

BUSINESS ANALYTICS

PRINCIPLES, CONCEPTS,
AND APPLICATIONS



WHAT, WHY, and HOW

MARC J. SCHNIEDERJANS • DARA G. SCHNIEDERJANS • CHRISTOPHER M. STARKEY

Business Analytics
Principles, Concepts, and
Applications

This page intentionally left blank

Business Analytics

Principles, Concepts, and Applications

What, Why, and How

Marc J. Schniederjans

Dara G. Schniederjans

Christopher M. Starkey

Pearson

Associate Publisher: Amy Neidlinger
Executive Editor: Jeanne Glasser Levine
Operations Specialist: Jodi Kemper
Cover Designer: Alan Clements
Cover Image: Alan McHugh
Managing Editor: Kristy Hart
Senior Project Editor: Lori Lyons
Copy Editor: Gill Editorial Services
Proofreader: Katie Matejka
Indexer: Erika Millen
Senior Compositor: Gloria Schurick
Manufacturing Buyer: Dan Uhrig

© 2014 by Marc J. Schniederjans, Dara G. Schniederjans, and Christopher M. Starkey
Pearson Education, Inc.
Upper Saddle River, New Jersey 07458

For information about buying this title in bulk quantities, or for special sales opportunities (which may include electronic versions; custom cover designs; and content particular to your business, training goals, marketing focus, or branding interests), please contact our corporate sales department at corpsales@pearsoned.com or (800) 382-3419.

For government sales inquiries, please contact governmentsales@pearsoned.com.

For questions about sales outside the U.S., please contact international@pearsoned.com.

Company and product names mentioned herein are the trademarks or registered trademarks of their respective owners.

All rights reserved. No part of this book may be reproduced, in any form or by any means, without permission in writing from the publisher.

Printed in the United States of America

First Printing: April 2014

ISBN-10: 0-13-355218-7

ISBN-13: 978-0-13-355218-8

Pearson Education LTD.

Pearson Education Australia PTY, Limited.

Pearson Education Singapore, Pte. Ltd.

Pearson Education Asia, Ltd.

Pearson Education Canada, Ltd.

Pearson Educación de Mexico, S.A. de C.V.

Pearson Education—Japan

Pearson Education Malaysia, Pte. Ltd.

Library of Congress Control Number: 2014931049

*This book is dedicated to Miles Starkey.
He is what brings purpose to our lives
and gives us a future.*

This page intentionally left blank

Contents-at-a-Glance

| | | |
|-------------------|--|------------|
| | Preface | xvi |
| PART I: | What Are Business Analytics | 1 |
| Chapter 1: | What Are Business Analytics? | 3 |
| PART II: | Why Are Business Analytics Important. | 15 |
| Chapter 2: | Why Are Business Analytics Important? | 17 |
| Chapter 3: | What Resource Considerations Are Important to Support Business Analytics? | 29 |
| PART III: | How Can Business Analytics Be Applied | 43 |
| Chapter 4: | How Do We Align Resources to Support Business Analytics within an Organization? | 45 |
| Chapter 5: | What Are Descriptive Analytics? | 63 |
| Chapter 6: | What Are Predictive Analytics? | 93 |
| Chapter 7: | What Are Prescriptive Analytics? | 119 |
| Chapter 8: | A Final Case Study Illustration | 139 |
| PART IV: | Appendixes | 165 |
| A: | Statistical Tools | 167 |
| B: | Linear Programming | 201 |
| C: | Duality and Sensitivity Analysis in Linear Programming | 241 |

D: Integer Programming263

E: Forecasting.....271

F: Simulation295

G: Decision Theory.....303

Index.....335

Table of Contents

| | | |
|-------------------|--|-----------|
| | Preface | xvi |
| PART I: | What Are Business Analytics | 1 |
| Chapter 1: | What Are Business Analytics?..... | 3 |
| | 1.1 Terminology | 3 |
| | 1.2 Business Analytics Process..... | 7 |
| | 1.3 Relationship of BA Process and Organization Decision-Making Process | 10 |
| | 1.4 Organization of This Book | 12 |
| | Summary..... | 13 |
| | Discussion Questions..... | 13 |
| | References | 14 |
| PART II: | Why Are Business Analytics Important | 15 |
| Chapter 2: | Why Are Business Analytics Important? | 17 |
| | 2.1 Introduction | 17 |
| | 2.2 Why BA Is Important: Providing Answers to Questions | 18 |
| | 2.3 Why BA Is Important: Strategy for Competitive Advantage | 20 |
| | 2.4 Other Reasons Why BA Is Important | 23 |
| | 2.4.1 Applied Reasons Why BA Is Important | 23 |
| | 2.4.2 The Importance of BA with New Sources of Data | 24 |
| | Summary..... | 26 |
| | Discussion Questions..... | 26 |
| | References | 26 |
| Chapter 3: | What Resource Considerations Are Important to Support Business Analytics? | 29 |
| | 3.1 Introduction | 29 |
| | 3.2 Business Analytics Personnel..... | 30 |
| | 3.3 Business Analytics Data..... | 33 |
| | 3.3.1 Categorizing Data | 33 |
| | 3.3.2 Data Issues..... | 35 |
| | 3.4 Business Analytics Technology | 36 |
| | Summary..... | 41 |
| | Discussion Questions..... | 41 |
| | References | 42 |

- PART III: How Can Business Analytics Be Applied 43**
- Chapter 4: How Do We Align Resources to Support Business Analytics within an Organization? 45**
 - 4.1 Organization Structures Aligning Business Analytics 45
 - 4.1.1 Organization Structures 46
 - 4.1.2 Teams 51
 - 4.2 Management Issues 54
 - 4.2.1 Establishing an Information Policy 54
 - 4.2.2 Outsourcing Business Analytics 55
 - 4.2.3 Ensuring Data Quality 56
 - 4.2.4 Measuring Business Analytics Contribution 58
 - 4.2.5 Managing Change 58
 - Summary 60
 - Discussion Questions 61
 - References 61
- Chapter 5: What Are Descriptive Analytics? 63**
 - 5.1 Introduction 63
 - 5.2 Visualizing and Exploring Data 64
 - 5.3 Descriptive Statistics 67
 - 5.4 Sampling and Estimation 72
 - 5.4.1 Sampling Methods 73
 - 5.4.2 Sampling Estimation 76
 - 5.5 Introduction to Probability Distributions 78
 - 5.6 Marketing/Planning Case Study Example: Descriptive Analytics Step in the BA Process 80
 - 5.6.1 Case Study Background 81
 - 5.6.2 Descriptive Analytics Analysis 82
 - Summary 91
 - Discussion Questions 91
 - Problems 92
- Chapter 6: What Are Predictive Analytics? 93**
 - 6.1 Introduction 93
 - 6.2 Predictive Modeling 94
 - 6.2.1 Logic-Driven Models 94
 - 6.2.2 Data-Driven Models 96
 - 6.3 Data Mining 97

| | |
|---|------------|
| 6.3.1 A Simple Illustration of Data Mining | 98 |
| 6.3.2 Data Mining Methodologies | 99 |
| 6.4 Continuation of Marketing/Planning Case Study Example: Prescriptive Analytics Step in the BA Process | 102 |
| 6.4.1 Case Study Background Review | 103 |
| 6.4.2 Predictive Analytics Analysis | 104 |
| Summary | 114 |
| Discussion Questions | 115 |
| Problems | 115 |
| References | 117 |
| Chapter 7: What Are Prescriptive Analytics? | 119 |
| 7.1 Introduction | 119 |
| 7.2 Prescriptive Modeling | 120 |
| 7.3 Nonlinear Optimization | 122 |
| 7.4 Continuation of Marketing/Planning Case Study Example: Prescriptive Step in the BA Analysis | 129 |
| 7.4.1 Case Background Review | 129 |
| 7.4.2 Prescriptive Analysis | 129 |
| Summary | 134 |
| Addendum | 134 |
| Discussion Questions | 135 |
| Problems | 135 |
| References | 137 |
| Chapter 8: A Final Business Analytics Case Problem | 139 |
| 8.1 Introduction | 139 |
| 8.2 Case Study: Problem Background and Data | 140 |
| 8.3 Descriptive Analytics Analysis | 141 |
| 8.4 Predictive Analytics Analysis | 147 |
| 8.4.1 Developing the Forecasting Models | 147 |
| 8.4.2 Validating the Forecasting Models | 155 |
| 8.4.3 Resulting Warehouse Customer Demand Forecasts | 157 |
| 8.5 Prescriptive Analytics Analysis | 158 |
| 8.5.1 Selecting and Developing an Optimization Shipping Model | 158 |
| 8.5.2 Determining the Optimal Shipping Schedule | 159 |
| 8.5.3 Summary of BA Procedure for the Manufacturer | 161 |
| 8.5.4 Demonstrating Business Performance Improvement | 162 |

Summary163
Discussion Questions164
Problems164

PART IV: Appendixes 165

A: Statistical Tools167

A.1 Introduction167
A.2 Counting.167
A.3 Probability Concepts171
A.4 Probability Distributions177
A.5 Statistical Testing.193

B: Linear Programming201

B.1 Introduction201
B.2 Types of Linear Programming Problems/Models201
B.3 Linear Programming Problem/Model Elements202
B.4 Linear Programming Problem/Model Formulation
Procedure207
B.5 Computer-Based Solutions for Linear Programming
Using the Simplex Method217
B.6 Linear Programming Complications.227
B.7 Necessary Assumptions for Linear Programming Models.232
B.8 Linear Programming Practice Problems233

**C: Duality and Sensitivity Analysis in Linear
Programming.241**

C.1 Introduction241
C.2 What Is Duality?241
C.3 Duality and Sensitivity Analysis Problems243
C.4 Determining the Economic Value of a
Resource with Duality.258
C.5 Duality Practice Problems.259

D: Integer Programming263

D.1 Introduction.263
D.2 Solving IP Problems/Models264
D.3 Solving Zero-One Programming Problems/Models268
D.4 Integer Programming Practice Problems.270

| | | |
|-----------|--|------------|
| E: | Forecasting | 271 |
| | E.1 Introduction | 271 |
| | E.2 Types of Variation in Time Series Data | 272 |
| | E.3 Simple Regression Model | 276 |
| | E.4 Multiple Regression Models | 281 |
| | E.5 Simple Exponential Smoothing | 284 |
| | E.6 Smoothing Averages | 286 |
| | E.7 Fitting Models to Data | 288 |
| | E.8 How to Select Models and Parameters for Models | 291 |
| | E.9 Forecasting Practice Problems | 292 |
| F: | Simulation | 295 |
| | F.1 Introduction | 295 |
| | F.2 Types of Simulation | 295 |
| | F.3 Simulation Practice Problems | 302 |
| G: | Decision Theory | 303 |
| | G.1 Introduction | 303 |
| | G.2 Decision Theory Model Elements | 304 |
| | G.3 Types of Decision Environments | 304 |
| | G.4 Decision Theory Formulation | 305 |
| | G.5 Decision-Making Under Certainty | 306 |
| | G.6 Decision-Making Under Risk | 307 |
| | G.7 Decision-Making under Uncertainty | 311 |
| | G.8 Expected Value of Perfect Information | 315 |
| | G.9 Sequential Decisions and Decision Trees | 317 |
| | G.10 The Value of Imperfect Information: Bayes's Theorem | 321 |
| | G.11 Decision Theory Practice Problems | 328 |
| | Index | 335 |

About the Authors



Marc J. Schniederjans is the C. Wheaton Battey Distinguished Professor of Business in the College of Business Administration at the University of Nebraska-Lincoln and has served on the faculty of three other universities. Professor Schniederjans is a Fellow of the Decision Sciences Institute (DSI) and in 2014–2015 will serve as DSI's President. His prior experience includes owning and operating his own truck leasing business. He is currently a member of the Institute of Supply Management (ISM), the Production and Operations Management Society (POMS), and Decision Sciences Institute (DSI). Professor Schniederjans has taught extensively in operations management and management science. He has won numerous teaching awards and is an honorary member of the Golden Key honor society and the Alpha Kappa Psi business honor society. He has published more than a hundred journal articles and has authored or coauthored twenty books in the field of management. The title of his most recent book is *Reinventing the Supply Chain Life Cycle*, and his research has encompassed a wide range of operations management and decision science topics. He has also presented more than one hundred research papers at academic meetings. Professor Schniederjans is serving on five journal editorial review boards, including *Computers & Operations Research*, *International Journal of Information & Decision Sciences*, *International Journal of Information Systems in the Service Sector*, *Journal of Operations Management*, *Production and Operations Management*. He is also serving as an area editor for the journal *Operations Management Research* and as an associate editor for the *International Journal of Strategic Decision Sciences* and *International Journal of the Society Systems Science and Management Review: An International Journal* (Korea). Professor Schniederjans has also served as a consultant and trainer to various business and government agencies.



Dara G. Schniederjans is an assistant professor of Supply Chain Management at the University of Rhode Island, College of Business Administration. She has published articles in journals such as *Decision Support Systems*, *Journal of the Operational Research Society*, and *Business Process Management Journal*. She has also coauthored two text books and coedited a readings book. She has contributed chapters to readings utilizing quantitative and statistical methods. Dara has served as a guest coeditor for a special issue on *Business Ethics in Social Sciences* in the *International Journal of Society Systems Science*. She has also served as a website coordinator for Decisions Sciences Institute. She currently teaches courses in Supplier Relationship Management and Operations Management.



Christopher M. Starkey is an economics student at the University of Connecticut-Storrs. He has presented papers at the Academy of Management and Production and Operations Management Society meetings. He currently teaches courses in Principles of Microeconomics and has taught Principles of Macroeconomics. His current research interests include macroeconomic and monetary policy, as well as other decision-making methodologies.

Preface

Like the face on the cover of this book, we are bombarded by information every day. We do our best to sort out and use the information to help us get by, but sometimes we are overwhelmed by the abundance of data. This can lead us to draw wrong conclusions and make bad decisions. When you are a global firm collecting millions of transactions and customer behavior data from all over the world, the size of the data alone can make the task of finding useful information about customers almost impossible. For that firm and even smaller businesses, the solution is to apply *business analytics* (BA). BA helps sort out large data files (called “big data”), find patterns of behavior useful in predicting the future, and allocate resources to optimize decision-making. BA involves a step-wise process that aids firms in managing big data in a systematic procedure to glean useful information, which can solve problems and pinpoint opportunities for enhanced business performance.

This book has been written to provide a basic education in BA that can serve both academic and practitioner markets. In addition to bringing BA up-to-date with literature and research, this book explains the BA process in simple terms and supporting methodologies useful in its application. Collectively, the statistical and quantitative tools presented in this book do not need substantial prerequisites other than basic high school algebra. To support both markets, a substantial number of solved problems are presented along with some case study applications to train readers in the use of common BA tools and software. Practitioners will find the treatment of BA methodologies useful review topics. Academic users will find chapter objectives and discussion questions helpful for serving their needs while also having an opportunity to obtain an Instructor’s Guide with chapter-end problem solutions and exam questions.

The purpose of this book is to explain what BA is, why it is important to know, and how to do it. To achieve this purpose, the book presents conceptual content, software familiarity, and some analytic tools.

Conceptual Content

The conceptual material is presented in the first eight chapters of the book. (See Section 1.4 in Chapter 1 for an explanation of the book’s organization.) The conceptual content covers much more than what BA is about. The book explains why BA is important in terms of proving answers to questions, how it can be used to achieve

competitive advantage, and how to align an organization to make best use of it. The book explains the managerial aspects of creating a BA presence in an organization and the skills BA personnel are expected to possess. The book also describes data management issues such as data collection, outsourcing, data quality, and change management as they relate to BA.

Having created a managerial foundation explaining “what” and “why” BA is important, the remaining chapters focus on “how” to do it. Embodied in a three-step process, BA is explained to have descriptive, predictive, and prescriptive analytic steps. For each of these steps, this book presents a series of strategies and best practice guides to aid in the BA process.

Software

Much of what BA is about involves the use of software. Unfortunately, no single software covers all aspects of BA. Many institutions prefer one type of software over others. To provide flexibility, this book’s use of software provides some options and can be used by readers who are not even interested in running computer software. In this book, SPSS®, Excel®, and Lingo® software are utilized to model and solve problems. The software treatment is mainly the output of these software systems, although some input and instructions on their use is provided. For those not interested in running software applications, the exposure to the printouts provides insight into their informational value. This book recognizes that academic curriculums prefer to uniquely train students in the use of software and does not duplicate basic software usage. As a prerequisite to using this book, it is recommended that those interested in running software applications for BA become familiar with and are instructed on the use of whatever software is desired.

Analytic Tools

The analytic tool materials are chiefly contained in this book’s appendixes. BA is a statistical, management information systems (MIS) and quantitative methods tools-oriented subject. While the conceptual content in the book overviews how to undertake the BA process, the implementation of how to actually do BA requires quantitative tools. Because some practitioners and academic programs are less interested in the technical aspects of BA, the bulk of the quantitative material is presented

in the appendixes. These appendixes provide an explanation and illustration of a substantial body of BA tools to support a variety of analyses. Some of the statistical tools that are explained and illustrated in this book include statistical counting (permutations, combinations, repetitions), probability concepts (approaches to probability, rules of addition, rules of multiplication, Bayes' Theorem), probability distributions (binomial, Poisson, normal, exponential), confidence intervals, sampling methods, simple and multiple regression, charting, and hypothesis testing. Although management information systems are beyond the scope of this book, the software applications previously mentioned are utilized to illustrate search, clustering, and typical data mining applications of MIS technology. In addition, quantitative methods tools explained and illustrated in this book include linear programming, duality and sensitivity analysis, integer programming, zero-one programming, forecasting modeling, nonlinear optimization, simulation analysis, breakeven analysis, and decision theory (certainty, risk, uncertainty, expected value opportunity loss analysis, expected value of perfect information, expected value of imperfect information).

We want to acknowledge the help of individuals who provided needed support for the creation of this book. First, we really appreciate the support of our editor, Jeanne Glasser Levine, and the outstanding staff at Financial Times Press/Pearson. They made creating this book a pleasure and worked with us to improve the final product. Decades of writing books with other publishers permitted us to recognize how a top-tier publisher like ours makes a difference. We thank Alan McHugh, who developed the image on our book cover. His constant willingness to explore and be innovative with ideas made a significant contribution to our book. We also want to acknowledge the great editing help we received from Jill Schniederjans. Her skill has reduced the wordiness and enhanced the content (making parts less boring to read). Finally, we would like to acknowledge the help of Miles Starkey, whose presence and charm have lifted our spirits and kept us on track to meet completion deadlines.

While many people have assisted in preparing this book, its accuracy and completeness are our responsibility. For all errors that this book may contain, we apologize in advance.

Marc J. Schniederjans

Dara G. Schniederjans

Christopher M. Starkey

This page intentionally left blank

1

What Are Business Analytics?

Chapter objectives:

- Define business analytics.
- Explain the relationship of analytics and business intelligence to the subject of business analytics.
- Describe the three steps of the business analytics process.
- Describe four data classification measurement scales.
- Explain the relationship of the business analytics process with the organization decision-making process.

1.1 Terminology

Business analytics begins with a *data set* (a simple collection of data or a data file) or commonly with a *database* (a collection of data files that contain information on people, locations, and so on). As databases grow, they need to be stored somewhere. Technologies such as *computer clouds* (hardware and software used for data remote storage, retrieval, and computational functions) and *data warehousing* (a collection of databases used for reporting and data analysis) store data. Database storage areas have become so large that a new term was devised to describe them. *Big data* describes the collection of data sets that are so large and complex that software systems are hardly able to process them (Isson and Harriott, 2013, pp. 57–61). Isson and Harriott (2013, p. 61) define *little data* as anything that is not big data. Little data describes the smaller data segments or files that help individual businesses keep track of customers. As a means of sorting through data to find useful information, the application of analytics has found new purpose.

Three terms in business literature are often related to one another: analytics, business analytics, and business intelligence. *Analytics* can be defined as a process that involves the use of statistical techniques (measures of central tendency, graphs, and so on), information system software (data mining, sorting routines), and operations research methodologies (linear programming) to explore, visualize, discover and communicate patterns or trends in data. Simply, analytics convert data into useful information. Analytics is an older term commonly applied to all disciplines, not just business. A typical example of the use of analytics is the weather measurements collected and converted into statistics, which in turn predict weather patterns.

There are many types of analytics, and there is a need to organize these types to understand their uses. We will adopt the three categories (*descriptive*, *predictive*, and *prescriptive*) that the *Institute of Operations Research and Management Sciences* (INFORMS) organization (www.informs.org) suggests for grouping the types of analytics (see Table 1.1). These types of analytics can be viewed independently. For example, some firms may only use descriptive analytics to provide information on decisions they face. Others may use a combination of analytic types to glean insightful information needed to plan and make decisions.

Table 1.1 Types of Analytics

| Type of Analytics | Definition |
|-------------------|---|
| Descriptive | The application of simple statistical techniques that describes what is contained in a data set or database. Example: An age bar chart is used to depict retail shoppers for a department store that wants to target advertising to customers by age. |
| Predictive | An application of advanced statistical, information software, or operations research methods to identify predictive variables and build predictive models to identify trends and relationships not readily observed in a descriptive analysis. Example: Multiple regression is used to show the relationship (or lack of relationship) between age, weight, and exercise on diet food sales. Knowing that relationships exist helps explain why one set of independent variables influences dependent variables such as business performance. |
| Prescriptive | An application of decision science, management science, and operations research methodologies (applied mathematical techniques) to make best use of allocable resources. Example: A department store has a limited advertising budget to target customers. Linear programming models can be used to optimally allocate the budget to various advertising media. |

The purposes and methodologies used for each of the three types of analytics differ, as can be seen in Table 1.2. It is these differences that distinguish *analytics* from *business analytics*. Whereas analytics is focused on generating insightful information

from data sources, business analytics goes the extra step to leverage analytics to create an improvement in measurable business performance. Whereas the process of analytics can involve any one of the three types of analytics, the major components of business analytics include all three used in combination to generate new, unique, and valuable information that can aid business organization decision-making. In addition, the three types of analytics are applied sequentially (descriptive, then predictive, then prescriptive). Therefore, *business analytics* (BA) can be defined as a process beginning with business-related data collection and consisting of sequential application of descriptive, predictive, and prescriptive major analytic components, the outcome of which supports and demonstrates business decision-making and organizational performance. Stubbs (2011, p. 11) believes that BA goes beyond plain analytics, requiring a clear relevancy to business, a resulting insight that will be implementable, and performance and value measurement to ensure a successful business result.

Table 1.2 Analytic Purposes and Tools

| Type of Analytics | Purpose | Examples of Methodologies |
|-------------------|---|--|
| Descriptive | To identify possible trends in large data sets or databases. The purpose is to get a rough picture of what generally the data looks like and what criteria might have potential for identifying trends or future business behavior. | Descriptive statistics, including measures of central tendency (mean, median, mode), measures of dispersion (standard deviation), charts, graphs, sorting methods, frequency distributions, probability distributions, and sampling methods. |
| Predictive | To build predictive models designed to identify and predict future trends. | Statistical methods like multiple regression and ANOVA. Information system methods like data mining and sorting. Operations research methods like forecasting models. |
| Prescriptive | To allocate resources optimally to take advantage of predicted trends or future opportunities. | Operations research methodologies like linear programming and decision theory. |

Business intelligence (BI) can be defined as a set of processes and technologies that convert data into meaningful and useful information for business purposes. While some believe that BI is a broad subject that encompasses analytics, business analytics, and information systems (Bartlett, 2013, p.4), others believe it is mainly focused on collecting, storing, and exploring large database organizations for information useful to decision-making and planning (Negash, 2004). One function that is generally accepted as a major component of BI involves storing an organization's data in computer cloud storage or in data warehouses. Data warehousing is not an analytics or business analytics function, although the data can be used for analysis. In application,

BI is focused on querying and reporting, but it can include reported information from a BA analysis. BI seeks to answer questions such as what is happening now and where, and also what business actions are needed based on prior experience. BA, on the other hand, can answer questions like why something is happening, what new trends may exist, what will happen next, and what is the best course for the future.

In summary, BA includes the same procedures as in plain analytics but has the additional requirement that the outcome of the analytic analysis must make a measurable impact on business performance. BA includes reporting results like BI but seeks to explain why the results occur based on the analysis rather than just reporting and storing the results, as is the case with BI. Analytics, BA, and BI will be mentioned throughout this book. A review of characteristics to help differentiate these terms is presented in Table 1.3.

Table 1.3 Characteristics of Analytics, Business Analytics, and Business Intelligence

| Characteristics | Analytics | Business Analytics (BA) | Business Intelligence (BI) |
|--|--|---|---|
| Business performance planning role | What is happening, and what will be happening? | What is happening now, what will be happening, and what is the best strategy to deal with it? | What is happening now, and what have we done in the past to deal with it? |
| Use of descriptive analytics as a major component of analysis | Yes | Yes | Yes |
| Use of predictive analytics as a major component of analysis | Yes | Yes | No (only historically) |
| Use of prescriptive analytics as a major component of analysis | Yes | Yes | No (only historically) |
| Use of all three in combination | No | Yes | No |
| Business focus | Maybe | Yes | Yes |
| Focus of storing and maintaining data | No | No | Yes |
| Required focus of improving business value and performance | No | Yes | No |

1.2 Business Analytics Process

The complete *business analytic process* involves the three major component steps applied sequentially to a source of data (see Figure 1.1). The outcome of the business analytic process must relate to business and seek to improve business performance in some way.

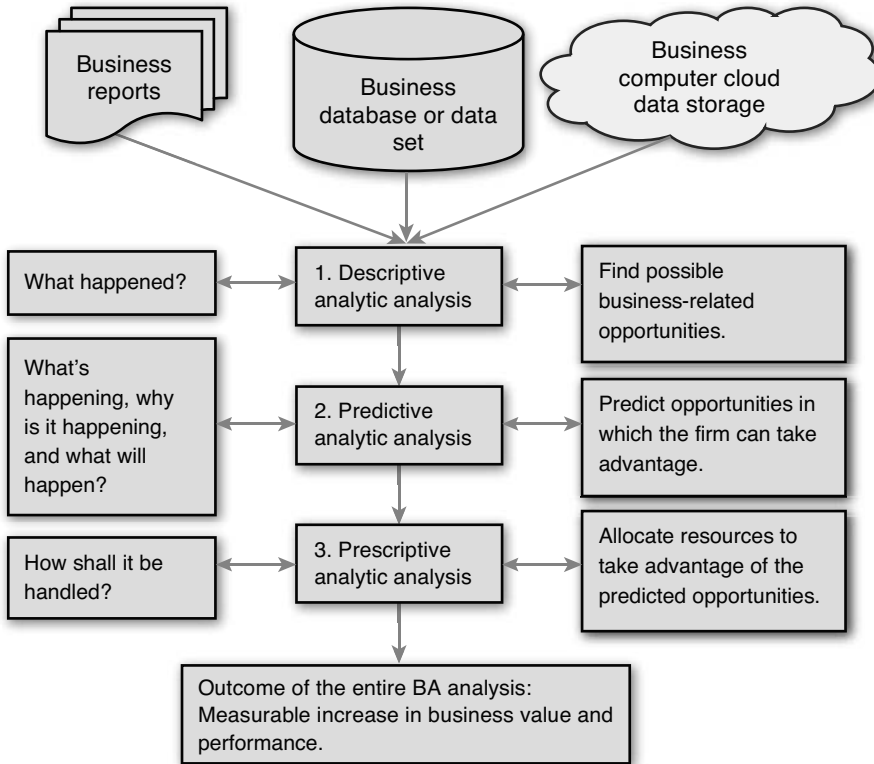


Figure 1.1 Business analytic process

The logic of the BA process in Figure 1.1 is initially based on a question: What valuable or problem-solving information is locked up in the sources of data that an organization has available? At each of the three steps that make up the BA process, additional questions need to be answered, as shown in Figure 1.1. Answering all these questions requires mining the information out of the data via the three steps of analysis that comprise the BA process. The analogy of digging in a mine is appropriate for the BA process because finding new, unique, and valuable information that can lead to a successful strategy is just as good as finding gold in a mine. SAS, a major analytic

corporation (www.sas.com), actually has a step in its BA process, *Query Drilldown*, which refers to the mining effort of questioning and finding answers to pull up useful information in the BA analysis. Many firms routinely undertake BA to solve specific problems, while other firms undertake BA to explore and discover new knowledge to guide organizational planning and decision-making to improve business performance.

The size of some data sources can be unmanageable, overly complex, and generally confusing. Sorting out data and trying to make sense of its informational value requires the application of descriptive analytics as a first step in the BA process. One might begin simply by sorting the data into groups using the four possible classifications presented in Table 1.4. Also, incorporating some of the data into spreadsheets like Excel and preparing cross tabulations and contingency tables are means of restricting the data into a more manageable data structure. Simple measures of central tendency and dispersion might be computed to try to capture possible opportunities for business improvement. Other descriptive analytic summarization methods, including charting, plotting, and graphing, can help decision makers visualize the data to better understand content opportunities.

Table 1.4 Types of Data Measurement Classification Scales

| Type of Data | Description |
|--------------------------|---|
| Measurement Scale | Description |
| Categorical Data | Data that is grouped by one or more characteristics. Categorical data usually involves cardinal numbers counted or expressed as percentages. Example 1: Product markets that can be characterized by categories of “high-end” products or “low-income” products, based on dollar sales. It is common to use this term to apply to data sets that contain items identified by categories as well as observations summarized in cross-tabulations or contingency tables. |
| Ordinal Data | Data that is ranked or ordered to show relational preference. Example 1: Football team rankings not based on points scored but on wins. Example 2: Ranking of business firms based on product quality. |
| Interval Data | Data that is arranged along a scale where each value is equally distant from others. It is ordinal data. Example 1: A temperature gauge. Example 2: A survey instrument using a Likert scale (that is, 1, 2, 3, 4, 5, 6, 7), where 1 to 2 is perceived as equidistant to the interval from 2 to 3, and so on. Note: In ordinal data, the ranking of firms might vary greatly from first place to second, but in interval data, they would have to be relationally proportional. |
| Ratio Data | Data expressed as a ratio on a continuous scale. Example 1: The ratio of firms with green manufacturing programs is twice that of firms without such a program. |

From Step 1 in the *Descriptive Analytic analysis* (see Figure 1.1), some patterns or variables of business behavior should be identified representing targets of business opportunities and possible (but not yet defined) future trend behavior. Additional effort (more mining) might be required, such as the generation of detailed statistical reports narrowly focused on the data related to targets of business opportunities to explain what is taking place in the data (what happened in the past). This is like a statistical search for predictive variables in data that may lead to patterns of behavior a firm might take advantage of if the patterns of behavior occur in the future. For example, a firm might find in its general sales information that during economic downturns, certain products are sold to customers of a particular income level if certain advertising is undertaken. The sales, customers, and advertising variables may be in the form of any of the measurable scales for data in Table 1.4, but they have to meet the three conditions of BA previously mentioned: clear relevancy to business, an implementable resulting insight, and performance and value measurement capabilities.

To determine whether observed trends and behavior found in the relationships of the descriptive analysis of Step 1 actually exist or hold true and can be used to forecast or predict the future, more advanced analysis is undertaken in Step 2, *Predictive Analytic analysis*, of the BA process. There are many methods that can be used in this step of the BA process. A commonly used methodology is multiple regression. (See Appendix A, “Statistical Tools,” and Appendix E, “Forecasting,” for a discussion on multiple regression and ANOVA testing.) This methodology is ideal for establishing whether a statistical relationship exists between the predictive variables found in the descriptive analysis. The relationship might be to show that a dependent variable is predictively associated with business value or performance of some kind. For example, a firm might want to determine which of several promotion efforts (independent variables measured and represented in the model by dollars in TV ads, radio ads, personal selling, and/or magazine ads) is most efficient in generating customer sale dollars (the dependent variable and a measure of business performance). Care would have to be taken to ensure the multiple regression model was used in a valid and reliable way, which is why ANOVA and other statistical confirmatory analyses are used to support the model development. Exploring a database using advanced statistical procedures to verify and confirm the best predictive variables is an important part of this step in the BA process. This answers the questions of what is currently happening and why it happened between the variables in the model.

A single or multiple regression model can often forecast a trend line into the future. When regression is not practical, other forecasting methods (exponential smoothing, smoothing averages) can be applied as predictive analytics to develop needed forecasts of business trends. (See Appendix E.) The identification of future

trends is the main output of Step 2 and the predictive analytics used to find them. This helps answer the question of what will happen.

If a firm knows where the future lies by forecasting trends as they would in Step 2 of the BA process, it can then take advantage of any possible opportunities predicted in that future state. In Step 3, *Prescriptive Analytics analysis*, operations research methodologies can be used to optimally allocate a firm's limited resources to take best advantage of the opportunities it found in the predicted future trends. Limits on human, technology, and financial resources prevent any firm from going after all opportunities they may have available at any one time. Using prescriptive analytics allows the firm to allocate limited resources to optimally achieve objectives as fully as possible. For example, *linear programming* (a constrained optimization methodology) has been used to maximize the profit in the design of supply chains (Paksoy et al., 2013). (Note: Linear programming and other optimization methods are presented in Appendixes B, "Linear Programming," C, "Duality and Sensitivity Analysis in Linear Programming," and D, "Integer Programming.") This third step in the BA process answers the question of how best to allocate and manage decision-making in the future.

In summary, the three major components of descriptive, predictive, and prescriptive analytics arranged as steps in the BA process can help a firm find opportunities in data, predict trends that forecast future opportunities, and aid in selecting a course of action that optimizes the firm's allocation of resources to maximize value and performance. The BA process, along with various methodologies, will be detailed in Chapters 5 through 10.

1.3 Relationship of BA Process and Organization Decision-Making Process

The BA process can solve problems and identify opportunities to improve business performance. In the process, organizations may also determine strategies to guide operations and help achieve competitive advantages. Typically, solving problems and identifying strategic opportunities to follow are organization decision-making tasks. The latter, identifying opportunities, can be viewed as a problem of strategy choice requiring a solution. It should come as no surprise that the BA process described in Section 1.2 closely parallels classic organization decision-making processes. As depicted in Figure 1.2, the business analytic process has an inherent relationship to the steps in typical organization decision-making processes.

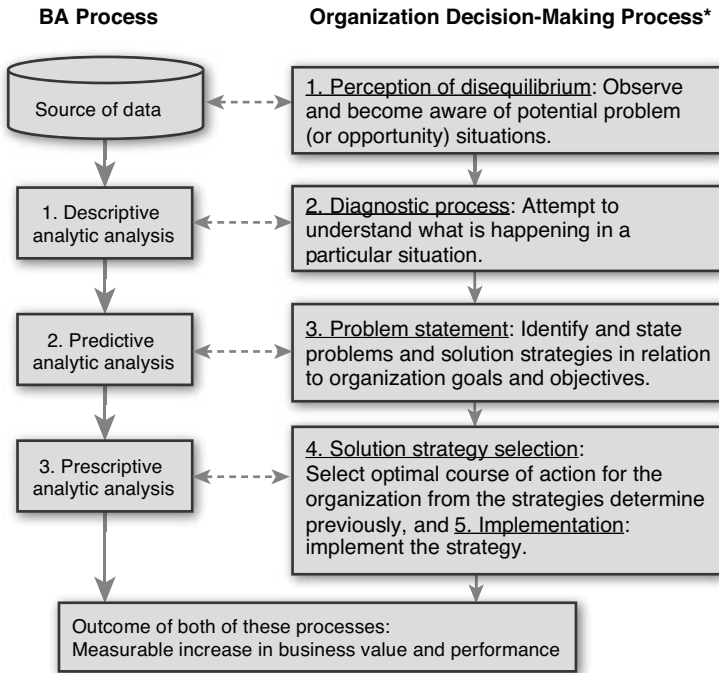


Figure 1.2 Comparison of business analytics and organization decision-making processes

*Source: Adapted from Figure 1 in Elbing (1970), pp. 12–13.

The *organization decision-making process* (ODMP) developed by Elbing (1970) and presented in Figure 1.2 is focused on decision making to solve problems but could also be applied to finding opportunities in data and deciding what is the best course of action to take advantage of them. The five-step ODMP begins with the perception of disequilibrium, or the awareness that a problem exists that needs a decision. Similarly, in the BA process, the first step is to recognize that databases may contain information that could both solve problems and find opportunities to improve business performance. Then in Step 2 of the ODMP, an exploration of the problem to determine its size, impact, and other factors is undertaken to diagnose what the problem is. Likewise, the BA descriptive analytic analysis explores factors that might prove useful in solving problems and offering opportunities. The ODMP problem statement step is similarly structured to the BA predictive analysis to find strategies, paths, or trends that clearly define a problem or opportunity for an organization to solve problems. Finally, the ODMP's last steps of strategy selection and implementation involve the same kinds of tasks that the BA process requires in the final prescriptive step (make an

optimal selection of resource allocations that can be implemented for the betterment of the organization).

The decision-making foundation that has served ODMF for many decades parallels the BA process. The same logic serves both processes and supports organization decision-making skills and capacities.

1.4 Organization of This Book

This book is designed to answer three questions about BA:

- What is it?
- Why is it important?
- How do you do it?

To answer these three questions, the book is divided into three parts. In Part I, “What Are Business Analytics?”, Chapter 1 answers the “what” question. In Part II, the “why” question is answered in Chapter 2, “Why Are Business Analytics Important?” and Chapter 3, “What Resource Considerations Are Important to Support Business Analytics?”

Knowing the importance of explaining how BA is undertaken, the rest of the book’s chapters and appendixes are devoted to answering that question. Chapter 4, “How Do We Align Resources to Support Business Analytics within an Organization?”, explains how an organization needs to support BA. Chapter 5, “What Are Descriptive Analytics?”, Chapter 6, “What Are Predictive Analytics?”, and Chapter 7, “What Are Prescriptive Analytics?”, detail and illustrate the three respective steps in the BA process. To further illustrate how to conduct a BA analysis, Chapter 8, “A Final Case Study Illustration,” provides an example of BA. Supporting the analytic discussions is a series of analytically oriented appendixes that follow Chapter 8.

Part III includes quantitative analyses utilizing computer software. In an effort to provide some diversity of software usage, SPSS, Excel, and LINGO software output are presented. SPSS and LINGO can be used together to duplicate the analysis in this book, or only Excel with the necessary add-ins can be used. Because of the changing nature of software and differing educational backgrounds, this book does not provide extensive software explanation.

In addition to the basic content that makes up the body of the chapters, there are pedagogy enhancements that can aid learning. All chapters begin with chapter objectives and end with a summary, discussion questions, and, where needed, references.

In addition, Chapters 5 through 8 have sample problems with solutions, as well as additional assignment problems.

Some of the more detailed explanations of methodologies are presented in the appendixes. Their positioning in the appendixes is designed to enhance content flow and permit more experienced readers a flexible way to select only the technical content they might want to use. An extensive index allows quick access to terminology.

Summary

This chapter has introduced important terminology and defined business analytics in terms of a unique process useful in securing information on which decisions can be made and business opportunities seized. Data classification measurement scales were also briefly introduced to aid in understanding the types of measures that can be employed in BA. The relationship of the BA process and the organization decision-making process was explained in terms of how they complement each other. This chapter ended with a brief overview of this book's organization and how it is structured to aid learning.

Knowing *what* business analytics are about is important, but equally important is knowing *why* they are important. Chapter 2 begins to answer the question.

Discussion Questions

1. What is the difference between analytics and business analytics?
2. What is the difference between business analytics and business intelligence?
3. Why are the steps in the business analytics process sequential?
4. How is the business analytics process similar to the organization decision-making process?
5. Why does interval data have to be relationally proportional?

References

- Bartlett, R. (2013) *A Practitioner's Guide to Business Analytics*. McGraw-Hill, New York, NY.
- Elbing, A.O. (1970) *Behavioral Decisions in Organizations*. Scott Foresman and Company, Glenview, IL.
- Isson, J.P., Harriott, J.S. (2013) *Win with Advanced Business Analytics*. John Wiley & Sons, Hoboken, NJ.
- Negash, S. (2004) "Business Intelligence." *Communications of the Association of Information Systems*. Vol. 13, pp. 177–195.
- Paksoy, T., Ozkeylan, E., Weber, G.W. (2013) "Profit-Oriented Supply Chain Network Optimization." *Central European Journal of Operations Research*. Vol. 21, No. 2, pp. 455–478.
- Stubbs, E. (2011) *The Value of Business Analytics*. John Wiley & Sons, Hoboken, NJ.

This page intentionally left blank

Index

A

- addition, rules of, 173-174
- additive time series model, 274
- additivity in LP (Linear Programming) models, 232
- administrators, 31
- aligning business analytics, 45-46
 - management issues, 54
 - change management*, 58-59
 - ensuring data quality*, 55-57
 - establishing information policy*, 54
 - measuring business analytics contribution*, 58
 - outsourcing business analytics*, 55
 - organization structures, 46-50
 - centralized BA organization structure*, 49-50
 - functional organization structure*, 48
 - hierarchical relationships*, 46
 - matrix organization structure*, 48
 - project structure*, 47-48
 - reasons for BA initiative and organization failure*, 51-50
 - teams, 50-53
 - collaboration*, 50-53
 - participant roles*, 52
 - reasons for team failures*, 53
- alternatives (DT), 304
- Analysis ToolPak, 39
- analytics. *See also* DT (decision theory)
 - alignment. *See* business analytics alignment
 - analytic purposes and tools, 5
 - business analytics personnel, 30-33
 - administrators*, 31
 - BAP (Business Analytics Professional) exam*, 30-31
 - designers*, 31
 - developers*, 31
 - skills and competency requirements*, 32-33
 - solution experts*, 31
 - technical specialists*, 31
 - business analytics process
 - data measurement scales*, 8
 - explained*, 7-10
 - relationship with organization decision-making process (ODMP)*, 10-12
 - characteristics of, 6
 - decision analysis. *See* DT (decision theory)
 - definition of, 3-4
 - descriptive analytics
 - analytic purposes and tools*, 5
 - confidence intervals*, 76-77
 - definition of*, 4
 - descriptive statistics*, 67-72
 - marketing/planning case study example*, 80-90
 - overview*, 63-64
 - probability distributions*, 78-80
 - sampling estimation*, 76-77
 - sampling methods*, 73-75
 - statistical charts*, 64-67
 - supply chain shipping problem case study*, 141-145

- predictive analytics
 - analytic purposes and tools*, 5
 - data mining*, 97-102
 - data-driven models*, 96-97
 - definition of*, 4
 - logic-driven models*, 94-96
 - marketing/planning case study example*, 102-114
 - methodologies*, 119-120
 - overview*, 93-94
 - prescriptive modeling*, 120-122
 - supply chain shipping problem case study*, 147-157
 - prescriptive analytics
 - analytic purposes and tools*, 5
 - definition of*, 4
 - integer programming*. *See IP (integer programming)*
 - regression analysis, 97
 - Durbin-Watson Autocorrelation Test*, 284
 - multiple regression models*, 281-284
 - simple regression model*, 276-281
 - sensitivity analysis
 - economic value of resources, determining*, 258-259
 - overview*, 242-243
 - primal maximization problems*, 243-251
 - primal minimization problems*, 251-258
 - analytics analysts, 51
 - analytics modelers, 51
 - analytics process designers, 51
 - ANOVA testing, 9, 195
 - applications of business analytics to
 - enhance decision-making, 23-24
 - applied LP (Linear Programming) model, 202
 - area charts, 65
 - artificial variables, 219
 - assessing probability
 - Frequency Theory, 171-172
 - Principle of Insufficient Reason, 172
 - rules of addition, 173-174
 - rules of multiplication, 174-177
 - associations, 39, 99
 - averages, smoothing, 286-288
- ## B
- BA team heads, 51
 - backward decision method, 317-320
 - BAP (Business Analytics Professional) exam, 30-31
 - bar charts, 65
 - Bayes's theorem, 321-328
 - belief of physical proximity, 51
 - BI (business intelligence), 5-6
 - billing and reminder systems, 34
 - binomial probability distribution, 179-181
 - binomial tests, 199
 - blending formulations, 230
 - branch-and-bound method, 264-267
 - business analytics alignment, 45-46
 - management issues, 54
 - change management*, 58-59
 - ensuring data quality*, 55-57
 - establishing information policy*, 54
 - measuring business analytics contribution*, 58
 - outsourcing business analytics*, 55
 - organization structures, 46-50
 - centralized BA organization structure*, 49-50
 - functional organization structure*, 48
 - hierarchical relationships*, 46
 - matrix organization structure*, 48
 - project structure*, 47-48
 - reasons for BA initiative and organization failure*, 51-50
 - teams, 50-53
 - collaboration*, 50-53
 - participant roles*, 52
 - reasons for team failures*, 53
 - business analytics personnel, 30-33
 - administrators, 31
 - BAP (Business Analytics Professional) exam, 30-31
 - designers, 31
 - developers, 31

skills and competency requirements, 32-33

solution experts, 31

technical specialists, 31

business analytics process

- data measurement scales, 8
- explained, 7-10
- relationship with organization
 - decision-making process (ODMP), 10-12

Business Analytics Professional (BAP)

- exam, 30-31

business domain experts, 52

business intelligence (BI), 5-6

business performance tracking, 24

butcher problem example (LP), 208-210

C

CAP (Certified Analytic Professional), 30

case studies

- explained, 121
- marketing/planning case study example.
 - See* marketing/planning case study example
- supply chain shipping problem case study
 - descriptive analytics analysis, 141-145*
 - predictive analytics analysis, 147-157*
 - prescriptive analysis, 158-163*
 - problem background and data, 139-140*

categorical data, 8

categorizing data, 33-35

cause-and-effect diagrams, 95

centralized BA organization structure, 49-50

certainty

- decision-making under certainty, 306
 - maximax criterion, 306*
 - maximin criterion, 307*
- explained, 304
- in LP (Linear Programming) models, 232

Certified Analytic Professional (CAP), 30

championing change, 59

change management, 58-59

- best practices, 59-60
- targets, 59

charts

- marketing/planning case study example
 - case study background, 81*
 - descriptive analytics analysis, 82-90*
- statistical charts, 65-67

CHISQ.TEST, 199

Chi-Square tests, 199

Claritas, 35

Clarke Special Parts problem example, 214-215

classification, 39, 99

clearly stated goals, 59

cluster random sampling, 73

clustering

- data mining, 39, 99
- hierarchical clustering, 100
- K-mean clustering, 100-102

coding, checking for, 57

coefficient of kurtosis, 68

coefficient of skewedness, 68

Cognizure BAP (Business Analytics Professional) exam, 30-31

collaboration

- lack of, 50
- in teams, 50-53

column charts, 65

combinations, 169

communication

- good communication, 59
- lack of, 53

competency requirements for business analytics personnel, 32-33

competition data sources, 35

competitive advantage

- achieving with business analytics, 20-21
- innovation, 21
- operations efficiency, 21
- price leadership, 21
- product differentiation, 21
- service effectiveness, 21
- sustainability, 21

completeness, checking for, 57

computer simulation methods, 301
 conditional probabilities, 176
 confidence coefficient, 79
 confidence intervals, 76-77
 constrained optimization models, 128-129
 constraints
 formulating, 130-131
 LP (Linear Programming), 204-206
 continuous probability distributions, 185-192
 exponential probability distribution, 190-192
 normal probability distribution, 186-189
 correlation analysis, 97
 counting, 167
 combinations, 169
 permutations, 167-168
 repetitions, 170
 credit union example of business analysis, 19
 CRM (customer relationship management) systems, 34
 culture as target of change management, 59
 current data, checking for, 57
 Curve Estimation (SPSS), 288-289
 curve fitting
 explained, 123-129
 SPSS Curve Estimation, 288-289
 supply chain shipping problem case study, 147-154
 customer demographics, 35
 customer internal data, 34
 customer profitability, increasing, 23
 customer relationship management (CRM) systems, 34
 customer satisfaction, 35
 customer service problem example (LP), 213-214
 cyclical variation, 275

D

data inspection items, 57
 data management technology, 37
 data managers, 52
 data marts, 38
 data measurement scales, 8
 data mining, 38-40, 97-98
 methodologies, 99-102
 discriminant analysis, 100
 hierarchical clustering, 100
 K-mean clustering, 100-102
 logistic regression, 100
 neural networks, 100
 types of information, 99
 simple illustration of, 98-99
 data privacy, 36
 data quality
 ensuring, 55-57
 overview, 35-36
 data sets, 3
 data sources
 categorizing data, 33-35
 data privacy, 35-36
 data quality, 35-36
 external sources, 34-35
 internal sources, 34
 new sources of data, applying business analytics to, 23-25
 data visualization
 marketing/planning case study example
 case study background, 81
 descriptive analytics analysis, 82-90
 statistical charts, 64-67
 data warehouses, 38
 database management systems (DBMS), 37-36
 databases, 3
 database encyclopedia content, 36
 DBMS (database management systems), 37-36
 data-driven models, 96-97
 DBMS (database management systems), 37-36
 decision environments. *See also* DT (decision theory)
 certainty
 decision-making under certainty, 306-307
 explained, 304

- risk
 - decision-making under risk, 307-311*
 - explained, 304*
- uncertainty
 - decision-making under uncertainty, 311-315*
 - explained, 305*
- decision theory. *See* DT (decision theory)
- decision trees, 317-320
- decision variables, defining, 130
- delegation of responsibility, 51
- descriptive analytics
 - analytic purposes and tools, 5
 - confidence intervals, 76-77
 - definition of, 4
 - descriptive statistics, 67-72
 - marketing/planning case study
 - example, 80
 - case study background, 81*
 - descriptive analytics analysis, 82-90*
 - overview, 63-64
 - probability distributions, 78-80
 - sampling estimation, 76-77
 - sampling methods, 73-75
 - statistical charts, 65-67
 - supply chain shipping problem case study, 141-145
 - actual monthly customer demand in motors, 143*
 - Chicago customer demand (graph), 143*
 - estimated shipping costs per motor, 141*
 - Excel summary statistics of actual monthly customer demand in motors, 144*
 - Houston customer demand (graph), 143*
 - Kansas City customer demand (graph), 145*
 - Little Rock customer demand (graph), 145*
 - Oklahoma City customer demand (graph), 145*
 - Omaha customer demand (graph), 145*
 - problem background and data, 140*
 - SPSS summary statistics of actual monthly customer demand in motors, 144*
- designers, 31
- deterministic simulation, 295-296
- developers, 31
- diagrams
 - cause-and-effect diagrams, 95
 - influence diagrams, 95-96
- diet problem example (LP), 210-212
- differential calculus, 134
- digital analytics, 23-25
- discrete probability distributions, 178-184
 - binomial probability distribution, 179-181
 - geometric probability distribution, 184
 - hypergeometric probability distribution, 184
 - Poisson probability distribution, 182-184
- discriminant analysis, 100
- divisibility in LP (Linear Programming) models, 232
- downloading LINGO, 220
- DT (decision theory)
 - Bayes's theorem, 321-328
 - decision-making under certainty, 306
 - maximax criterion, 306*
 - maximin criterion, 307*
 - decision-making under risk, 307
 - EV (expected value) criterion, 308-309*
 - expected opportunity loss criterion, 309-311*
 - origin of probabilities, 308*
 - decision-making under uncertainty, 311
 - Hurwicz criterion, 312-313*
 - Laplace criterion, 311-312*
 - maximax criterion, 312*
 - maximin criterion, 312*
 - minimax criterion, 313-315*
 - enhancing decision-making with business analytics, 23-24
 - EVPI (expected value of perfect information), 315

- model elements, 304
- model formulation, 305-306
- overview, 122, 303
- practice problems, 328-333
- sequential decisions and decision trees, 317-320
- types of decision environments, 304-305
- duality**
 - duality practice problems, 259-261
 - economic value of resources, determining, 258-259
 - informational value of, 242
 - overview, 241
 - primal maximization problems, 243-251
 - primal minimization problems, 251-258
- Dun & Bradstreet, 35**
- duplication, checking for, 57**
- Durbin-Watson Autocorrelation Test, 284**

E

- economic data sources, 35
- economic value of resources, determining, 258-259
- ensuring data quality, 55-57
- enterprise resource planning (ERP) systems, 34
- Equifax, 35
- ERP (enterprise resource planning) systems, 34
- errors
 - confidence intervals, 76-77
 - error metrics, 291-292
- establishing information policy, 54
- estimating sampling, 76-77
- EV (expected value) criterion, 308-309
- EVPI (expected value of perfect information), 315
- Excel
 - computer-based solution with simplex method, 224-227
 - LP (Linear Programming) solutions
 - infeasible solutions*, 229
 - practice problems*, 233-238
 - unbounded solutions*, 227-228

- marketing/planning case study example
 - case study background*, 81, 103
 - descriptive analytics analysis*, 82-90
 - predictive analytics analysis*, 104-114
 - solution for LP marketing/planning model*, 132-133
- primal maximization problems, 243-251
- primal minimization problems, 251-258
- simple regression model, 277-280
- supply chain shipping problem case study, 144
- t-test statistics, 197
- ZOP (zero-one programming) problems/models, solving, 268-269
- executive sponsorship, lack of, 51
- expected opportunity loss criterion, 309-311
- expected value (EV) criterion, 308-309
- expected value of perfect information (EVPI), 315
- experiments, 177
- exponential probability distribution, 190-192
- exponential smoothing
 - example of, 285
 - simple model, 284-285
 - smoothing averages, 286-288
- external data sources, 34-35

F

- factorials, 168
- failures
 - failure to deliver, 53
 - failure to provide value, 53
 - reasons for BA initiative and organization failure, 50-51
 - reasons for team failures, 53
- farming problem example (LP), 212-213
- Federal Division problem example (LP), 215-217
- finiteness in LP (Linear Programming) models, 232
- fitting models to data, 288-289
- forecasting
 - additive time series model, 274
 - data mining, 39, 99

exponential smoothing
 example of, 285
 simple model, 284-285
 fitting models to data, 288-289
 forecasting accuracy statistics, 291-292
 MAD (*mean absolute deviation*),
 291-292
 MAPE (*mean absolute percentage error*), 292
 MSE (*mean square error*), 291-292
 forecasting methods, 275-276
 marketing/planning case study
 example, 112
 multiple regression models, 281
 application, 282-283
 limitations in forecasting time series data, 283-284
 multiplicative time series model, 274
 overview, 97, 271
 practice problems, 292-293
 simple regression model
 computer-based solution, 277-280
 model for trend, 276
 statistical assumptions and rules,
 280-281
 smoothing averages, 286-288
 supply chain shipping problem case study
 developing forecasting models,
 147-154
 resulting warehouse customer demand forecasts, 157
 validating forecasting models,
 155-157
 time series data, variation in
 cyclical variation, 275
 random variation, 275
 seasonal variation, 274
 trend variation, 274
 variation in time series data, 272-274
 formulating DT (decision theory) models,
 305-306
 F-ratio statistic, 110
 Frequency Theory, 171-172
 F-Test Two-Sample for Variances tool, 195
 functional organization structure, 48
 functions, objective, 203-204

G

generalized LP (Linear Programming)
 model, 202
 geometric probability distribution, 184
 given requirements, stating, 131, 206
 goals, 59
 Google Insights for Search, 39
 Google Trends, 39

H

hardware, 37
 hierarchical clustering, 100
 hierarchical relationships, 46
 histograms, 66
 human resources
 decisions, 23
 human resources data, 34
 lack of, 51
 Hurwicz criterion, 312-313
 hypergeometric probability
 distribution, 184
 hypothesis testing, 193-199

I

IBM's SPSS software, 40
 IMF (International Monetary Fund), 35
 implementation specialists, 52
 importance of business analytics
 applications to enhance decision-making,
 23-24
 new sources of data, 23-25
 overview, 17-18
 providing answers to questions, 18-20
 strategy for competitive advantage, 20-21
 inability to delegate responsibility, 51
 inability to prove success, 53
 inconsistent values, checking for, 57
 increasing customer profitability, 24
 infeasible solutions, 229
 influence diagrams, 95-96
 information policy, establishing, 54
 information technology (IT)
 computer hardware, 36

computer software, 36
 data management technology, 37
 data marts, 38
 data mining, 38-40
 data warehouses, 38
 database encyclopedia content, 36
 DBMS (database management systems),
 37-36
 infrastructure, 37
 networking and telecommunications
 technology, 37
INFORMS, 30
 innovation, achieving with business
 analytics, 21
Insufficient Reason, Principle of, 172
integer programming. *See* IP (integer
 programming)
integrated processes, lack of, 51
internal data sources, 34
International Monetary Fund (IMF), 35
interval data, 8
IP (integer programming), 121, 263
 explained, 263-264
 IP problems/models, solving, 264
 maximization IP problem, 265-266
 minimization IP problem, 266-267
 practice problems, 270
 ZOP (zero-one programming)
 explained, 264
 problems/models, solving, 268-269
IT (information technology)
 computer hardware, 37
 computer software, 37
 data management technology, 37
 data marts, 38
 data mining, 38-40
 data warehouses, 38
 database encyclopedia content, 36
 DBMS (database management systems),
 37-36
 infrastructure, 37
 networking and telecommunications
 technology, 37

J-K

judgment sampling, 74
justification, lack of, 53
K-mean clustering, 101-102
Kolmogorov-Smirnov (One-Way) tests, 199
Kurtosis, 69

L

Laplace criterion, 311-312
leadership, lack of, 50
limited context perception, 50
Lindo Systems LINGO. *See* LINGO
line charts
 explained, 66
 marketing/planning case study example
 case study background, 81
 descriptive analytics analysis, 82-90
Linear Programming. *See* LP (Linear
 Programming)
linearity in LP (Linear Programming)
 models, 232
LINGO, 40
 downloading, 220
 IP problems/models, solving
 maximization IP problem, 265-266
 minimization IP problem, 266-267
 LP (Linear Programming) solutions
 computer-based solution with
 simplex method, 220-224
 infeasible solutions, 229
 marketing/planning case study
 example, 132-133
 practice problems, 233-238
 unbounded solutions, 227-228
 overview, 40
 primal maximization problems, 243-251
 primal minimization problems, 251-258
 supply chain shipping problem case
 study, 159-161
 trial versions, 220
 ZOP (zero-one programming) problems/
 models, solving, 268-269

little data, 3
 logic-driven models, 94-96
 logistic regression, 100
 loss values, expected opportunity loss
 criterion, 309-311
LP (Linear Programming)
 applied LP model, 121, 202
 blending formulations, 230
 computer-based solutions with simplex
 method, 217-218
 Excel solution, 224-227
 LINGO solution, 220-224
 simplex variables, 218-220
 constraints, 204-206
 duality
 duality practice problems, 259-261
 economic value of resources,
 determining, 258-259
 informational value of, 242
 overview, 241
 primal maximization problems,
 243-251
 primal minimization problems,
 251-258
 sensitivity analysis, 242-243
 generalized LP model, 202
 infeasible solutions, 229
 maximization models, 201-202
 minimization models, 201-202
 multidimensional decision variable
 formulations, 231
 necessary assumptions, 232
 nonnegativity and given
 requirements, 206
 objective function, 203-204
 overview, 201-202
 practice problems, 233-238
 problem/model formulation
 butcher problem example, 208-210
 Clarke Special Parts problem
 example, 214-215
 customer service problem example,
 213-214
 diet problem example, 210-212

farming problem example, 212-213
 Federal Division problem example,
 215-217
 stepwise procedure, 207-208
 unbounded solutions, 227-228

M

MAD (mean absolute deviation), 155-157,
 162, 291-292
management issues, 54
 change management, 58-59
 best practices, 59
 targets, 59
 ensuring data quality, 55-57
 establishing information policy, 54
 measuring business analytics
 contribution, 58
 outsourcing business analytics, 55
 advantages of, 55
 disadvantages of, 56
**MAPE (mean absolute percentage
 error)**, 292
marginal probability, 321
marketing/planning case study example,
 80-90
 case study background, 81, 103, 129
 descriptive analytics analysis, 82-90
 predictive analytics analysis, 104-114
 Excel best variable combination
 regression model and statistics, 113
 Excel POS regression model, 108
 Excel radio regression model, 109
 Excel TV regression model, 109
 forecasting model, 112
 F-ratio statistic, 110
 R-Square statistics, 110-111
 SPSS best variable combination
 regression model and statistics, 106
 SPSS Pearson correlation coefficients,
 104
 SPSS POS regression model, 106
 SPSS radio regression model, 107
 SPSS TV regression model, 108

- prescriptive analysis, 102-103, 129-134
 - final comments*, 133-134
 - formulation of LP marketing/ planning model*, 130-131
 - solution for LP marketing/planning model*, 132-133
- matrix organization structure, 48-49
- maximax criterion, 306, 312
- maximin criterion, 307, 312
- maximization IP problem, solving, 265-266
- maximization models
 - LP (Linear Programming), 201-202
 - primal maximization problems, 243-251
- maximum/minimum, 68
- mean, 68
- mean absolute deviation (MAD), 155-157, 162, 291-292
- mean absolute percentage error (MAPE), 292
- mean square error (MSE), 291-292
- measured performance, 59
- measuring business analytics
 - contribution, 58
- median, 68
- merchandise strategy optimization, 23
- methods, sampling, 73-75
- Microsoft Excel, 39
- minimax criterion, 313-315
- minimization IP problem, solving, 266-267
- minimization models
 - LP (Linear Programming), 201-202
 - primal minimization problems, 251-258
- minimum/maximum, 68
- mobile analytics, 25
- mode, 68
- modeling
 - constrained optimization models, 128-129
 - DT (decision theory)
 - decision environments*, 304-305
 - model elements*, 304
 - model formulation*, 305-306
 - overview*, 303
 - exponential smoothing
 - example of*, 285
 - simple model*, 284-285
 - forecasting models
 - developing*, 147-154
 - exponential smoothing*, 284-285
 - fitting models to data*, 288-289
 - forecasting accuracy statistics*, 291-292
 - forecasting methods*, 275-276
 - multiple regression models*, 281-284
 - practice problems*, 292-293
 - sample warehouse customer demand forecasts*, 157
 - simple regression model*, 276-281
 - smoothing averages*, 286-288
 - statistical assumptions and rules*, 280-281
 - validating*, 155-157
- LP (Linear Programming)
 - applied LP model*, 202
 - blending formulations*, 230
 - computer-based solutions with simplex method*, 217-227
 - constraints*, 204-206
 - generalized LP model*, 202
 - infeasible solutions*, 229
 - maximization models*, 201-202
 - minimization models*, 201-202
 - multidimensional decision variable formulations*, 231
 - necessary assumptions*, 232
 - nonnegativity and given requirements*, 206
 - objective function*, 203-204
 - problem/model formulation*, 207-217
 - unbounded solutions*, 227-228
- predictive modeling
 - data-driven models*, 96-97
 - logic-driven models*, 94-96
- prescriptive modeling, 120-122
 - case studies*, 122
 - decision analysis*, 122
 - integer programming*. *See integer programming*
 - linear programming*. *See LP (Linear Programming)*
 - nonlinear optimization*, 121, 122-129
 - other methodologies*, 122

- simulation, 122, 295
 - deterministic simulation*, 295-296
 - practice problems*, 301
 - probabilistic simulation*, 296-301
 - variation in time series data
 - additive time series model*, 274
 - cyclical variation*, 275
 - multiplicative time series model*, 274
 - random variation*, 275
 - seasonal variation*, 274
 - trend variation*, 274
 - monitoring analysts, 52
 - Monte Carlo simulation method
 - application, 298-301
 - procedure, 296-298
 - MSE (mean square error), 291-292
 - multidimensional decision variable
 - formulations, 231
 - multiple regression models, 9, 281
 - application, 282-283
 - limitations in forecasting time series data, 283-284
 - multiplication, rules of, 174-177
 - multiplicative time series model, 274
- N**
- N function, 67
 - need for business analytics
 - applications to enhance decision-making, 23-24
 - new sources of data, 23-25
 - overview, 17-18
 - providing answers to questions, 18-20
 - strategy for competitive advantage, 20-21
 - networking and telecommunications
 - technology, 37
 - neural networks, 100
 - new sources of data, applying business analytics to, 23-25
 - Nielsen data, 35
 - nonlinear optimization, 121, 122-129
 - calculus methods, 129
 - curve fitting, 123-129, 288-289
 - quadratic programming, 128-129
 - nonnegativity, 131, 206
 - nonparametric hypothesis testing, 200-199
 - normal probability distribution, 186-189
- O**
- objective function, 203-204
 - ODMP (organization decision-making process), 10-12
 - operations efficiency, achieving with business analytics, 21
 - optimization, nonlinear, 121, 122-129
 - calculus methods, 129
 - curve fitting, 123-129, 288-289
 - quadratic programming, 128-129
 - ordinal data, 8
 - organization decision-making process (ODMP), 10-12
 - organization structures, 45-50
 - centralized BA organization structure, 49-50
 - functional organization structure, 48
 - hierarchical relationships, 46
 - matrix organization structure, 48
 - project structure, 47-48
 - reasons for BA initiative and organization failure, 51-50
 - as target of change management, 59
 - organizational planning, 20
 - origin of probabilities, 308
 - outcomes, 177
 - outliers, checking for, 57
 - outsourcing business analytics, 55
 - advantages of, 55
 - disadvantages of, 55-56
- P**
- parametric hypothesis testing, 195-197
 - payoffs (DT), 304
 - period sampling, 74
 - permutations, 167-168
 - personnel, 30-33
 - administrators, 31
 - BAP (Business Analytics Professional) exam, 30-31

- designers, 31
- developers, 31
- skills and competency requirements, 32-33
- solution experts, 31
 - as target of change management, 59
 - technical specialists, 31
- physical proximity, belief of, 50**
- pie charts, 66**
- planning, organizational, 20**
- Poisson probability distribution, 182-184**
- policy, information policy, 54**
- practice problems**
 - DT (decision theory), 328-333
 - forecasting, 292-293
 - IP (integer programming), 270
 - LP (Linear Programming), 233-238
 - simulation, 301
- predictive analytics**
 - analytic purposes and tools, 5
 - data mining, 97-98
 - methodologies, 99-102*
 - simple illustration of, 98-99*
 - data-driven models, 96-97
 - definition of, 4
 - logic-driven models, 94-96
 - marketing/planning case study
 - example, 102
 - case study background, 103*
 - predictive analytics analysis, 104-114*
 - overview, 93-94
 - supply chain shipping problem case study, 147-157
 - developing forecasting models, 147-154*
 - problem background and data, 140*
 - resulting warehouse customer demand forecasts, 157*
 - validating forecasting models, 155-157*
- predictive modeling, logic-driven models, 94-96**
- prescriptive analytics**
 - analytic purposes and tools, 5
 - definition of, 4
- marketing/planning case study example
 - case study background, 129*
 - prescriptive analysis, 129-134*
- methodologies, 119-120
- prescriptive modeling, 120-122
 - case studies, 122*
 - decision analysis, 122*
 - integer programming. See integer programming*
 - linear programming. See LP (Linear Programming)*
 - nonlinear optimization, 121, 122-129*
 - other methodologies, 122*
 - simulation, 122*
- supply chain shipping problem case study, 158-163
 - demonstrating business performance improvement, 162-163*
 - determining optimal shipping schedule, 159-161*
 - problem background and data, 140*
 - selecting and developing optimization shipping model, 158-159*
 - summary of BA procedure for manufacturer, 161-162*
- prescriptive modeling, 120-122**
 - case studies, 122
 - decision analysis, 122
 - integer programming, 122
 - IP (integer programming)
 - explained, 263-264*
 - IP problems/models, solving, 264-267*
 - practice problems, 270*
 - ZOP (zero-one programming) problems/models, solving, 264, 268-269*
 - linear programming. *See LP (Linear Programming)*
 - nonlinear optimization, 121, 122-129
 - calculus methods, 129*
 - curve fitting, 123-129, 288-289*
 - quadratic programming, 128-129*
 - other methodologies, 122
 - simulation, 122
- price leadership, achieving with business analytics, 21**

primal maximization problems, 243-251

primal minimization problems, 251-258

Principle of Insufficient Reason, 172

privacy (data), 35-36

probabilistic simulation

Monte Carlo simulation method

application, 298-301

procedure, 296-298

overview, 296

probability. *See also* DT (decision theory)

Bayes's theorem, 321-328

marginal probability, 321

Monte Carlo simulation method,

application, 298-301

origin of probabilities, 308

probabilistic simulation, 296

Monte Carlo simulation method

procedure, 296-298

overview, 296

probability concepts, 171

Frequency Theory, 171-172

Principle of Insufficient Reason, 172

rules of addition, 173-174

rules of multiplication, 174-177

probability distributions, 177-178

binomial probability distribution,

179-181

exponential probability distribution,

190-192

geometric probability distribution,

184

hypergeometric probability

distribution, 184

normal probability distribution,

186-189

Poisson probability distribution,

182-184

random variables, 177

probability distributions, 78-80, 97,

177-178

continuous probability distributions,

185-192

exponential probability distribution,

190-

normal probability distribution,

186-189

discrete probability distributions, 178-184

binomial probability distribution,

179-181

geometric probability distribution,

184

hypergeometric probability

distribution, 184

Poisson probability distribution,

182-184

random variables, 177

process of business analytics

data measurement scales, 8

explained, 7-10

integrated processes, lack of, 51

relationship with organization

decision-making process (ODMP),

10-12

product data, 34

product differentiation, achieving with

business analytics, 21

production data, 34

profit, calculating, 96

project structure, 47-48

providing answers to questions, 18-20

Q

quadratic programming, 127-129

quality of data

ensuring, 56-57

overview, 35-36

Query Drilldown, 8

questionnaires, 34

questions business analytics seeks to

answer, 18

quota sampling, 74

R

random variables, 177

random variation, 275

range, 68

ratio data, 8

reducing risk, 24

regression analysis, 97

Durbin-Watson Autocorrelation Test, 284

- multiple regression models, 281
 - application*, 282-283
 - limitations in forecasting time series data*, 283-284
 - simple regression model
 - computer-based solution*, 277-280
 - model for trend*, 276-281
 - statistical assumptions and rules*, 280-281
 - relevance, checking for, 57
 - repetitions, 170
 - responsibility, inability to delegate, 51
 - risk
 - decision-making under risk, 307
 - EV (expected value) criterion*, 308-309
 - expected opportunity loss criterion*, 309-311
 - origin of probabilities*, 308
 - explained, 304
 - risk reduction, 23
 - roles (team), 52
 - R-Square statistics, 110-111
 - rules of addition, 173-174
 - rules of multiplication, 174-177
 - run testing, 199
- S**
- sampling
 - sample variance, 69
 - sampling estimation, 76-77, 97
 - sampling methods, 73-75
 - SAS Analytics Pro, 7, 40
 - scatter charts, 66
 - seasonal variation, 274
 - sensitivity analysis
 - economic value of resources,
 - determining, 258-259
 - overview, 242-243
 - primal maximization problems, 243-251
 - primal minimization problems, 251-258
 - sequences
 - data mining, 39, 99
 - sequential decisions and decision trees, 317-320
 - sequential decisions, 317-320
 - senior management support, 59
 - service effectiveness, achieving with
 - business analytics, 21
 - simple random sampling, 73
 - simple regression model
 - computer-based solution, 277-280
 - model for trend, 276-281
 - statistical assumptions and rules, 280-281
 - simplex method, 217-218
 - Excel, 224-227
 - LINGO, 220-224
 - simplex variables, 218-220
 - artificial variables*, 219
 - slack variables*, 218-219
 - surplus variables*, 219
 - simplex variables, 218-220
 - artificial variables, 219
 - slack variables, 218-219
 - surplus variables, 219
 - simulation, 97, 122, 295
 - computer simulation methods, 301
 - deterministic simulation, 295-296
 - practice problems, 301
 - probabilistic simulation, 296
 - Monte Carlo simulation method*, 296-298
 - Monte Carlo simulation method application*, 298-301
 - skewedness, 69
 - skill requirements for business analytics
 - personnel, 32-33
 - slack variables, 218-219
 - smoothing averages, 286-288
 - social media analytics, 23-25
 - software, 37. *See also* specific software
 - solution experts, 31
 - Solver, 39
 - SPSS software, 40
 - Curve Estimation, 288-289
 - Curve Fitting, 123-129, 148-153
 - K-Mean cluster software, 101-102
 - marketing/planning case study example
 - case study background*, 81, 103
 - descriptive analytics analysis*, 82-90
 - predictive analytics analysis*, 104-114

- simple regression model, 277-280
- supply chain shipping problem case study, 138
 - t-test statistics, 197
- standard deviation, 68
- standard error, 69
- standard normal probability distribution, 78
- states of nature (DT), 304
- statistical charts, 65-67
- statistical testing, 193-199
- statistical tools, 167
 - counting, 167
 - combinations, 169
 - permutations, 167-168
 - repetitions, 170
 - descriptive statistics, 67-72
 - probability
 - rules of addition, 173-174
 - rules of multiplication, 174-177
 - probability concepts, 171
 - conditional probabilities, 176
 - Frequency Theory, 171-172
 - Principle of Insufficient Reason, 172
 - probability distributions, 177-178
 - binomial probability distribution, 179-181
 - exponential probability distribution, 190-192
 - geometric probability distribution, 184
 - hypergeometric probability distribution, 184
 - normal probability distribution, 186-189
 - Poisson probability distribution, 182-184
 - random variables, 177
 - statistical charts, 64-67
 - statistical testing, 193-199
- strategy for competitive advantage, 20-21
- stratified random sampling, 73
- structured data analytics, 25
- success, proving, 53
- sum, 67
- supply chain shipping problem case study
 - descriptive analytics analysis, 141-145
 - actual monthly customer demand in motors, 143
 - Chicago customer demand (graph), 143
 - estimated shipping costs per motor, 141
 - Excel summary statistics of actual monthly customer demand in motors, 144
 - Houston customer demand (graph), 143
 - Kansas City customer demand (graph), 145
 - Little Rock customer demand (graph), 145
 - Oklahoma City customer demand (graph), 145
 - Omaha customer demand (graph), 145
 - SPSS summary statistics of actual monthly customer demand in motors, 144
 - predictive analytics analysis, 147-157
 - developing forecasting models, 147-154
 - resulting warehouse customer demand forecasts, 157
 - validating forecasting models, 155-157
 - prescriptive analysis, 158-163
 - demonstrating business performance improvement, 162-163
 - determining optimal shipping schedule, 159-161
 - selecting and developing optimization shipping model, 158-159
 - summary of BA procedure for manufacturer, 161-162
 - problem background and data, 139-140
- support, lack of, 50
- surplus variables, 219
- sustainability, achieving with business analytics, 21
- systematic random sampling, 73

T

- targets of change management, 59
- tasks as target of change management, 59
- teams, 51-53
 - collaboration, 51-53
 - participant roles, 52
 - reasons for team failures, 53
- technical specialists, 31
- technology as target of change management, 59
- testing
 - Durbin-Watson Autocorrelation Test, 284
 - statistical testing, 193-199
- text analytics, 23-25
- time series data
 - exponential smoothing
 - example of*, 285
 - simple model*, 284-285
 - smoothing averages*, 286-288
 - multiple regression models, 283-284
 - simple regression model
 - additive model*, 274
 - cyclical variation*, 275
 - multiplicative model*, 274
 - overview*, 272-274
 - random variation*, 275
 - seasonal variation*, 274
 - trend variation*, 274
- trend
 - simple regression model, 276-281
 - trend variation, 274
- trials, 177
- t-test: Paired Two Sample Means, 195

U

- unbounded solutions, 227-228
- uncertainty
 - decision-making under uncertainty, 311
 - Hurwicz criterion*, 312-313
 - laplace criterion*, 311-312

- maximax criterion*, 312
- maximin criterion*, 312
- minimax criterion*, 313-315

explained, 305

U.S. Census, 35

V

- validating forecasting models, 155-157
- value
 - EV (expected value) criterion, 308-309
 - EVPI (expected value of perfect information), 315
 - expected opportunity loss criterion, 309-311
 - failure to provide value, 53
 - inconsistent values, checking for, 57
- variables
 - slack variables, 218-219
 - surplus variables, 219
- variance, 68, 219
- variation in time series data, 272-274
- visualizing data
 - marketing/planning case study example
 - case study background*, 81
 - descriptive analytics analysis*, 82-90
 - statistical charts, 65-67

W

- warehouses (data), 38
- web logs, 34
- web mining, 39
- Wilcoxon Signed-Rank tests, 199

X-Y-Z

- Z values, 78-79
- zero-one programming (ZOP) model
 - explained, 264
 - problems/models, solving, 268-269