CAMM • COCHRAN • FRY • OHLMANN
ANDERSON • SWEENEY • WILLIAMS

Essentials of Business Analytics

CENGAGE Learning

Essentials of Business Analytics,
First Edition

Camm, Cochran, Fry, Ohlmann, Anderson, Sweeney, Williams

Senior Vice President, Global Product Manager, Higher Education: Jack W. Calhoun

Vice President, General Manager, Social Science & Qualitative Business: Erin Joyner

Product Director: Joe Sabatino

Content Developer: Maggie Kubale

Product Assistant: Brad Sullender

Content Project Manager: Cliff Kallemeyn

Media Developer: Chris Valentine

Manufacturing Planner: Ron Montgomery

Marketing Director: Natalie King

Marketing Manager: Heather Mooney

Associate Market Development Manager: Roy Rosa

Production Service: MPS Limited

Sr. Art Director: Stacy Jenkins Shirley

Internal Designer: Mike Stratton/ Chris Miller

Cover Designer: Craig Ramsdell

Cover Image: © Kutlayev Dmitry/ Shutterstock

Sr. Rights Acquisitions Specialist: John Hill

Library of Congress Control Number: 2013944175

ISBN-13: 978-1-285-18727-3

ISBN-10: 1-285-18727-X

Cengage Learning
200 First Stamford Place, 4th Floor
Stamford, CT 06902
USA

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

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

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

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

Printed in the United States of America
1 2 3 4 5 6 7 17 16 15 14 13

Brief Contents

Contents

Chapter 12 Decision Analysis 550

About the Authors

Jeffrey D. Camm Jeffrey D. Camm is Professor of Quantitative Analysis, Head of the Department of Operations, Business Analytics, and Information Systems and College of Business Research Fellow in the Carl H. Lindner College of Business at the University of Cincinnati. Born in Cincinnati, Ohio, he holds a B.S. from Xavier University and a Ph.D. from Clemson University. He has been at the University of Cincinnati since 1984, and has been a visiting scholar at Stanford University and a visiting professor of business administration at the Tuck School of Business at Dartmouth College.

Dr. Camm has published over 30 papers in the general area of optimization applied to problems in operations management. He has published his research in *Science, Management Science, Operations Research, Interfaces* and other professional journals. At the University of Cincinnati, he was named the Dornoff Fellow of Teaching Excellence and he was the 2006 recipient of the INFORMS Prize for the Teaching of Operations Research Practice. A firm believer in practicing what he preaches, he has served as an operations research consultant to numerous companies and government agencies. From 2005 to 2010 he served as editor-in-chief of *Interfaces,* and is currently on the editorial board of *INFORMS Transactions on Education*.

James J. Cochran James J. Cochran is the Bank of Ruston Barnes, Thompson, & Thurmon Endowed Research Professor of Quantitative Analysis at Louisiana Tech University. Born in Dayton, Ohio, he holds a B.S., an M.S., and an M.B.A. from Wright State University and a Ph.D. from the University of Cincinnati. He has been at Louisiana Tech University since 2000, and has been a visiting scholar at Stanford University, Universidad de Talca, and the University of South Africa.

Professor Cochran has published over two dozen papers in the development and application of operations research and statistical methods. He has published his research in *Management Science, The American Statistician, Communications in Statistics—Theory and Methods, European Journal of Operational Research, Journal of Combinatorial Optimization,* and other professional journals. He was the 2008 recipient of the INFORMS Prize for the Teaching of Operations Research Practice and the 2010 recipient of the Mu Sigma Rho Statistical Education Award. Professor Cochran was elected to the International Statistics Institute in 2005 and named a Fellow of the American Statistical Association in 2011. A strong advocate for effective operations research and statistics education as a means of improving the quality of applications to real problems, Professor Cochran has organized and chaired teaching effectiveness workshops in Montevideo, Uruguay; Cape Town, South Africa; Cartagena, Colombia; Jaipur, India; Buenos Aires, Argentina; and Nairobi, Kenya. He has served as an operations research consultant to numerous companies and not-for-profit organizations. He currently serves as editor-in-chief of *INFORMS Transactions on Education* and is on the editorial board of *Interfaces,* the *Journal of the Chilean Institute of Operations Research,* and *ORiON*.

Michael J. Fry Michael J. Fry is Associate Professor of Operations, Business Analytics, and Information Systems in the Carl H. Lindner College of Business at the University of Cincinnati. Born in Killeen, Texas, he earned a B.S. from Texas A&M University, and M.S.E. and Ph.D. degrees from the University of Michigan. He has been at the University of Cincinnati since 2002, and he has been a visiting professor at The Johnson School at Cornell University and the Sauder School of Business at the University of British Columbia.

Professor Fry has published over a dozen research papers in journals such as *Operations Research, M&SOM, Transportation Science, Naval Research Logistics,* and *Interfaces*. His research interests are in applying quantitative management methods to the areas of supply chain analytics, sports analytics, and public-policy operations. He has worked with many different organizations for his research, including Dell, Inc., Copeland Corporation, Starbucks Coffee Company, the Cincinnati Fire Department, the State of Ohio Election Commission, the Cincinnati Bengals, and the Cincinnati Zoo. In 2008, he was named a finalist for the Daniel H. Wagner Prize for Excellence in Operations Research Practice, and he has been recognized for both his research and teaching excellence at the University of Cincinnati.

Jeffrey W. Ohlmann Jeffrey W. Ohlmann is Associate Professor of Management Sciences in the Tippie College of Business at the University of Iowa. Born in Valentine, Nebraska, he earned a B.S. from the University of Nebraska, and M.S. and Ph.D. degrees from the University of Michigan. He has been at the University of Iowa since 2003.

Professor Ohlmann's research on the modeling and solution of decision-making problems has produced over a dozen research papers in journals such as *Mathematics of Operations Research, INFORMS Journal on Computing, Transportation Science,* and *Interfaces*. He has collaborated with companies such as Transfreight, LeanCor, Cargill, the Hamilton County Board of Elections, and the Cincinnati Bengals. Due to the relevance of his work to industry, he was bestowed the George B. Dantzig Dissertation Award and was recognized as a finalist for the Daniel H. Wagner Prize for Excellence in Operations Research Practice.

David R. Anderson David R. Anderson is Professor Emeritus of Quantitative Analysis in the Carl H. Lindner College of Business at the University of Cincinnati. Born in Grand Forks, North Dakota, he earned his B.S., M.S., and Ph.D. degrees from Purdue University. Professor Anderson has served as Head of the Department of Quantitative Analysis and Operations Management and as Associate Dean of the College of Business Administration. In addition, he was the coordinator of the College's first Executive Program.

At the University of Cincinnati, Professor Anderson has taught introductory statistics for business students as well as graduate-level courses in regression analysis, multivariate analysis, and management science. He has also taught statistical courses at the Department of Labor in Washington, D.C. He has been honored with nominations and awards for excellence in teaching and excellence in service to student organizations.

Professor Anderson has coauthored 10 textbooks in the areas of statistics, management science, linear programming, and production and operations management. He is an active consultant in the field of sampling and statistical methods.

Dennis J. Sweeney Dennis J. Sweeney is Professor Emeritus of Quantitative Analysis and Founder of the Center for Productivity Improvement at the University of Cincinnati. Born in Des Moines, Iowa, he earned a B.S.B.A. degree from Drake University and his M.B.A. and D.B.A. degrees from Indiana University, where he was an NDEA Fellow. During 1978–1979, Professor Sweeney worked in the management science group at Procter & Gamble; during 1981–1982, he was a visiting professor at Duke University. Professor Sweeney served as Head of the Department of Quantitative Analysis and as Associate Dean of the College of Business Administration at the University of Cincinnati.

Professor Sweeney has published more than 30 articles and monographs in the areas of management science and statistics. The National Science Foundation, IBM, Procter & Gamble, Federated Department Stores, Kroger, and Cincinnati Gas & Electric have funded his research, which has been published in *Management Science, Operations Research, Mathematical Programming, Decision Sciences,* and other journals.

Professor Sweeney has coauthored 10 textbooks in the areas of statistics, management science, linear programming, and production and operations management.

Thomas A. Williams Thomas A. Williams is Professor Emeritus of Management Science in the College of Business at Rochester Institute of Technology. Born in Elmira, New York, he earned his B.S. degree at Clarkson University. He did his graduate work at Rensselaer Polytechnic Institute, where he received his M.S. and Ph.D. degrees.

Before joining the College of Business at RIT, Professor Williams served for seven years as a faculty member in the College of Business Administration at the University of Cincinnati, where he developed the undergraduate program in Information Systems and then served as its coordinator. At RIT he was the first chairman of the Decision Sciences Department. He teaches courses in management science and statistics, as well as graduate courses in regression and decision analysis.

Professor Williams is the coauthor of 11 textbooks in the areas of management science, statistics, production and operations management, and mathematics. He has been a consultant for numerous *Fortune* 500 companies and has worked on projects ranging from the use of data analysis to the development of large-scale regression models.

Preface

Essentials of Business Analytics is designed to introduce the concept of business analytics to undergraduate and graduate students. This textbook contains one of the first collections of materials that are essential to the growing field of business analytics. In Chapter 1 we present an overview of business analytics and our approach to the material in this textbook. In simple terms, business analytics helps business professionals make better decisions based on data. We discuss models for summarizing, visualizing, and understanding useful information from historical data in Chapters 2 through 6. Chapter 7 covers the use of spreadsheets for examining data and building decision models. In Chapters 8 through 10 we discuss optimization models to help decision makers choose the best decision based on the available data. Chapter 10 presents material that some may consider more advanced forms of optimization (nonlinear optimization models), although these models are extremely useful and widely applicable to many business situations. In any case, some instructors may choose to omit covering Chapter 10. In Chapter 11 we introduce the concept of simulation models for understanding the effect of uncertainty on decisions. Chapter 12 is an overview of decision analysis approaches for incorporating a decision maker's views about risk into decision making. In Appendix A we present optional material for students who need to learn the basics of using Microsoft Excel. The use of databases and manipulating data in Microsoft Access is discussed in Appendix B.

This textbook can be used by students who have previously taken a course on basic statistical methods as well as students who have not had a prior course in statistics. This textbook introduces basic statistical concepts in enough detail to support their use in business analytics tools. For the student who has not had a prior statistics course, these concepts are sufficient to prepare the student for more advanced business analytics methods. For students who have had a previous statistics class, the material will provide a good review. All statistical concepts contained in this textbook are presented from a business analytics perspective using practical business examples. For those instructors who wish to skip the introductory statistics material, Chapters 2 and 4 can be considered optional.

Features and Pedagogy

The style and format of this textbook is based on the other classic textbooks written by the Anderson, Sweeney, and Williams (ASW) team. Some of the specific features that we use in this textbook are listed below.

- Integration of Microsoft Excel: Excel has been thoroughly integrated throughout this textbook. For many methodologies, we provide instructions for how to perform calculations both by hand and with Excel. In other cases where realistic models are practical only with the use of a spreadsheet, we focus on the use of Excel to describe the methods to be used.
- Use of Excel 2013: The material presented for Excel in this textbook is fully compatible with Excel 2013. In most cases, Excel 2013 can be considered a relatively minor update from previous Excel versions as it relates to business analytics. However, the data visualization abilities of Excel have been greatly enhanced in Excel 2013. It is much easier to create, modify and analyze charts in Excel 2013.

Recognizing that many students and instructors may not have access to Excel 2013 at this time, we also provide instructions for using previous versions of Excel whenever possible.

- Use of Analytics Solver Platform and XLMiner: This textbook incorporates the use of two very powerful Microsoft Excel Add-ins: Analytics Solver Platform and XLMiner, both created by Frontline Systems. Analytics Solver Platform provides additional optimization and simulation features for Excel. XLMiner incorporates sophisticated data mining algorithms into Excel and allows for additional data visualization and data exploration. In most chapters we place the use of Analytics Solver Platform and XLMiner in the chapter appendix so that the instructor can choose whether or not to cover this material. However, because these tools are essential to performing simulation and data mining methods, we integrate XLMiner throughout Chapter 6 on data mining and we utilize Analytics Solver Platform in Sections 11.3 and 11.4 for simulation.
- Notes and Comments: At the end of many sections, we provide Notes and Comments to give the student additional insights about the methods presented in that section. These insights include comments on the limitations of the presented methods, recommendations for applications, and other matters. Additionally, margin notes are used throughout the textbook to provide additional insights and tips related to the specific material being discussed.
- Analytics in Action: Each chapter contains an Analytics in Action article. These articles present interesting examples of the use of business analytics in practice. The examples are drawn from many different organizations in a variety of areas including healthcare, finance, manufacturing, marketing, and others.
- WEBfiles: All data sets used as examples and in student exercises are also provided online as files available for download by the student. The names of the WEBfiles are called out in margin notes throughout the textbook.
- Problems and Cases: With the exception of Chapter 1, each chapter contains more than 20 problems to help the student master the material presented in that chapter. The problems vary in difficulty and most relate to specific examples of the use of business analytics in practice. Answers to even-numbered problems are provided in an online supplement for student access. With the exception of Chapter 1, each chapter also includes an in-depth case study that connects many of the different methods introduced in the chapter. The case studies are designed to be more open-ended than the chapter problems, but enough detail is provided to give the student some direction in solving the cases.

Acknowledgements

We would like to acknowledge the work of our reviewers, who provided comments and suggestions for improvement of this text. Thanks to:

Matthew D. Bailey
Bucknell University

Phillip Beaver
Daniels College of Business
University of Denver

M. Khurrum S. Bhutta
Ohio University

Q B. Chung
Villanova University

Elizabeth A. Denny
University of Kentucky

Mike Taein Eom
University of Portland

Yvette Njan Essounga
Fayetteville State University

Lawrence V. Fulton
Texas State University

James F. Hoelscher
Lincoln Memorial University

Eric Huggins
Fort Lewis College

Faizul Huq
Ohio University

Marco Lam
York College of Pennsylvania

Ram Pakath
University of Kentucky

Susan Palocsay
James Madison University

Dothan Truong
Embry-Riddle Aeronautical University

Kai Wang
Wake Technical Community College

We are indebted to our product director Joe Sabatino and our product manager, Aaron Arnsparger; our marketing director, Natalie King, our marketing manager, Heather Mooney, and our associate marketing development manager, Roy Rosa; our content developer, Maggie Kubale; our senior content project manager, Cliff Kallemeyn; our media developer, Chris Valentine; and others at Cengage Learning for their counsel and support during the preparation of this text.

Jeffrey D. Camm
James J. Cochran
Michael J. Fry
Jeffrey W. Ohlmann
David R. Anderson
Dennis J. Sweeney
Thomas A. Williams

CHAPTER 1

Introduction

CONTENTS

You apply for a loan for the first time. How does the bank assess the riskiness of the loan it might make to you? How does Amazon.com know which books and other products to recommend to you when you log in to their Web site? How do airlines determine what price to quote to you when you are shopping for a plane ticket? How can doctors better diagnose and treat you when you are ill or injured?

Even though you are applying for a loan for the first time, millions of people around the world have applied for loans. Many of these loan recipients have paid back their loans in full and on time, but some of them have not. The bank wants to know whether you are more like those who have paid back their loans or more like those who defaulted. By comparing your credit history, financial situation, and other factors to the vast database of previous loan recipients, the bank can effectively assess how likely you are to default on a loan.

Similarly, Amazon.com has access to millions of previous purchases made by customers on its Web site. Amazon.com examines your previous purchases, the products you have viewed, and any product recommendations you have provided. Amazon.com then searches through its huge database for customers who are similar to you in terms of product purchases, recommendations, and interests. Once similar customers have been identified, their purchases form the basis of the recommendations given to you.

Prices for airline tickets are frequently updated. The price quoted to you for a flight between New York and San Francisco today could be very different from the price quoted tomorrow. These changes happen because airlines use a pricing strategy known as revenue management. Revenue management works by examining vast amounts of data on past airline customer purchases and using these data to forecast future purchases. These forecasts are then fed into sophisticated optimization algorithms that determine the optimal price to charge for a particular flight and when to change that price. Revenue management has resulted in substantial increases in airline revenues.

Finally, consider the case of being evaluated by a doctor for a potentially serious medical issue. Hundreds of medical papers may describe research studies done on patients facing similar diagnoses and thousands of data points exist on their outcomes. However, it is extremely unlikely that your doctor has read every one of these research papers or is aware of all previous patient outcomes. Instead of relying only on her medical training and knowledge gained from her limited set of previous patients, wouldn't it be better for your doctor to have access to the expertise and patient history of thousands of doctors around the world?

In 2007, a group of IBM computer scientists initiated a project to develop a new decision technology to help in answering these types of questions. That technology is called Watson, named after the founder of IBM, Thomas J. Watson. The team at IBM focused on one aim: how the vast amounts of data now available on the Internet can be used to make more data-driven, smarter decisions.

Watson became a household name in 2011, when it famously won the television game show, *Jeopardy!* Since that proof of concept in 2011, IBM has reached agreements with the health insurance provider WellPoint, the financial services company Citibank, and Memorial Sloan-Kettering Cancer Center to apply Watson to the decision problems that they face.

Watson is a system of computing hardware, high-speed data processing, and analytical algorithms that are combined to make data-based recommendations. As more and more data are collected, Watson has the capability to learn over time. In simple terms, according to IBM, Watson gathers hundreds of thousands of possible solutions from a huge data bank, evaluates them using analytical techniques, and proposes only the best solutions for consideration. Watson provides not just a single solution, but a range of good solutions with a confidence level for each.

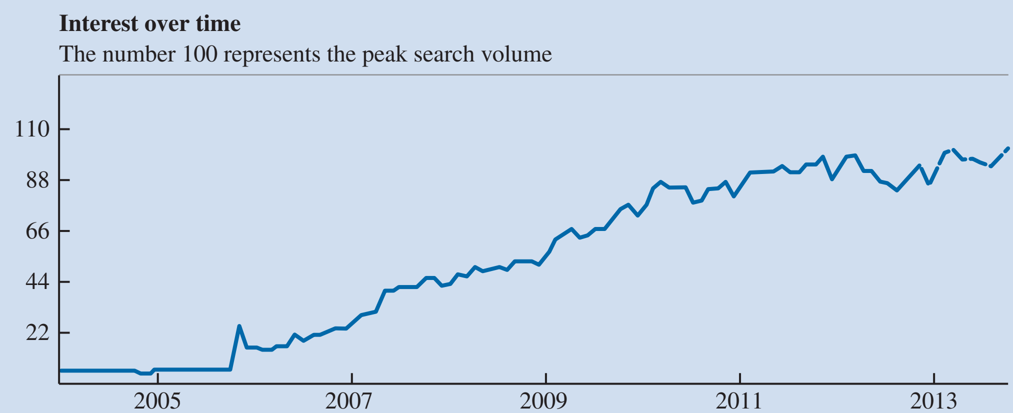
For example, at WellPoint's Virginia data center, to the delight of doctors and patients, Watson is already being used to speed up the approval of medical procedures. Citibank is beginning to explore how to use Watson to better serve its customers, and Sloan-Kettering is launching a pilot study to assess the effectiveness of Watson in assisting with the diagnosis and treatment of patients.[1]

This book is concerned with data-driven decision making and the use of analytical approaches in the decision-making process. Three developments spurred recent explosive growth in the use of analytical methods in business applications. First, technological advances, such as improved point-of-sale scanner technology and the collection of data through e-commerce, Internet social networks, and data generated from personal electronic devices, produce incredible amounts of data for businesses. Naturally, businesses want to use these data to improve the efficiency and profitability of their operations, better understand their customers, price their products more effectively, and gain a competitive advantage. Second, ongoing research has resulted in numerous methodological developments, including advances in computational approaches to effectively handle and explore massive amounts of data, faster algorithms for optimization and simulation, and more effective approaches for visualizing data. Third, these methodological developments were paired with an explosion in computing power and storage capability. Better computing hardware, parallel computing, and, more recently, cloud computing (the remote use of hardware and software over the Internet) have enabled businesses to solve big problems faster and more accurately than ever before.

In summary, the availability of massive amounts of data, improvements in analytic methodologies, and substantial increases in computing power have all come together to result in a dramatic upsurge in the use of analytical methods in business and a reliance on the discipline that is the focus of this text: business analytics. Figure 1.1, a graph generated by Google Trends, displays the search volume for the word *analytics* from 2004 to 2013 (projected) on a percentage basis from the peak. The figure clearly illustrates the recent increase in interest in analytics.

Business analytics is a crucial area of study for students looking to enhance their employment prospects. By 2018, it is predicted that there will be a shortage of more than 1.5 million business managers with adequate training in analytics in the United States

FIGURE 1.1 GOOGLE TRENDS GRAPH OF SEARCHES ON THE TERM *ANALYTICS*



[1]"IBM's Watson Is Learning Its Way to Saving Lives," Fastcompany Web site, December 8, 2012.

alone.[2] As stated in the Preface, the purpose of this text is to provide students with a sound conceptual understanding of the role that business analytics plays in the decision-making process. To reinforce the applications orientation of the text and to provide a better understanding of the variety of applications in which analytical methods have been used successfully, Analytics in Action articles are presented throughout the book. Each Analytics in Action article summarizes an application of analytical methods in practice. For example, the first Analytics in Action, *Procter & Gamble Uses Business Analytics to Redesign Its Supply Chain* (later in this chapter) describes how analytics was used to drive efficiency in Procter & Gamble's North American supply chain.

1.1 Decision Making

It is the responsibility of managers to plan, coordinate, organize, and lead their organizations to better performance. Ultimately, managers' responsibilities require that they make strategic, tactical, or operational decisions. **Strategic decisions** involve higher-level issues concerned with the overall direction of the organization; these decisions define the organization's overall goals and aspirations for the future. Strategic decisions are usually the domain of higher-level executives and have a time horizon of three to five years. **Tactical decisions** concern how the organization should achieve the goals and objectives set by its strategy, and they are usually the responsibility of midlevel management. Tactical decisions usually span a year and thus are revisited annually or even every six months. **Operational decisions** affect how the firm is run from day to day; they are the domain of operations managers, who are the closest to the customer.

Consider the case of the Thoroughbred Running Company (TRC). Historically, TRC had been a catalog-based retail seller of running shoes and apparel. TRC sales revenue grew quickly as it changed its emphasis from catalog-based sales to Internet-based sales. Recently, TRC decided that it should also establish retail stores in the malls and downtown areas of major cities. This is a strategic decision that will take the firm in a new direction that it hopes will complement its Internet-based strategy. TRC middle managers will therefore have to make a variety of tactical decisions in support of this strategic decision, including how many new stores to open this year, where to open these new stores, how many distribution centers will be needed to support the new stores, and where to locate these distribution centers. Operations managers in the stores will need to make day-to-day decisions regarding, for instance, how many pairs of each model and size of shoes to order from the distribution centers and how to schedule their sales personnel.

Regardless of the level within the firm, *decision making* can be defined as the following process:

1. Identify and define the problem
2. Determine the criteria that will be used to evaluate alternative solutions
3. Determine the set of alternative solutions
4. Evaluate the alternatives
5. Choose an alternative

If I were given one hour to save the planet, I would spend 59 minutes defining the problem and one minute resolving it.
—Albert Einstein

Step 1 of decision making, identifying and defining the problem, is the most critical. Only if the problem is well-defined, with clear metrics of success or failure (step 2), can a proper approach for solving the problem (steps 3 and 4) be devised. Decision making concludes with the choice of an alternative (step 5).

[2] J. Manyika et al., "Big Data: The Next Frontier for Innovation, Competition and Productivity," McKinsey Global Institute Report, 2011.

There are a number of approaches to making decisions: tradition ("We've always done it this way"), intuition ("gut feeling"), and rules of thumb ("As the restaurant owner, I schedule twice the number of waiters and cooks on holidays"). The power of each of these approaches should not be underestimated. Managerial experience and intuition are valuable inputs to making decisions, but what if relevant data were available to help us make more informed decisions? With the vast amounts of data now generated and stored electronically, it is estimated that the amount of data stored by businesses more than doubles every two years. How can managers convert these data into knowledge that they can use to be more efficient and effective in managing their businesses?

1.2 Business Analytics Defined

What makes decision making difficult and challenging? Uncertainty is probably the number one challenge. If we knew how much the demand will be for our product, we could do a much better job of planning and scheduling production. If we knew exactly how long each step in a project will take to be completed, we could better predict the project's cost and completion date. If we knew how stocks will perform, investing would be a lot easier.

Another factor that makes decision making difficult is that we often face such an enormous number of alternatives that we cannot evaluate them all. What is the best combination of stocks to help me meet my financial objectives? What is the best product line for a company that wants to maximize its market share? How should an airline price its tickets so as to maximize revenue?

Some firms and industries use the simpler term, analytics. *Analytics is often thought of as a broader category than business analytics, encompassing the use of analytical techniques in the sciences and engineering as well. In this text, we use* business analytics *and* analytics *synonymously.*

Business analytics is the scientific process of transforming data into insight for making better decisions.[3] Business analytics is used for data-driven or fact-based decision making, which is often seen as more objective than other alternatives for decision making.

As we shall see, the tools of business analytics can aid decision making by creating insights from data, by improving our ability to more accurately forecast for planning, by helping us quantify risk, and by yielding better alternatives through analysis and optimization. Indeed, a recent study based on a large sample of firms that was conducted by researchers at MIT's Sloan School of Management and the University of Pennsylvania, concluded that firms guided by data-driven decision making have higher productivity and market value and increased output and profitability.[4]

1.3 A Categorization of Analytical Methods and Models

Business analytics can involve anything from simple reports to the most advanced optimization techniques (methods for finding the best course of action). Analytics is generally thought to comprise three broad categories of techniques: descriptive analytics, predictive analytics, and prescriptive analytics.

Descriptive Analytics

Descriptive analytics encompasses the set of techniques that describes what has happened in the past. Examples are data queries, reports, descriptive statistics, data visualization

[3]We adopt the definition of analytics developed by the Institute for Operations Research and the Management Sciences (INFORMS).

[4]E. Brynjolfsson, L. M. Hitt, and H. H. Kim, "Strength in Numbers: How Does Data-Driven Decisionmaking Affect Firm Performance?" (April 18, 2013). Available at SSRN http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1819486.

including data dashboards, some data-mining techniques, and basic what-if spreadsheet models.

Appendix B at the end of this book describes how to use Microsoft Access to conduct data queries.

A **data query** is a request for information with certain characteristics from a database. For example, a query to a manufacturing plant's database might be for all records of shipments to a particular distribution center during the month of March. This query provides descriptive information about these shipments: the number of shipments, how much was included in each shipment, the date each shipment was sent, and so on. A report summarizing relevant historical information for management might be conveyed by the use of descriptive statistics (means, measures of variation, etc.) and data visualization tools (tables, charts, and maps). Simple descriptive statistics and data visualization techniques can be used to find patterns or relationships in a large database.

Data dashboards are collections of tables, charts, maps, and summary statistics that are updated as new data become available. Dashboards are used to help management monitor specific aspects of the company's performance related to their decision-making responsibilities. For corporate-level managers, daily data dashboards might summarize sales by region, current inventory levels, and other company-wide metrics; front-line managers may view dashboards that contain metrics related to staffing levels, local inventory levels, and short-term sales forecasts.

Predictive Analytics

Predictive analytics consists of techniques that use models constructed from past data to predict the future or ascertain the impact of one variable on another. For example, past data on product sales may be used to construct a mathematical model to predict future sales, which can factor in the product's growth trajectory and seasonality based on past patterns. A packaged food manufacturer may use point-of-sale scanner data from retail outlets to help in estimating the lift in unit sales due to coupons or sales events. Survey data and past purchase behavior may be used to help predict the market share of a new product. All of these are applications of predictive analytics.

Linear regression, time series analysis, some data-mining techniques, and simulation, often referred to as risk analysis, all fall under the banner of predictive analytics. We discuss all of these techniques in greater detail later in this text.

Data mining, techniques used to find patterns or relationships among elements of the data in a large database, is often used in predictive analytics. For example, a large grocery store chain might be interested in developing a new targeted marketing campaign that offers a discount coupon on potato chips. By studying historical point-of-sale data, the store may be able to use data mining to predict which customers are the most likely to respond to an offer on discounted chips by purchasing higher-margin items such as beer or soft drinks in addition to the chips, thus increasing the store's overall revenue.

Simulation involves the use of probability and statistics to construct a computer model to study the impact of uncertainty on a decision. For example, banks often use simulation to model investment and default risk in order to stress test financial models. Simulation is also often used in the pharmaceutical industry to assess the risk of introducing a new drug.

Prescriptive Analytics

Prescriptive analytics differ from descriptive or predictive analytics in that **prescriptive analytics** indicate a best course of action to take; that is, the output of a prescriptive model is a best decision. The airline industry's use of revenue management is an example of a prescriptive analytics. Airlines use past purchasing data as inputs into a model that recommends the best pricing strategy across all flights for maximizing revenue.

Other examples of prescriptive analytics are portfolio models in finance, supply network design models in operations, and price markdown models in retailing. Portfolio models use historical investment return data to determine the mix of investments that yield the highest expected return while controlling or limiting exposure to risk. Supply network design models provide the cost-minimizing plant and distribution center locations subject to meeting the customer service requirements. Given historical data, retail price markdown models yield revenue-maximizing discount levels and the timing of discount offers when goods have not sold as planned. All of these models are known as **optimization models**, that is, models that give the best decision subject to constraints of the situation.

Another type of modeling in the prescriptive analytics category is **simulation optimization**, which combines the use of probability and statistics to model uncertainty with optimization techniques to find good decisions in highly complex and highly uncertain settings. Finally, the techniques of **decision analysis** can be used to develop an optimal strategy when a decision maker is faced with several decision alternatives and an uncertain set of future events. Decision analysis also employs **utility theory**, which assigns values to outcomes based on the decision maker's attitude toward risk, loss, and other factors.

In this text we cover all three areas of business analytics: descriptive, predictive, and prescriptive. Table 1.1 shows how the chapters cover the three categories.

TABLE 1.1 COVERAGE OF BUSINESS ANALYTICS TOPICS IN THIS TEXT

Chapter	Title	Descriptive	Predictive	Prescriptive
1	Introduction	●	●	●
2	Descriptive Statistics	●		
3	Data Visualization	●		
4	Linear Regression	●	●	
5	Time Series Analysis & Forecasting		●	
6	Data Mining	●	●	
7	Spreadsheet Models	●		
8	Linear Optimization Models			●
9	Integer Linear Optimization Models			●
10	Nonlinear Optimization Models			●
11	Monte Carlo Simulation		●	●
12	Decision Analysis			●

ANALYTICS *in* ACTION

PROCTER & GAMBLE USES BUSINESS ANALYTICS TO REDESIGN ITS SUPPLY CHAIN[5]

Consumer goods giant Procter & Gamble (P&G), the maker of such well-known brands as Tide, Olay, Crest, Bounty, and Pampers, sells its products in over 180 countries around the world. Supply chain coordination and efficiency are critical to the company's profitability. After many years of acquisitions and growth, P&G embarked on a effort known as Strengthening Global Effectiveness. A major piece of that effort was the North American Supply Chain Study, whose purpose was to make the supply chain in North America as efficient as possible, while ensuring that customer service requirements were met.

A team of P&G analysts and managers partnered with a group of analytics faculty at the University of Cincinnati to create a system to help managers redesign

[5]J. Camm, T. Chorman, F. Dill, J. Evans, D. Sweeney, and G. Wegryn, "Blending OR/MS, Judgment and GIS: Restructuring P&G's Supply Chain," *Interfaces* 27, no. 1 (1997): 83–97.

the supply effort in North America. The fundamental questions to be answered were: (1) Which plants should make which product families? (2) Where should the distribution centers be located? (3) Which plants should serve which distribution centers? (4) Which customers should be served by each distribution center? The team's approach utilized all three categories of business analytics: descriptive, predictive, and prescriptive.

At the start of the study, data had to be collected from all aspects of the supply chain. These included demand by product family, fixed and variable production costs by plant, and freight costs and handling charges at the distribution centers. Data queries and descriptive statistics were utilized to acquire and better understand the current supply chain data. Data visualization, in the form of a geographic information system, allowed the proposed solutions to be displayed on a map for more intuitive interpretation by management. Because the supply chain had to be redesigned for the future, predictive analytics was used to forecast product family demand by three-digit zip code for ten years into the future. This future demand was then input, along with projected freight and other relevant costs, into an interactive optimization model, that minimized cost subject to service constraints. The suite of analytical models was aggregated into a single system that could be run quickly on a laptop computer. P&G product category managers made over a thousand runs of the system before reaching consensus on a small set of alternative designs. Each proposed design in this selected set was then subjected to a risk analysis using computer simulation, ultimately leading to a single go-forward design.

The chosen redesign of the supply chain was implemented over time and led to a documented savings in excess of $250 million per year in P&G's North American supply chain. The system of models was later utilized to streamline the supply chains in Europe and Asia, and P&G has become a world leader in the use of analytics in supply chain management.

1.4 Big Data

Like the explosion of interest in analytics, interest in what is known as big data has recently increased dramatically. **Big data** is simply a set of data that cannot be managed, processed, or analyzed with commonly available software in a reasonable amount of time. Walmart handles over one million purchase transactions per hour. Facebook processes more than 250 million picture uploads per day. Five billion cell-phone owners around the world generate vast amounts of data by calling, texting, tweeting and browsing the web on a daily basis.[6] As Google CEO Eric Schmidt has noted,[7] the amount of data currently created every 48 hours is equivalent to the entire amount of data created from the dawn of civilization until the year 2003. Perhaps it is not surprising that 90 percent of the data in the world today has been created in the last two years.[8]

Businesses are interested in understanding and using data to gain a competitive advantage. Although big data represents opportunities, it also presents analytical challenges from a processing point of view and consequently has itself led to an increase in the use of analytics. More companies are hiring data scientists who know how to process and analyze massive amounts of data. However, it is important to understand that in some sense big data issues are a subset of analytics and that many very valuable applications of analytics do not involve big data.

[6]SAS White Paper, "Big Data Meets Big Data Analytics," SAS Institute, 2012.

[7]E. Schmidt, Panel discussion at Technomy Conference, Lake Tahoe, CA, August 4, 2010.

[8]"Bringing Big Data to the Enterprise," IBM Website. Available at http://www-01.ibm.com/software/data/bigdata/, retrieved December 1, 2012.

1.5 Business Analytics in Practice

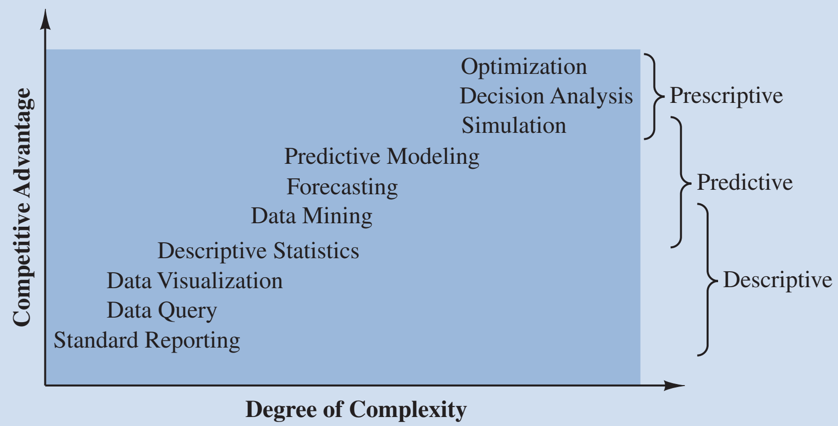
Business analytics involves tools as simple as reports and graphs, as well as some that are as sophisticated as optimization, data mining, and simulation. In practice, companies that apply analytics often follow a trajectory similar to that shown in Figure 1.2. Organizations start with basic analytics in the lower left. As they realize the advantages of these analytic techniques, they often progress to more sophisticated techniques in an effort to reap the derived competitive advantage. Predictive and prescriptive analytics are sometimes therefore referred to as **advanced analytics**. Not all companies reach that level of usage, but those that embrace analytics as a competitive strategy often do.

Analytics has been applied in virtually all sectors of business and government. Organizations such as Procter & Gamble, IBM, UPS, Netflix, Amazon.com, Google, the Internal Revenue Service, and General Electric have embraced analytics to solve important problems or to achieve competitive advantage. In this section, we briefly discuss some of the types of applications of analytics by application area.

Financial Analytics

Applications of analytics in finance are numerous and pervasive. Predictive models are used to forecast future financial performance, to assess the risk of investment portfolios and projects, and to construct financial instruments such as derivatives. Prescriptive models are used to construct optimal portfolios of investments, to allocate assets, and to create optimal capital budgeting plans. For example, GE Asset Management uses optimization models to decide how to invest its own cash received from insurance policies and other financial products, as well as the cash of its clients such as Genworth Financial. The estimated benefit from the optimization models was $75 million over a five-year period.[9] Simulation is also often used to assess risk in the financial sector; one example is the deployment by Hypo Real Estate International of simulation models to successfully manage commercial real estate risk.[10]

FIGURE 1.2 THE SPECTRUM OF BUSINESS ANALYTICS



Source: Adapted from SAS.

[9]L. C. Chalermkraivuth et al., "GE Asset Management, Genworth Financial, and GE Insurance Use a Sequential-Linear Programming Algorithm to Optimize Portfolios," *Interfaces* 35, no. 5 (September–October 2005): 370–80.
[10]Y. Jafry, C. Marrison, and U. Umkehrer-Neudeck, "Hypo International Strengthens Risk Management with a Large-Scale, Secure Spreadsheet-Management Framework," *Interfaces* 38, no. 4 (July–August 2008): 281–88.