Watch the video that appears on YouTube. Essentially, you have to assume the role of a customer service center professional. An incoming flight is running late, and several passengers are likely to miss their connecting flights. There are seats on one outgoing flight that can accommodate two of the four passengers. Which two passengers should be given

TABLE 1.3 Business Value of BI Analytical Applications

Analytic Application	Business Question	Business Value
Customer segmentation	What market segments do my customers fall into, and what are their characteristics?	Personalize customer relationships for higher satisfaction and retention.
Propensity to buy	Which customers are most likely to respond to my promotion?	Target customers based on their need to increase their loyalty to your product line. Also, increase campaign profitability by focusing on the most likely to buy.
Customer profitability	What is the lifetime profitability of my customer?	Make individual business interaction decisions based on the overall profitability of customers.
Fraud detection	How can I tell which transactions are likely to be fraudulent?	Quickly determine fraud and take immediate action to minimize cost.
Customer attrition	Which customer is at risk of leaving?	Prevent loss of high-value customers and let go of lower-value customers.
Channel optimization	What is the best channel to reach my customer in each segment?	Interact with customers based on their preference and your need to manage cost.

Source: A. Ziama and J. Kasher, *Data Mining Primer for the Data Warehousing Professional*. Teradata, Dayton, OH, 2004.

Application Case 1.1

Sabre Helps Its Clients Through Dashboards and Analytics

Sabre is one of the world leaders in the travel industry, providing both business-to-consumer services as well as business-to-business services. It serves travelers, travel agents, corporations, and travel suppliers through its four main companies: Travelocity, Sabre Travel Network, Sabre Airline Solutions, and Sabre Hospitality Solutions. The current volatile global economic environment poses significant competitive challenges to the airline industry. To stay ahead of the competition, Sabre Airline Solutions recognized that airline executives needed enhanced tools for managing their business decisions by eliminating the traditional, manual, time-consuming process of collecting and aggregating financial and other information needed for actionable initiatives. This enables real-time decision support at airlines throughout the world that maximize their (and, in turn, Sabre's) return on information by driving insights, actionable intelligence, and value for customers from the growing data.

Sabre developed an Enterprise Travel Data Warehouse (ETDW) using Teradata to hold its massive reservations data. ETDW is updated in near-real time with batches that run every 15 minutes, gathering data from all of Sabre's businesses. Sabre uses its ETDW to create Sabre Executive Dashboards that provide near–real-time executive insights using a Cognos 8 BI platform with Oracle Data Integrator and Oracle Goldengate technology infrastructure. The Executive Dashboards offer their client airlines' top-level managers and decision makers a timely, automated, user-friendly solution, aggregating critical performance metrics in a succinct way and providing at a glance a 360-degree view of the overall health of the airline. At one airline, Sabre's Executive Dashboards provide senior management with a daily and intra-day snapshot of key performance indicators in a single application, replacing the once-a-week, 8-hour process of generating the same report from various data sources. The use of dashboards is not limited to the external customers; Sabre also uses them for their assessment of internal operational performance.

The dashboards help Sabre's customers to have a clear understanding of the data through the visual displays that incorporate interactive drill-down capabilities. It replaces flat presentations and allows for more focused review of the data with less effort and

(*Continued*)

Application Case 1.1 (Continued)

time. This facilitates team dialog by making the data/metrics pertaining to sales performance, including ticketing, seats sold and flown, operational performance such as data on flight movement and tracking, customer reservations, inventory, and revenue across an airline's multiple distribution channels, available to many stakeholders. The dashboard systems provide scalable infrastructure, graphical user interface (GUI) support, data integration, and data aggregation that empower airline executives to be more proactive in taking actions that lead to positive impacts on the overall health of their airline.

With its ETDW, Sabre could also develop other Web-based analytical and reporting solutions that leverage data to gain customer insights through analysis of customer profiles and their sales interactions to calculate customer value. This enables better customer segmentation and insights for value-added services.

Questions for Discussion

1. What is traditional reporting? How is it used in organizations?
2. How can analytics be used to transform traditional reporting?
3. How can interactive reporting assist organizations in decision making?

What We Can Learn from This Application Case

This Application Case shows that organizations that earlier used reporting only for tracking their internal business activities and meeting compliance requirements set out by the government are now moving toward generating actionable intelligence from their transactional business data. Reporting has become broader as organizations are now trying to analyze archived transactional data to understand underlying hidden trends and patterns that would enable them to make better decisions by gaining insights into problematic areas and resolving them to pursue current and future market opportunities. Reporting has advanced to interactive online reports that enable users to pull and quickly build custom reports as required and even present the reports aided by visualization tools that have the ability to connect to the database, providing the capabilities of digging deep into summarized data.

Source: Teradata.com, "Sabre Airline Solutions," **teradata.com/t/case-studies/Sabre-Airline-Solutions-EB6281** (accessed February 2013).

priority? You are given information about customers' profiles and relationship with the airline. Your decisions might change as you learn more about those customers' profiles.

Watch the video, pause it as appropriate, and answer the questions on which passengers should be given priority. Then resume the video to get more information. After the video is complete, you can see the slides related to this video and how the analysis was prepared on a slide set at **teradatauniversitynetwork.com/templates/Download.aspx?ContentItemId=891.** Please note that access to this content requires initial registration.

This multimedia excursion provides an example of how additional information made available through an enterprise data warehouse can assist in decision making.

The DSS–BI Connection

By now, you should be able to see some of the similarities and differences between DSS and BI. First, their architectures are very similar because BI evolved from DSS. However, BI implies the use of a data warehouse, whereas DSS may or may not have such a feature. BI is, therefore, more appropriate for large organizations (because data warehouses are expensive to build and maintain), but DSS can be appropriate to any type of organization.

Second, most DSS are constructed to *directly* support specific decision making. BI systems, in general, are geared to provide accurate and timely information, and they support decision support *indirectly*. This situation is changing, however, as more and more decision support tools are being added to BI software packages.

Third, BI has an executive and strategy orientation, especially in its BPM and dashboard components. DSS, in contrast, is oriented toward analysts.

Fourth, most BI systems are constructed with commercially available tools and components that are fitted to the needs of organizations. In building DSS, the interest may be in constructing solutions to very unstructured problems. In such situations, more programming (e.g., using tools such as Excel) may be needed to customize the solutions.

Fifth, DSS methodologies and even some tools were developed mostly in the academic world. BI methodologies and tools were developed mostly by software companies. (See Zaman, 2005, for information on how BI has evolved.)

Sixth, many of the tools that BI uses are also considered DSS tools. For example, data mining and predictive analysis are core tools in both areas.

Although some people equate DSS with BI, these systems are not, at present, the same. It is interesting to note that some people believe that DSS is a part of BI—one of its analytical tools. Others think that BI is a special case of DSS that deals mostly with reporting, communication, and collaboration (a form of data-oriented DSS). Another explanation (Watson, 2005) is that BI is a result of a continuous revolution and, as such, DSS is one of BI's original elements. In this book, we separate DSS from BI. However, we point to the DSS–BI connection frequently. Further, as noted in the next section onward, in many circles BI has been subsumed by the new term *analytics* or *data science*.

SECTION 1.7 REVIEW QUESTIONS

1. Define *BI*.

2. List and describe the major components of BI.

3. What are the major similarities and differences of DSS and BI?

1.8 BUSINESS ANALYTICS OVERVIEW

The word "analytics" has replaced the previous individual components of computerized decision support technologies that have been available under various labels in the past. Indeed, many practitioners and academics now use the word *analytics* in place of BI. Although many authors and consultants have defined it slightly differently, one can view analytics as the process of developing actionable decisions or recommendation for actions based upon insights generated from historical data. The Institute for Operations Research and Management Science (INFORMS) has created a major initiative to organize and promote analytics. According to INFORMS, analytics represents the combination of computer technology, management science techniques, and statistics to solve real problems. Of course, many other organizations have proposed their own interpretations and motivation for analytics. For example, SAS Institute Inc. proposed eight levels of analytics that begin with standardized reports from a computer system. These reports essentially provide a sense of what is happening with an organization. Additional technologies have enabled us to create more customized reports that can be generated on an ad hoc basis. The next extension of reporting takes us to online analytical processing (OLAP)–type queries that allow a user to dig deeper and determine the specific source of concern or opportunities. Technologies available today can also automatically issue alerts for a decision maker when performance issues warrant such alerts. At a consumer level we see such alerts for weather or other issues. But similar alerts can also be generated in specific settings when sales fall above or below a certain level within a certain time period or when the inventory for a specific product is running low. All of these applications are made possible through analysis and queries on data being collected by an organization. The next level of analysis might entail statistical analysis to better understand patterns. These can then be taken a step further to develop forecasts or models for predicting how customers might respond to

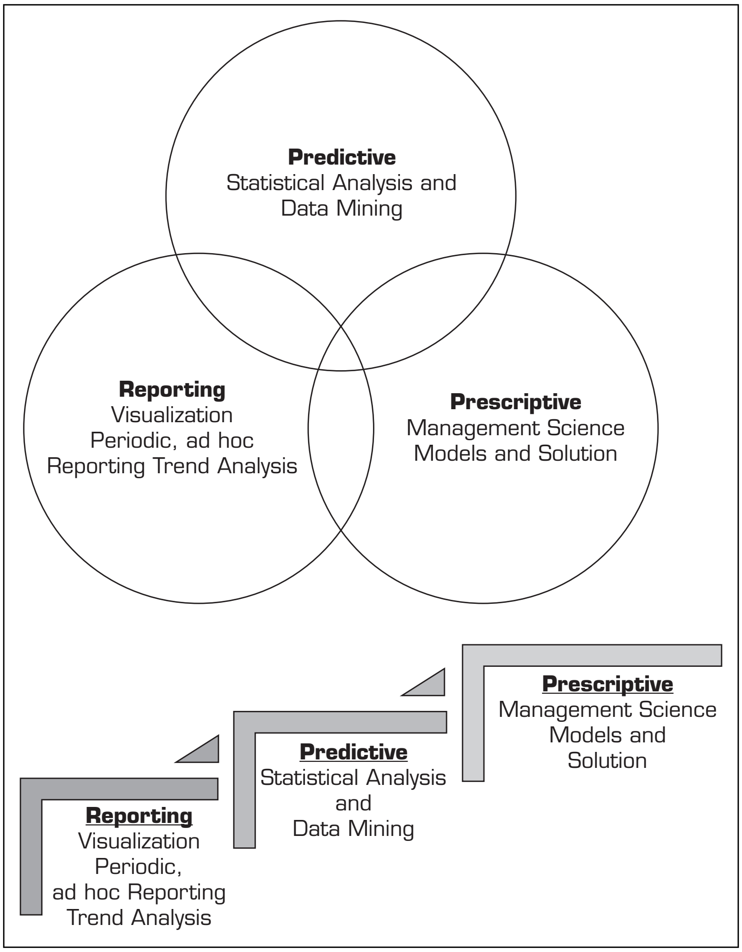

FIGURE 1.5 Three Types of Analytics.

a specific marketing campaign or ongoing service/product offerings. When an organization has a good view of what is happening and what is likely to happen, it can also employ other techniques to make the best decisions under the circumstances. These eight levels of analytics are described in more detail in a white paper by SAS (**sas.com/news/sascom/analytics_levels.pdf**).

This idea of looking at all the data to understand what is happening, what will happen, and how to make the best of it has also been encapsulated by INFORMS in proposing three levels of analytics. These three levels are identified (**informs.org/Community/Analytics**) as descriptive, predictive, and prescriptive. Figure 1.5 presents two graphical views of these three levels of analytics. One view suggests that these three are somewhat independent steps (of a ladder) and one type of analytics application leads to another. The interconnected circles view suggests that there is actually some overlap across these three types of analytics. In either case, the interconnected nature of different types of analytics applications is evident. We next introduce these three levels of analytics.

Descriptive Analytics

Descriptive or reporting analytics refers to knowing what is happening in the organization and understanding some underlying trends and causes of such occurrences. This involves, first of all, consolidation of data sources and availability of all relevant data in a form that enables appropriate reporting and analysis. Usually development of this data infrastructure is part of data warehouses, which we study in Chapter 3. From this data infrastructure we can develop appropriate reports, queries, alerts, and trends using various reporting tools and techniques. We study these in Chapter 4.

A significant technology that has become a key player in this area is visualization. Using the latest visualization tools in the marketplace, we can now develop powerful insights into the operations of our organization. Application Cases 1.2 and 1.3 highlight some such applications in the healthcare domain. Color renderings of such applications are available on the companion Web site and also on Tableau's Web site. Chapter 4 covers visualization in more detail.

Application Case 1.2

Eliminating Inefficiencies at Seattle Children's Hospital

Seattle Children's was the seventh highest ranked children's hospital in 2011, according to *U.S. News & World Report*. For any organization that is committed to saving lives, identifying and removing the inefficiencies from systems and processes so that more resources become available to cater to patient care become very important. At Seattle Children's, management is continuously looking for new ways to improve the quality, safety, and processes from the time a patient is admitted to the time they are discharged. To this end, they spend a lot of time in analyzing the data associated with the patient visits.

To quickly turn patient and hospital data into insights, Seattle Children's implemented Tableau Software's business intelligence application. It provides a browser based on easy-to-use analytics to the stakeholders; this makes it intuitive for individuals to create visualizations and to understand what the data has to offer. The data analysts, business managers, and financial analysts as well as clinicians, doctors, and researchers are all using descriptive analytics to solve different problems in a much faster way. They are developing visual systems on their own, resulting in dashboards and scorecards that help in defining the standards, the current performance achieved measured against the standards, and how these systems will grow into the future. Through the use of monthly and daily dashboards, day-to-day decision making at Seattle Children's has improved significantly.

Seattle Children's measures patient wait-times and analyzes them with the help of visualizations to discover the root causes and contributing factors for patient waiting. They found that early delays cascaded during the day. They focused on on-time appointments of patient services as one of the solutions to improving patient overall waiting time and increasing the availability of beds. Seattle Children's saved about $3 million from the supply chain, and with the help of tools like Tableau, they are finding new ways to increase savings while treating as many patients as possible by making the existing processes more efficient.

Questions for Discussion

1. Who are the users of the tool?
2. What is a dashboard?
3. How does visualization help in decision making?
4. What are the significant results achieved by the use of Tableau?

What We Can Learn from This Application Case

This Application Case shows that reporting analytics involving visualizations such as dashboards can offer major insights into existing data and show how a variety of users in different domains and departments can contribute toward process and quality improvements in an organization. Furthermore, exploring the data visually can help in identifying the root causes of problems and provide a basis for working toward possible solutions.

Source: Tableausoftware.com, "Eliminating Waste at Seattle Children's," **tableausoftware.com/eliminating-waste-at-seattle-childrens** (accessed February 2013).

Application Case 1.3

Analysis at the Speed of Thought

Kaleida Health, the largest healthcare provider in western New York, has more than 10,000 employees, five hospitals, a number of clinics and nursing homes, and a visiting-nurse association that deals with millions of patient records. Kaleida's traditional reporting tools were inadequate to handle the growing data, and they were faced with the challenge of finding a business intelligence tool that could handle large data sets effortlessly, quickly, and with a much deeper analytic capability.

At Kaleida, many of the calculations are now done in Tableau, primarily pulling the data from Oracle databases into Excel and importing the data into Tableau. For many of the monthly analytic reports, data is directly extracted into Tableau from the data warehouse; many of the data queries are saved and rerun, resulting in time savings when dealing with millions of records—each having more than 40 fields per record. Besides speed, Kaleida also uses Tableau to merge different tables for generating extracts.

Using Tableau, Kaleida can analyze emergency room data to determine the number of patients who visit more than 10 times a year. The data often reveal that people frequently use emergency room and ambulance services inappropriately for stomachaches, headaches, and fevers. Kaleida can manage resource utilizations—the use and cost of supplies—which will ultimately lead to efficiency and standardization of supplies management across the system.

Kaleida now has its own business intelligence department and uses Tableau to compare itself to other hospitals across the country. Comparisons are made on various aspects, such as length of patient stay, hospital practices, market share, and partnerships with doctors.

Questions for Discussion

1. What are the desired functionalities of a reporting tool?
2. What advantages were derived by using a reporting tool in the case?

What We Can Learn from This Application Case

Correct selection of a reporting tool is extremely important, especially if an organization wants to derive value from reporting. The generated reports and visualizations should be easily discernible; they should help people in different sectors make sense out of the reports, identify the problematic areas, and contribute toward improving them. Many future organizations will require reporting analytic tools that are fast and capable of handling huge amounts of data efficiently to generate desired reports without the need for third-party consultants and service providers. A truly useful reporting tool can exempt organizations from unnecessary expenditure.

Source: Tableausoftware.com, "Kaleida Health Finds Efficiencies, Stays Competitive," **tableausoftware.com/learn/stories/user-experience-speed-thought-kaleida-health** (accessed February 2013).

Predictive Analytics

Predictive analytics aims to determine what is likely to happen in the future. This analysis is based on statistical techniques as well as other more recently developed techniques that fall under the general category of data mining. The goal of these techniques is to be able to predict if the customer is likely to switch to a competitor ("churn"), what the customer is likely to buy next and how much, what promotion a customer would respond to, or whether this customer is a creditworthy risk. A number of techniques are used in developing predictive analytical applications, including various classification algorithms. For example, as described in Chapters 5 and 6, we can use classification techniques such as decision tree models and neural networks to predict how well a motion picture will do at the box office. We can also use clustering algorithms for segmenting customers into different clusters to be able to target specific promotions to them. Finally, we can use association mining techniques to estimate relationships between different purchasing behaviors. That is, if a customer buys one product, what else is the customer likely to purchase? Such analysis can assist a retailer in recommending or promoting related products. For example, any product search on Amazon.com results in the retailer also suggesting other similar products that may interest a customer. We will study these techniques and their applications in Chapters 6 through 9. Application Cases 1.4 and 1.5 highlight some similar applications. Application Case 1.4 introduces a movie you may have heard of: *Moneyball*. It is perhaps one of the best examples of applications of predictive analysis in sports.

Application Case 1.4

Moneyball: Analytics in Sports and Movies

Moneyball, a biographical, sports, drama film, was released in 2011 and directed by Bennett Miller. The film was based on Michael Lewis's book, *Moneyball*. The movie gave a detailed account of the Oakland Athletics baseball team during the 2002 season and the Oakland general manager's efforts to assemble a competitive team.

The Oakland Athletics suffered a big loss to the New York Yankees in 2001 postseason. As a result, Oakland lost many of its star players to free agency and ended up with a weak team with unfavorable financial prospects. The general manager's efforts to reassemble a competitive team were denied because Oakland had limited payroll. The scouts for the Oakland Athletics followed the old baseball custom of making subjective decisions when selecting the team members. The general manager then met a young, computer whiz with an economics degree from Yale. The general manager decided to appoint him as the new assistant general manager.

The assistant general manager had a deep passion for baseball and had the expertise to crunch the numbers for the game. His love for the game made him develop a radical way of understanding baseball statistics. He was a disciple of Bill James, a marginal figure who offered rationalized techniques to analyze baseball. James looked at baseball statistics in a different way, crunching the numbers purely on facts and eliminating subjectivity. James pioneered the nontraditional analysis method called the Sabermetric approach, which derived from SABR—Society for American Baseball Research.

The assistant general manager followed the Sabermetric approach by building a prediction model to help the Oakland Athletics select players based on their "on-base percentage" (OBP), a statistic that measured how often a batter reached base for any reason other than fielding error, fielder's choice, dropped/uncaught third strike, fielder's obstruction, or catcher's interference. Rather than relying on the scout's experience and intuition, the assistant general manager selected players based almost exclusively on OBP.

Spoiler Alert: The new team beat all odds, won 20 consecutive games, and set an American League record.

Questions for Discussion

1. How is predictive analytics applied in *Moneyball*?
2. What is the difference between objective and subjective approaches in decision making?

What We Can Learn from This Application Case

Analytics finds its use in a variety of industries. It helps organizations rethink their traditional problem-solving abilities, which are most often subjective, relying on the same old processes to find a solution. Analytics takes the radical approach of using historical data to find fact-based solutions that will remain appropriate for making even future decisions.

Source: Wikipedia, "On-Base Percentage," **en.wikipedia.org/wiki/On_base_percentage** (accessed January 2013); Wikipedia, "Sabermetricsm," **wikipedia.org/wiki/Sabermetrics** (accessed January 2013).

Application Case 1.5

Analyzing Athletic Injuries

Any athletic activity is prone to injuries. If the injuries are not handled properly, then the team suffers. Using analytics to understand injuries can help in deriving valuable insights that would enable the coaches and team doctors to manage the team composition, understand player profiles, and ultimately aid in better decision making concerning which players might be available to play at any given time.

In an exploratory study, Oklahoma State University analyzed American football-related sport injuries by using reporting and predictive analytics. The project followed the CRISP-DM methodology to understand the problem of making recommendations on managing injuries, understanding the various data elements collected about injuries, cleaning the data, developing visualizations to draw various inferences, building predictive models to analyze the injury healing time period, and drawing sequence rules to predict the relationship among the injuries and the various body part parts afflicted with injuries.

The injury data set consisted of more than 560 football injury records, which were categorized into injury-specific variables—body part/site/laterality, action taken, severity, injury type, injury start and healing dates—and player/sport-specific variables—player ID, position played, activity, onset, and game location. Healing time was calculated for each record, which was classified into different sets of time periods: 0–1 month, 1–2 months, 2–4 months, 4–6 months, and 6–24 months.

Various visualizations were built to draw inferences from injury data set information depicting the healing time period associated with players' positions, severity of injuries and the healing time period, treatment offered and the associated healing time period, major injuries afflicting body parts, and so forth.

Neural network models were built to predict each of the healing categories using IBM SPSS Modeler. Some of the predictor variables were current status of injury, severity, body part, body site, type of injury, activity, event location, action taken, and position played. The success of classifying the healing category was quite good: Accuracy was 79.6 percent. Based on the analysis, many business recommendations were suggested, including employing more specialists' input from injury onset instead of letting the training room staff screen the injured players; training players at defensive positions to avoid being injured; and holding practice to thoroughly safety-check mechanisms.

Questions for Discussion

1. What types of analytics are applied in the injury analysis?
2. How do visualizations aid in understanding the data and delivering insights into the data?
3. What is a classification problem?
4. What can be derived by performing sequence analysis?

What We Can Learn from This Application Case

For any analytics project, it is always important to understand the business domain and the current state of the business problem through extensive analysis of the only resource—historical data. Visualizations often provide a great tool for gaining the initial insights into data, which can be further refined based on expert opinions to identify the relative importance of the data elements related to the problem. Visualizations also aid in generating ideas for obscure business problems, which can be pursued in building predictive models that could help organizations in decision making.

Prescriptive Analytics

The third category of analytics is termed **prescriptive analytics**. The goal of prescriptive analytics is to recognize what is going on as well as the likely forecast and make decisions to achieve the best performance possible. This group of techniques has historically been studied under the umbrella of operations research or management sciences and has generally been aimed at optimizing the performance of a system. The goal here is to provide a decision or a recommendation for a specific action. These recommendations can be in the forms of a specific yes/no decision for a problem, a specific amount (say, price for a specific item or airfare to charge), or a complete set of production plans. The decisions may be presented to a decision maker in a report or may directly be used in an automated decision rules system (e.g., in airline pricing systems). Thus, these types of analytics can also be termed **decision or normative analytics**. Application Case 1.6 gives an example of such prescriptive analytic applications. We will learn about some of these techniques and several additional applications in Chapters 10 through 12.

Application Case 1.6

Industrial and Commercial Bank of China (ICBC) Employs Models to Reconfigure Its Branch Network

The Industrial and Commercial Bank of China (ICBC) has more than 16,000 branches and serves over 230 million individual customers and 3.6 million corporate clients. Its daily financial transactions total about $180 million. It is also the largest publicly traded bank in the world in terms of market capitalization, deposit volume, and profitability. To stay competitive and increase profitability, ICBC was faced with the challenge to quickly adapt to the fast-paced economic growth, urbanization, and increase in personal wealth of the Chinese. Changes had to be implemented in over 300 cities with high variability in customer behavior and financial status. Obviously, the nature of the challenges in such a huge economy meant that a large-scale optimization solution had to be developed to locate branches in the right places, with right services, to serve the right customers.

With their existing method, ICBC used to decide where to open new branches through a scoring model in which different variables with varying weight were used as inputs. Some of the variables were customer flow, number of residential households, and number of competitors in the intended geographic region. This method was deficient in determining the customer distribution of a geographic area. The existing method was also unable to optimize the distribution of bank branches in the branch network. With support from IBM, a branch reconfiguration (BR) tool was developed. Inputs for the BR system are in three parts:

a. Geographic data with 83 different categories
b. Demographic and economic data with 22 different categories
c. Branch transactions and performance data that consisted of more than 60 million transaction records each day

These three inputs helped generate accurate customer distribution for each area and, hence, helped the bank optimize its branch network. The BR system consisted of a market potential calculation model, a branch network optimization model, and a branch site evaluation model. In the market potential model, the customer volume and value is measured based on input data and expert knowledge. For instance, expert knowledge would help determine if personal income should be weighted more than gross domestic product (GDP). The geographic areas are also demarcated into cells, and the preference of one cell over the other is determined. In the branch network optimization model, mixed integer programming is used to locate branches in candidate cells so that they cover the largest market potential areas. In the branch site evaluation model, the value for establishing bank branches at specific locations is determined.

Since 2006, the development of the BR has been improved through an iterative process. ICBC's branch reconfiguration tool has increased deposits by $21.2 billion since its inception. This increase in deposit is because the bank can now reach more customers with the right services by use of its optimization tool. In a specific example, when BR was implemented in Suzhou in 2010, deposits increased to $13.67 billion from an initial level of $7.56 billion in 2007. Hence, the BR tool assisted in an increase of deposits to the tune of $6.11 billion between 2007 and 2010. This project was selected as a finalist in the Edelman Competition 2011, which is run by INFORMS to promote actual applications of management science/operations research models.

(*Continued*)

Application Case 1.6 (Continued)

Questions for Discussion

1. How can analytical techniques help organizations to retain competitive advantage?
2. How can descriptive and predictive analytics help in pursuing prescriptive analytics?
3. What kinds of prescriptive analytic techniques are employed in the case study?
4. Are the prescriptive models once built good forever?

What We Can Learn from This Application Case

Many organizations in the world are now embracing analytical techniques to stay competitive and achieve growth. Many organizations provide consulting solutions to the businesses in employing prescriptive analytical solutions. It is equally important to have proactive decision makers in the organizations who are aware of the changing economic environment as well as the advancements in the field of analytics to ensure that appropriate models are employed. This case shows an example of geographic market segmentation and customer behavioral segmentation techniques to isolate the profitability of customers and employ optimization techniques to locate the branches that deliver high profitability in each geographic segment.

Source: X. Wang et al., "Branch Reconfiguration Practice Through Operations Research in Industrial and Commercial Bank of China," *Interfaces*, January/February 2012, Vol. 42, No. 1, pp. 33–44; DOI: 10.1287/inte.1110.0614.

Analytics Applied to Different Domains

Applications of analytics in various industry sectors have spawned many related areas or at least buzzwords. It is almost fashionable to attach the word *analytics* to any specific industry or type of data. Besides the general category of text analytics—aimed at getting value out of text (to be studied in Chapter 6)—or Web analytics—analyzing Web data streams (Chapter 7)—many industry- or problem-specific analytics professions/streams have come up. Examples of such areas are marketing analytics, retail analytics, fraud analytics, transportation analytics, health analytics, sports analytics, talent analytics, behavioral analytics, and so forth. For example, Application Case 1.1 could also be termed as a case study in airline analytics. Application Cases 1.2 and 1.3 would belong to health analytics; Application Cases 1.4 and 1.5 to sports analytics; Application Case 1.6 to bank analytics; and Application Case 1.7 to retail analytics. The End-of-Chapter Application Case could be termed insurance analytics. Literally, any systematic analysis of data in a specific sector is being labeled as "(fill-in-blanks)" Analytics. Although this may result in overselling the concepts of analytics, the benefit is that more people in specific industries are aware of the power and potential of analytics. It also provides a focus to professionals developing and applying the concepts of analytics in a vertical sector. Although many of the techniques to develop analytics applications may be common, there are unique issues within each vertical segment that influence how the data may be collected, processed, analyzed, and the applications implemented. Thus, the differentiation of analytics based on a vertical focus is good for the overall growth of the discipline.

Analytics or Data Science?

Even as the concept of analytics is getting popular among industry and academic circles, another term has already been introduced and is becoming popular. The new term is *data science*. Thus the practitioners of data science are data scientists. Mr. D. J. Patil of LinkedIn is sometimes credited with creating the term *data science*. There have been some attempts to describe the differences between data analysts and data scientists (e.g., see this study at **emc.com/collateral/about/news/emc-data-science-study-wp.pdf**). One view is that *data analyst* is just another term for professionals who were doing business intelligence in the form of data compilation, cleaning, reporting, and perhaps some visualization. Their skill sets included Excel, some SQL knowledge, and reporting. A reader of Section 1.8 would recognize that as descriptive or reporting analytics. In contrast, a data scientist is responsible for predictive analysis, statistical analysis, and more advanced analytical tools and algorithms. They may have a deeper knowledge of algorithms and may recognize them under various labels—data mining, knowledge discovery, machine learning, and so forth. Some of these professionals may also need deeper programming knowledge to be able to write code for data cleaning and analysis in current Web-oriented languages such as Java and Python. Again, our readers should recognize these as falling under the predictive and prescriptive analytics umbrella. Our view is that the distinction between analytics and data science is more of a degree of technical knowledge and skill sets than the functions. It may also be more of a distinction across disciplines. Computer science, statistics, and applied mathematics programs appear to prefer the data science label, reserving the analytics label for more business-oriented professionals. As another example of this, applied physics professionals have proposed using *network science* as the term for describing analytics that relate to a group of people—social networks, supply chain networks, and so forth. See **barabasilab.neu.edu/networksciencebook/downlPDF.html** for an evolving textbook on this topic.

Aside from a clear difference in the skill sets of professionals who only have to do descriptive/reporting analytics versus those who engage in all three types of analytics, the distinction is fuzzy between the two labels, at best. We observe that graduates of our analytics programs tend to be responsible for tasks more in line with data science professionals (as defined by some circles) than just reporting analytics. This book is clearly aimed at introducing the capabilities and functionality of all analytics (which includes data science), not just reporting analytics. From now on, we will use these terms interchangeably.

SECTION 1.8 REVIEW QUESTIONS

1. Define *analytics*.
2. What is descriptive analytics? What various tools are employed in descriptive analytics?
3. How is descriptive analytics different from traditional reporting?
4. What is a data warehouse? How can data warehousing technology help in enabling analytics?
5. What is predictive analytics? How can organizations employ predictive analytics?
6. What is prescriptive analytics? What kinds of problems can be solved by prescriptive analytics?
7. Define modeling from the analytics perspective.
8. Is it a good idea to follow a hierarchy of descriptive and predictive analytics before applying prescriptive analytics?
9. How can analytics aid in objective decision making?

1.9 BRIEF INTRODUCTION TO BIG DATA ANALYTICS

What Is Big Data?

Our brains work extremely quickly and are efficient and versatile in processing large amounts of all kinds of data: images, text, sounds, smells, and video. We process all different forms of data relatively easily. Computers, on the other hand, are still finding it hard to keep up with the pace at which data is generated—let alone analyze it quickly. We have the problem of Big Data. So what is Big Data? Simply put, it is data that cannot be stored in a single storage unit. Big Data typically refers to data that is arriving in many different forms, be they structured, unstructured, or in a stream. Major sources of such data are clickstreams from Web sites, postings on social media sites such as Facebook, or data from traffic, sensors, or weather. A Web search engine like Google needs to search and index billions of Web pages in order to give you relevant search results in a fraction of a second. Although this is not done in real time, generating an index of all the Web pages on the Internet is not an easy task. Luckily for Google, it was able to solve this problem. Among other tools, it has employed Big Data analytical techniques.

There are two aspects to managing data on this scale: storing and processing. If we could purchase an extremely expensive storage solution to store all the data at one place on one unit, making this unit fault tolerant would involve major expense. An ingenious solution was proposed that involved storing this data in chunks on different machines connected by a network, putting a copy or two of this chunk in different locations on the network, both logically and physically. It was originally used at Google (then called *Google File System*) and later developed and released as an Apache project as the Hadoop Distributed File System (HDFS).

However, storing this data is only half the problem. Data is worthless if it does not provide business value, and for it to provide business value, it has to be analyzed. How are such vast amounts of data analyzed? Passing all computation to one powerful computer does not work; this scale would create a huge overhead on such a powerful computer. Another ingenious solution was proposed: Push computation to the data, instead of pushing data to a computing node. This was a new paradigm, and it gave rise to a whole new way of processing data. This is what we know today as the MapReduce programming paradigm, which made processing Big Data a reality. MapReduce was originally developed at Google, and a subsequent version was released by the Apache project called Hadoop MapReduce.

Today, when we talk about storing, processing, or analyzing Big Data, HDFS and MapReduce are involved at some level. Other relevant standards and software solutions have been proposed. Although the major toolkit is available as open source, several companies have been launched to provide training or specialized analytical hardware or software services in this space. Some examples are HortonWorks, Cloudera, and Teradata Aster.

Over the past few years, what was called Big Data changed more and more as Big Data applications appeared. The need to process data coming in at a rapid rate added velocity to the equation. One example of fast data processing is algorithmic trading. It is the use of electronic platforms based on algorithms for trading shares on the financial market, which operates in the order of microseconds. The need to process different kinds of data added variety to the equation. Another example of the wide variety of data is sentiment analysis, which uses various forms of data from social media platforms and customer responses to gauge sentiments. Today Big Data is associated with almost any kind of large data that has the characteristics of volume, velocity, and variety. Application Case 1.7 illustrates one example of Big Data analytics. We will study Big Data characteristics in more detail in Chapters 3 and 13.

SECTION 1.9 REVIEW QUESTIONS

1. What is Big Data analytics?

2. What are the sources of Big Data?

3. What are the characteristics of Big Data?

4. What processing technique is applied to process Bi ta?

Application Case 1.7

Gilt Groupe's Flash Sales Streamlined by Big Data Analytics

Gilt Groupe is an online destination offering flash sales for major brands by selling their clothing and accessories. It offers its members exclusive discounts on high-end clothing and other apparel. After registering with Gilt, customers are sent e-mails containing a variety of offers. Customers are given a 36-48 hour window to make purchases using these offers. There are about 30 different sales each day. While a typical department store turns over its inventory two or three times a year, Gilt does it eight to 10 times a year. Thus, they have to manage their inventory extremely well or they could incur extremely high inventory costs. In order to do this, analytics software developed at Gilt keeps track of every customer click—ranging from what brands the customers click on, what colors they choose, what styles they pick, and what they end up buying. Then Gilt tries to predict what these customers are more likely to buy and stocks inventory according to these predictions. Customers are sent customized alerts to sale offers depending on the suggestions by the analytics software.

That, however, is not the whole process. The software also monitors what offers the customers choose from the recommended offers to make more accurate predictions and to increase the effectiveness of its personalized recommendations. Some customers do not check e-mail that often. Gilt's analytics software keeps track of responses to offers and sends the same offer 3 days later to those customers who haven't responded. Gilt also keeps track of what customers are saying in general about Gilt's products by analyzing Twitter feeds to analyze sentiment. Gilt's recommendation software is based on Teradata Aster's technology solution that includes Big Data analytics technologies.

Questions for Discussion

1. What makes this case study an example of Big Data analytics?
2. What types of decisions does Gilt Groupe have to make?

What We Can Learn From this Application Case

There is continuous growth in the amount of structured and unstructured data, and many organizations are now tapping these data to make actionable decisions. Big Data analytics is now enabled by the advancements in technologies that aid in storage and processing of vast amounts of rapidly growing data.

Source: Asterdata.com, "Gilt Groupe Speaks on Digital Marketing Optimization," **asterdata.com/gilt_groupe_video.php** (accessed February 2013).

1.10 PLAN OF THE BOOK

The previous sections have given you an understanding of the need for using information technology in decision making; an IT-oriented view of various types of decisions; and the evolution of decision support systems into business intelligence, and now into analytics. In the last two sections we have seen an overview of various types of analytics and their applications. Now we are ready for a more detailed managerial excursion into these topics, along with some potentially deep hands-on experience in some of the technical topics. The 14 chapters of this book are organized into five parts, as shown in Figure 1.6.

Part I: Business Analytics: An Overview

In Chapter 1, we provided an introduction, definitions, and an overview of decision support systems, business intelligence, and analytics, including Big Data analytics. Chapter 2 covers the basic phases of the decision-making process and introduces decision support systems in more detail.

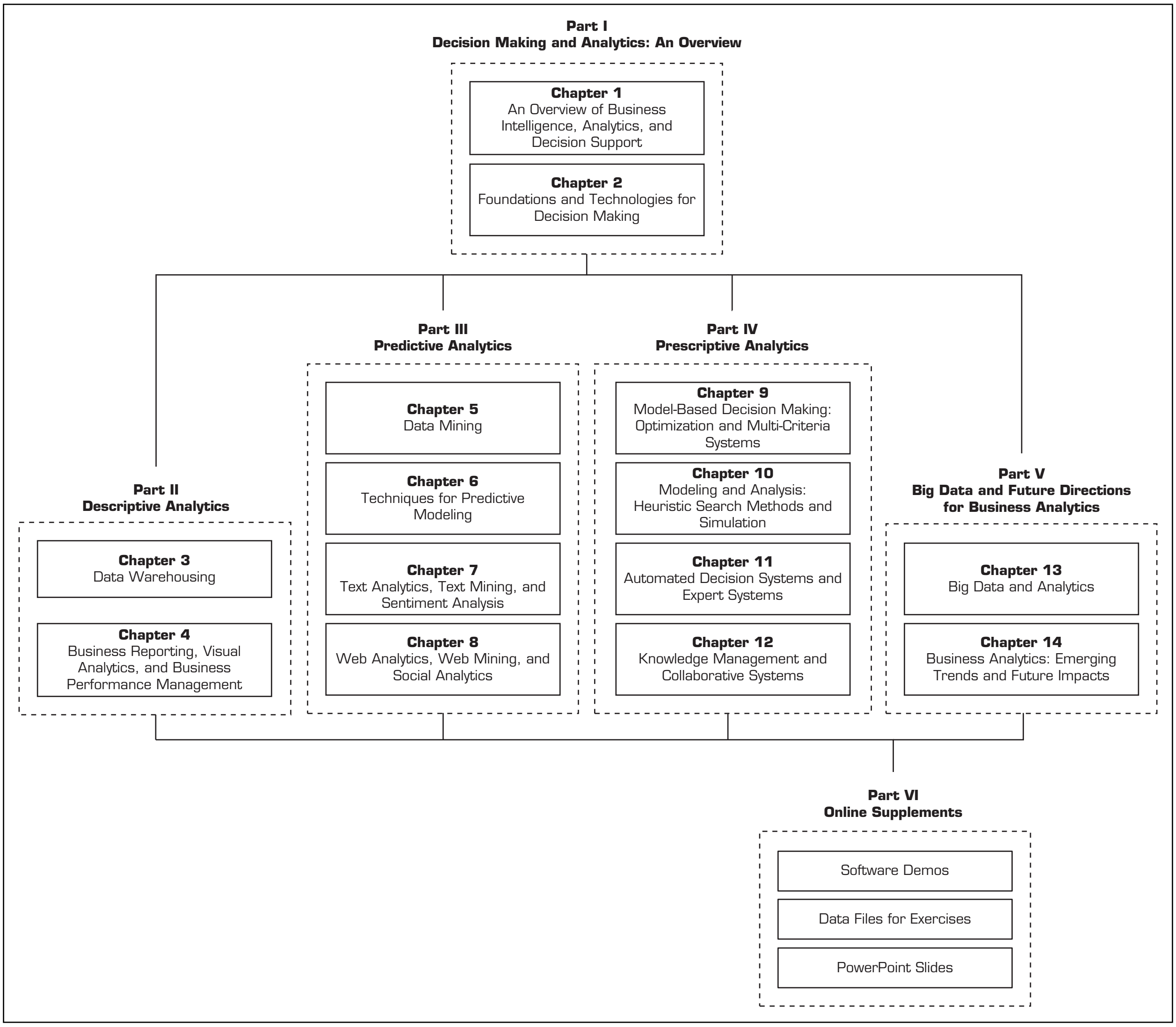

FIGURE 1.6 Plan of the Book.

Part II: Descriptive Analytics

Part II begins with an introduction to data warehousing issues, applications, and technologies in Chapter 3. Data represent the fundamental backbone of any decision support and analytics application. Chapter 4 describes business reporting, visualization technologies, and applications. It also includes a brief overview of business performance management techniques and applications, a topic that has been a key part of traditional BI.

Part III: Predictive Analytics

Part III comprises a large part of the book. It begins with an introduction to predictive analytics applications in Chapter 5. It includes many of the common application techniques: classification, clustering, association mining, and so forth. Chapter 6 includes a technical description of selected data mining techniques, especially neural network models. Chapter 7 focuses on text mining applications. Similarly, Chapter 8 focuses on Web analytics, including social media analytics, sentiment analysis, and other related topics.

Part IV: Prescriptive Analytics

Part IV introduces decision analytic techniques, which are also called prescriptive analytics. Specifically, Chapter 9 covers selected models that may be implemented in spreadsheet environments. It also covers a popular multi-objective decision technique—analytic hierarchy processes.

Chapter 10 then introduces other model-based decision-making techniques, especially heuristic models and simulation. Chapter 11 introduces automated decision systems including expert systems. This part concludes with a brief discussion of knowledge management and group support systems in Chapter 12.

Part V: Big Data and Future Directions for Business Analytics

Part V begins with a more detailed coverage of Big Data and analytics in Chapter 13.

Chapter 14 attempts to integrate all the material covered in this book and concludes with a discussion of emerging trends, such as how the ubiquity of wireless and GPS devices and other sensors is resulting in the creation of massive new databases and unique applications. A new breed of data mining and BI companies is emerging to analyze these new databases and create a much better and deeper understanding of customers' behaviors and movements. The chapter also covers cloud-based analytics, recommendation systems, and a brief discussion of security/privacy dimensions of analytics. It concludes the book by also presenting a discussion of the analytics ecosystem. An understanding of the ecosystem and the various players in the analytics industry highlights the various career opportunities for students and practitioners of analytics.

1.11 RESOURCES, LINKS, AND THE TERADATA UNIVERSITY NETWORK CONNECTION

The use of this chapter and most other chapters in this book can be enhanced by the tools described in the following sections.

Resources and Links

We recommend the following major resources and links:

- The Data Warehousing Institute (**tdwi.org**)
- Information Management (**information-management.com**)
- DSS Resources (**dssresources.com**)
- Microsoft Enterprise Consortium (**enterprise.waltoncollege.uark.edu/mec.asp**)

Vendors, Products, and Demos

Most vendors provide software demos of their products and applications. Information about products, architecture, and software is available at **dssresources.com**.

Periodicals

We recommend the following periodicals:

- *Decision Support Systems*
- *CIO Insight* (**cioinsight.com**)
- *Technology Evaluation* (**technologyevaluation.com**)
- *Baseline Magazine* (**baselinemag.com**)

The Teradata University Network Connection

This book is tightly connected with the free resources provided by Teradata University Network (TUN; see **teradatauniversitynetwork.com**). The TUN portal is divided into two major parts: one for students and one for faculty. This book is connected to the TUN portal via a special section at the end of each chapter. That section includes appropriate links for the specific chapter, pointing to relevant resources. In addition, we provide hands-on exercises, using software and other material (e.g., cases) available at TUN.

The Book's Web Site

This book's Web site, **pearsonglobaleditions.com/turban**, contains supplemental textual material organized as Web chapters that correspond to the printed book's chapters. The topics of these chapters are listed in the online chapter table of contents. Other content is also available on an independent Web site (**dssbibook.com)**.[2]

Chapter Highlights

- The business environment is becoming complex and is rapidly changing, making decision making more difficult.
- Businesses must respond and adapt to the changing environment rapidly by making faster and better decisions.
- The time frame for making decisions is shrinking, whereas the global nature of decision making is expanding, necessitating the development and use of computerized DSS.
- Computerized support for managers is often essential for the survival of an organization.
- An early decision support framework divides decision situations into nine categories, depending on the degree of structuredness and managerial activities. Each category is supported differently.
- Structured repetitive decisions are supported by standard quantitative analysis methods, such as MS, MIS, and rule-based automated decision support.
- DSS use data, models, and sometimes knowledge management to find solutions for semistructured and some unstructured problems.
- BI methods utilize a central repository called a data warehouse that enables efficient data mining, OLAP, BPM, and data visualization.
- BI architecture includes a data warehouse, business analytics tools used by end users, and a user interface (such as a dashboard).
- Many organizations employ descriptive analytics to replace their traditional flat reporting with interactive reporting that provides insights, trends, and patterns in the transactional data.
- Predictive analytics enable organizations to establish predictive rules that drive the business outcomes through historical data analysis of the existing behavior of the customers.
- Prescriptive analytics help in building models that involve forecasting and optimization techniques based on the principles of operations research and management science to help organizations to make better decisions.
- Big Data analytics focuses on unstructured, large data sets that may also include vastly different types of data for analysis.
- Analytics as a field is also known by industry-specific application names such as sports analytics. It is also known by other related names such as data science or network science.

[2]As this book went to press, we verified that all the cited Web sites were active and valid. However, URLs are dynamic. Web sites to which we refer in the text sometimes change or are discontinued because companies change names, are bought or sold, merge, or fail. Sometimes Web sites are down for maintenance, repair, or redesign. Many organizations have dropped the initial "www" designation for their sites, but some still use it. If you have a problem connecting to a Web site that we mention, please be patient and simply run a Web search to try to identify the possible new site. Most times, you can quickly find the new site through one of the popular search engines. We apologize in advance for this inconvenience.

Key Terms

business intelligence (BI)
dashboard
data mining
decision (or normative) analytics
decision support system (DSS)
descriptive (or reporting) analytics
predictive analytics
prescriptive analytics
semistructured problem
structured problem
unstructured problem

Questions for Discussion

1. Distinguish between strategic and tactical planning?
2. What is data mining and why is it classified under predictive analytics? Search the Web for an example of data mining in an organization of your choice and illustrate the way it is currently in use.
3. Prescriptive analytics is considered to be a step further ahead of predictive analysis and substantially different from it. Provide an example of each and outline their differences.
4. Provide a definition of BI.
5. Define managerial decision making. Discuss this concept in the context of the four-step approach to decision making.

Exercises

Teradata University Network (TUN) and Other Hands-On Exercises

1. Go to **teradatauniversitynetwork.com**. Using the registration your instructor provides, log on and learn the content of the site. You will receive assignments related to this site. Prepare a list of 20 items in the site that you think could be beneficial to you.
2. Enter the TUN site and select "cases, projects and assignments." Then select the case study: "Harrah's High Payoff from Customer Information." Answer the following questions about this case:
   a. What information does the data mining generate?
   b. How is this information helpful to management in decision making? (Be specific.)
   c. List the types of data that are mined.
   d. Is this a DSS or BI application? Why?
3. Go to **teradatauniversitynetwork.com** and find the paper titled "Data Warehousing Supports Corporate Strategy at First American Corporation" (by Watson, Wixom, and Goodhue). Read the paper and answer the following questions:
   a. What were the drivers for the DW/BI project in the company?
   b. What strategic advantages were realized?
   c. What operational and tactical advantages were achieved?
   d. What were the critical success factors (CSF) for the implementation?
4. Go to **analytics-magazine.org/issues/digital-editions** and find the January/February 2012 edition titled "Special Issue: The Future of Healthcare." Read the article "Predictive Analytics—Saving Lives and Lowering Medical Bills." Answer the following questions:
   a. What is the problem that is being addressed by applying predictive analytics?
   b. What is the FICO Medication Adherence Score?
   c. How is a prediction model trained to predict the FICO Medication Adherence Score? Did the prediction model classify FICO Medication Adherence Score?
   d. Zoom in on Figure 4 and explain what kind of technique is applied on the generated results.
   e. List some of the actionable decisions that were based on the results of the predictions.
5. Visit **http://www.ibm.com/analytics/us/en/what-is-smarter-analytics/big-data-analysis.html.** Read the section "Gain actionable insights from big data analysis," and watch the video "See how analytics can help organizations in all industries use big data to achieve significant outcomes" under Case Studies to answer the following questions:
   a. Explain *big data* and its importance in decision making with examples.
   b. Appraise the maxim "without analytics, big data is just noise."
   c. In which sectors and areas might big data analytics be most useful? Give examples.
   d. What are the suggested solutions to managing risks?
   e. Review and analyze a case study from IBM's Web site and explain how big data analytics provided solutions.
6. Business analytics and computerized data processing support managers and decision making. Keeping current

business environment challenges in mind, along with Mintzberg's 10 managerial roles (see Table 1.2), identify five such roles in companies/government's press release and communications.

7. Go to oracle.com, a leading company in BI. Make a map of their Web site illustrating their products and communication styles with available resources for business.
8. Search the Web for a company that uses the four major components of a standard BI system.
9. In the company identified in the previous question, illustrate their main products and style of BI and discuss the main tools used. Refer to the tools mentioned in this chapter.

End-of-Chapter Application Case

Nationwide Insurance Used BI to Enhance Customer Service

Nationwide Mutual Insurance Company, headquartered in Columbus, Ohio, is one of the largest insurance and financial services companies, with $23 billion in revenues and more than $160 billion in statutory assets. It offers a comprehensive range of products through its family of 100-plus companies with insurance products for auto, motorcycle, boat, life, homeowners, and farms. It also offers financial products and services including annuities, mortgages, mutual funds, pensions, and investment management.

Nationwide strives to achieve greater efficiency in all operations by managing its expenses along with its ability to grow its revenue. It recognizes the use of its strategic asset of information combined with analytics to outpace competitors in strategic and operational decision making even in complex and unpredictable environments.

Historically, Nationwide's business units worked independently and with a lot of autonomy. This led to duplication of efforts, widely dissimilar data processing environments, and extreme data redundancy, resulting in higher expenses. The situation got complicated when Nationwide pursued any mergers or acquisitions.

Nationwide, using enterprise data warehouse technology from Teradata, set out to create, from the ground up, a single, authoritative environment for clean, consistent, and complete data that can be effectively used for best-practice analytics to make strategic and tactical business decisions in the areas of customer growth, retention, product profitability, cost containment, and productivity improvements. Nationwide transformed its siloed business units, which were supported by stove-piped data environments, into integrated units by using cutting-edge analytics that work with clear, consolidated data from all of its business units. The Teradata data warehouse at Nationwide has grown from 400 gigabytes to more than 100 terabytes and supports 85 percent of Nationwide's business with more than 2,500 users.

Integrated Customer Knowledge

Nationwide's Customer Knowledge Store (CKS) initiative developed a customer-centric database that integrated customer, product, and externally acquired data from more than 48 sources into a single customer data mart to deliver a holistic view of customers. This data mart was coupled with Teradata's customer relationship management application to create and manage effective customer marketing campaigns that use behavioral analysis of customer interactions to drive customer management actions (CMAs) for target segments. Nationwide added more sophisticated customer analytics that looked at customer portfolios and the effectiveness of various marketing campaigns. This data analysis helped Nationwide to initiate proactive customer communications around customer lifetime events like marriage, birth of child, or home purchase and had significant impact on improving customer satisfaction. Also, by integrating customer contact history, product ownership, and payment information, Nationwide's behavioral analytics teams further created prioritized models that could identify which specific customer interaction was important for a customer at any given time. This resulted in one percentage point improvement in customer retention rates and significant improvement in customer enthusiasm scores. Nationwide also achieved 3 percent annual growth in incremental sales by using CKS. There are other uses of the customer database. In one of the initiatives, by integrating customer telephone data from multiple systems into CKS, the relationship managers at Nationwide try to be proactives in contacting customers in advance of a possible weather catastrophe, such as a hurricane or flood, to provide the primary policyholder information and explain the claims processes. These and other analytic insights now drive Nationwide to provide extremely personal customer service.

Financial Operations

A similar performance payoff from integrated information was also noted in financial operations. Nationwide's decentralized management style resulted in a fragmented financial reporting environment that included more than 14 general ledgers, 20 charts of accounts, 17 separate data repositories, 12 different reporting tools, and hundreds of thousands of spreadsheets. There was no common central view of the business, which resulted in labor-intensive slow and inaccurate reporting.

About 75 percent of the effort was spent on acquiring, cleaning, and consolidating and validating the data, and very little time was spent on meaningful analysis of the data.

The Financial Performance Management initiative implemented a new operating approach that worked on a single data and technology architecture with a common set of systems standardizing the process of reporting. It enabled Nationwide to operate analytical centers of excellence with world-class planning, capital management, risk assessment, and other decision support capabilities that delivered timely, accurate, and efficient accounting, reporting, and analytical services.

The data from more than 200 operational systems was sent to the enterprise-wide data warehouse and then distributed to various applications and analytics. This resulted in a 50 percent improvement in the monthly closing process with closing intervals reduced from 14 days to 7 days.

Postmerger Data Integration

Nationwide's Goal State Rate Management initiative enabled the company to merge Allied Insurance's automobile policy system into its existing system. Both Nationwide and Allied source systems were custom-built applications that did not share any common values or process data in the same manner. Nationwide's IT department decided to bring all the data from source systems into a centralized data warehouse, organized in an integrated fashion that resulted in standard dimensional reporting and helped Nationwide in performing what-if analyses. The data analysis team could identify previously unknown potential differences in the data environment where premiums rates were calculated differently between Nationwide and Allied sides. Correcting all of these benefited Nationwide's policyholders because they were safeguarded from experiencing wide premium rate swings.

Enhanced Reporting

Nationwide's legacy reporting system, which catered to the needs of property and casualty business units, took weeks to compile and deliver the needed reports to the agents. Nationwide determined that it needed better access to sales and policy information to reach its sales targets. It chose a single data warehouse approach and, after careful assessment of the needs of sales management and individual agents, selected a business intelligence platform that would integrate dynamic enterprise dashboards into its reporting systems, making it easy for the agents and associates to view policy information at a glance. The new reporting system, dubbed Revenue Connection, also enabled users to analyze the information with a lot of interactive and drill-down-to-details capabilities at various levels that eliminated the need to generate custom ad hoc reports. Revenue Connection virtually eliminated requests for manual policy audits, resulting in huge savings in time and money for the business and technology teams. The reports were produced in 4 to 45 seconds, rather than days or weeks, and productivity in some units improved by 20 to 30 percent.

Questions for Discussion

1. Why did Nationwide need an enterprise-wide data warehouse?
2. How did integrated data drive the business value?
3. What forms of analytics are employed at Nationwide?
4. With integrated data available in an enterprise data warehouse, what other applications could Nationwide potentially develop?

What We Can Learn from This Application Case

The proper use of integrated information in organizations can help achieve better business outcomes. Many organizations now rely on data warehousing technologies to perform the online analytical processes on the data to derive valuable insights. The insights are used to develop predictive models that further enable the growth of the organizations by more precisely assessing customer needs. Increasingly, organizations are moving toward deriving value from analytical applications in real time with the help of integrated data from real-time data warehousing technologies.

Source: Teradata.com, "Nationwide, Delivering an On Your Side Experience," **teradata.com/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=14714** (accessed February 2013).

References

Anthony, R. N. (1965). *Planning and Control Systems: A Framework for Analysis*. Cambridge, MA: Harvard University Graduate School of Business.

Asterdata.com. "Gilt Groupe Speaks on Digital Marketing Optimization." **www.asterdata.com/gilt_groupe_video.php** (accessed February 2013).

Barabasilab.neu.edu. "Network Science." **barabasilab.neu.edu/networksciencebook/downlPDF.html** (accessed February 2013).

Brooks, D. (2009, May 18). "In Praise of Dullness." *New York Times,* **nytimes.com/2009/05/19/opinion/19brooks.html** (accessed February 2013).

Centers for Disease Control and Prevention, Vaccines for Children Program, "Module 6 of the VFC Operations Guide." **cdc.gov/vaccines/pubs/pinkbook/vac-storage.html#storage** (accessed January 2013).

Eckerson, W. (2003). *Smart Companies in the 21st Century: The Secrets of Creating Successful Business Intelligent Solutions*. Seattle, WA: The Data Warehousing Institute.

Emc.com. "Data Science Revealed: A Data-Driven Glimpse into the Burgeoning New Field." **emc.com/collateral/about/news/emc-data-science-study-wp.pdf** (accessed February 2013).

Gorry, G. A., and M. S. Scott-Morton. (1971). "A Framework for Management Information Systems." *Sloan Management Review*, Vol. 13, No. 1, pp. 55–70.

INFORMS. "Analytics Section Overview." **informs.org/Community/Analytics** (accessed February 2013).

Keen, P. G. W., and M. S. Scott-Morton. (1978). *Decision Support Systems: An Organizational Perspective*. Reading, MA: Addison-Wesley.

Krivda, C. D. (2008, March). "Dialing Up Growth in a Mature Market." *Teradata Magazine*, pp. 1–3.

Magpiesensing.com. "MagpieSensing Cold Chain Analytics and Monitoring." **magpiesensing.com/wp-content/uploads/2013/01/ColdChainAnalyticsMagpieSensing-Whitepaper.pdf** (accessed January 2013).

Mintzberg, H. A. (1980). *The Nature of Managerial Work*. Englewood Cliffs, NJ: Prentice Hall.

Mintzberg, H. A. (1993). *The Rise and Fall of Strategic Planning*. New York: The Free Press.

Simon, H. (1977). *The New Science of Management Decision*. Englewood Cliffs, NJ: Prentice Hall.

Tableausoftware.com. "Eliminating Waste at Seattle Children's." **tableausoftware.com/eliminating-waste-at-seattle-childrens** (accessed February 2013).

Tableausoftware.com. "Kaleida Health Finds Efficiencies, Stays Competitive." **tableausoftware.com/learn/stories/user-experience-speed-thought-kaleida-health** (accessed February 2013).

Teradata.com. "Nationwide, Delivering an On Your Side Experience." **teradata.com/case-studies/delivering-on-your-side-experience** (accessed February 2013).

Teradata.com. "Sabre Airline Solutions." **teradata.com/t/case-studies/Sabre-Airline-Solutions-EB6281** (accessed February 2013).

Wang, X., et al. (2012, January/February). "Branch Reconfiguration Practice Through Operations Research in Industrial and Commercial Bank of China." *Interfaces*, Vol. 42, No. 1, pp. 33–44.

Watson, H. (2005, Winter). "Sorting Out What's New in Decision Support." *Business Intelligence Journal*.

Wikipedia. "On Base Percentage." **en.wikipedia.org/wiki/On_base_percentage** (accessed January 2013).

Wikipedia. "Sabermetrics." **en.wikipedia.org/wiki/Sabermetrics** (accessed January 2013).

Zaleski, A. (2012). "Magpie Analytics System Tracks Cold-Chain Products to Keep Vaccines, Reagents Fresh." **TechnicallyBaltimore.com** (accessed February 2013).

Zaman, M. (2009, April). "Business Intelligence: Its Ins and Outs." **technologyevaluation.com** (accessed February 2013).

Ziama, A., and J. Kasher. (2004). "Data Mining Primer for the Data Warehousing Professional." Dayton, OH: *Teradata*.

CHAPTER

2

Foundations and Technologies for Decision Making

LEARNING OBJECTIVES

- Understand the conceptual foundations of decision making
- Understand Simon's four phases of decision making: intelligence, design, choice, and implementation
- Understand the essential definition of DSS
- Understand important DSS classifications
- Learn how DSS support for decision making can be provided in practice
- Understand DSS components and how they integrate

Our major focus in this book is the support of decision making through computer-based information systems. The purpose of this chapter is to describe the conceptual foundations of decision making and how decision support is provided. This chapter includes the following sections:

2.1 OPENING VIGNETTE: Decision Modeling at HP Using Spreadsheets

HP is a major manufacturer of computers, printers, and many industrial products. Its vast product line leads to many decision problems. Olavson and Fry (2008) have worked on many spreadsheet models for assisting decision makers at HP and have identified several lessons from both their successes and their failures when it comes to constructing and applying spreadsheet-based tools. They define a *tool* as "a reusable, analytical solution designed to be handed off to nontechnical end users to assist them in solving a repeated business problem."

When trying to solve a problem, HP developers consider the three phases in developing a model. The first phase is problem framing, where they consider the following questions in order to develop the best solution for the problem:

- Will analytics solve the problem?
- Can an existing solution be leveraged?
- Is a tool needed?

The first question is important because the problem may not be of an analytic nature, and therefore, a spreadsheet tool may not be of much help in the long run without fixing the nonanalytical part of the problem first. For example, many inventory-related issues arise because of the inherent differences between the goals of marketing and supply chain groups. Marketing likes to have the maximum variety in the product line, whereas supply chain management focuses on reducing the inventory costs. This difference is partially outside the scope of any model. Coming up with nonmodeling solutions is important as well. If the problem arises due to "misalignment" of incentives or unclear lines of authority or plans, no model can help. Thus, it is important to identify the root issue.

The second question is important because sometimes an existing tool may solve a problem that then saves time and money. Sometimes modifying an existing tool may solve the problem, again saving some time and money, but sometimes a custom tool is necessary to solve the problem. This is clearly worthwhile to explore.

The third question is important because sometimes a new computer-based system is not required to solve the problem. The developers have found that they often use analytically derived decision guidelines instead of a tool. This solution requires less time for development and training, has lower maintenance requirements, and also provides simpler and more intuitive results. That is, after they have explored the problem deeper, the developers may determine that it is better to present decision rules that can be easily implemented as guidelines for decision making rather than asking the managers to run some type of a computer model. This results in easier training, better understanding of the rules being proposed, and increased acceptance. It also typically leads to lower development costs and reduced time for deployment.

If a model has to be built, the developers move on to the second phase—the actual design and development of the tools. Adhering to five guidelines tends to increase the probability that the new tool will be successful. The first guideline is to develop a prototype as quickly as possible. This allows the developers to test the designs, demonstrate various features and ideas for the new tools, get early feedback from the end users to see what works for them and what needs to be changed, and test adoption. Developing a prototype also prevents the developers from overbuilding the tool and yet allows them to construct more scalable and standardized software applications later. Additionally, by developing a prototype, developers can stop the process once the tool is "good enough," rather than building a standardized solution that would take longer to build and be more expensive.