

# Business Statistics

7th  
Edition

## A First Course



Levine

Szabat

Stephan

# A ROADMAP FOR SELECTING A STATISTICAL METHOD

Data Analysis Task	For Numerical Variables	For Categorical Variables
<b>Describing a group or several groups</b>	Ordered array, stem-and-leaf display, frequency distribution, relative frequency distribution, percentage distribution, cumulative percentage distribution, histogram, polygon, cumulative percentage polygon <b>(Sections 2.2, 2.4)</b> Mean, median, mode, quartiles, range, interquartile range, standard deviation, variance, coefficient of variation, skewness, kurtosis, boxplot, normal probability plot <b>(Sections 3.1, 3.2, 3.3, 6.3)</b>	Summary table, bar chart, pie chart, Pareto chart <b>(Sections 2.1 and 2.3)</b>
<b>Inference about one group</b>	Confidence interval estimate of the mean <b>(Sections 8.1 and 8.2)</b> $t$ test for the mean <b>(Section 9.2)</b>	Confidence interval estimate of the proportion <b>(Section 8.3)</b> $Z$ test for the proportion <b>(Section 9.4)</b>
<b>Comparing two groups</b>	Tests for the difference in the means of two independent populations <b>(Section 10.1)</b> Paired $t$ test <b>(Section 10.2)</b> $F$ test for the difference between two variances <b>(Section 10.4)</b>	$Z$ test for the difference between two proportions <b>(Section 10.3)</b> Chi-square test for the difference between two proportions <b>(Section 11.1)</b>
<b>Comparing more than two groups</b>	One-way analysis of variance for comparing several means <b>(Section 10.5)</b>	Chi-square test for differences among more than two proportions <b>(Section 11.2)</b>
<b>Analyzing the relationship between two variables</b>	Scatter plot, time series plot <b>(Section 2.5)</b> Covariance, coefficient of correlation <b>(Section 3.5)</b> Simple linear regression <b>(Chapter 12)</b> $t$ test of correlation <b>(Section 12.7)</b>	Contingency table, side-by-side bar chart, PivotTables <b>(Sections 2.1, 2.3, 2.6)</b> Chi-square test of independence <b>(Section 11.3)</b>
<b>Analyzing the relationship between two or more variables</b>	Multiple regression <b>(Chapter 13)</b>	Multidimensional contingency tables <b>(Section 2.6)</b>

*This page intentionally left blank*

# Business Statistics

## A First Course

SEVENTH EDITION

### David M. Levine

Department of Statistics and Computer Information Systems  
Zicklin School of Business, Baruch College, City University of New York

### Kathryn A. Szabat

Department of Business Systems and Analytics  
School of Business, La Salle University

### David F. Stephan

Two Bridges Instructional Technology

**PEARSON**

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

**Editorial Director:** Chris Hoag  
**Editor in Chief:** Deirdre Lynch  
**Acquisitions Editor:** Suzanna Bainbridge  
**Editorial Assistant:** Justin Billing  
**Program Manager:** Chere Bemelmans  
**Project Manager:** Sherry Berg  
**Program Management Team Lead:** Marianne Stepanian  
**Project Management Team Lead:** Peter Silvia  
**Media Producer:** Jean Choe  
**TestGen Content Manager:** John Flanagan  
**MathXL Content Developer:** Bob Carroll

**Marketing Manager:** Erin Kelly  
**Marketing Assistant:** Emma Sarconi  
**Senior Author Support/Technology Specialist:** Joe Vetere  
**Rights and Permissions Project Manager:** Diahanne Lucas Dowridge  
**Senior Procurement Specialist:** Carol Melville  
**Associate Director of Design:** Andrea Nix  
**Program Design Lead and Cover Design:** Barbara Atkinson  
**Text Design, Production Coordination, Composition, and Illustrations:**  
Lumina Datamatics  
**Cover Image:** Pressmaster/Shutterstock

Copyright © 2016, 2013, 2010 by Pearson Education, Inc. All Rights Reserved. Printed in the United States of America. This publication is protected by copyright, and permission should be obtained from the publisher prior to any prohibited reproduction, storage in a retrieval system, or transmission in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise. For information regarding permissions, request forms and the appropriate contacts within the Pearson Education Global Rights & Permissions department, please visit [www.pearsoned.com/permissions/](http://www.pearsoned.com/permissions/).

PEARSON, ALWAYS LEARNING, is an exclusive trademark in the U.S. and/or other countries, owned by Pearson Education, Inc. or its affiliates.

Unless otherwise indicated herein, any third-party trademarks that may appear in this work are the property of their respective owners and any references to third-party trademarks, logos or other trade dress are for demonstrative or descriptive purposes only. Such references are not intended to imply any sponsorship, endorsement, authorization, or promotion of Pearson's products by the owners of such marks, or any relationship between the owner and Pearson Education, Inc. or its affiliates, authors, licensees or distributors.

MICROSOFT® AND WINDOWS® ARE REGISTERED TRADEMARKS OF THE MICROSOFT CORPORATION IN THE U.S.A. AND OTHER COUNTRIES. THIS BOOK IS NOT SPONSORED OR ENDORSED BY OR AFFILIATED WITH THE MICROSOFT CORPORATION. ILLUSTRATIONS OF MICROSOFT EXCEL IN THIS BOOK HAVE BEEN TAKEN FROM MICROSOFT EXCEL 2013, UNLESS OTHERWISE INDICATED.

MICROSOFT AND/OR ITS RESPECTIVE SUPPLIERS MAKE NO REPRESENTATIONS ABOUT THE SUITABILITY OF THE INFORMATION CONTAINED IN THE DOCUMENTS AND RELATED GRAPHICS PUBLISHED AS PART OF THE SERVICES FOR ANY PURPOSE. ALL SUCH DOCUMENTS AND RELATED GRAPHICS ARE PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND. MICROSOFT AND/OR ITS RESPECTIVE SUPPLIERS HEREBY DISCLAIM ALL WARRANTIES AND CONDITIONS WITH REGARD TO THIS INFORMATION, INCLUDING ALL WARRANTIES AND CONDITIONS OF MERCHANTABILITY, WHETHER EXPRESS, IMPLIED OR STATUTORY, FITNESS FOR A PARTICULAR PURPOSE, TITLE AND NON-INFRINGEMENT. IN NO EVENT SHALL MICROSOFT AND/OR ITS RESPECTIVE SUPPLIERS BE LIABLE FOR ANY SPECIAL, INDIRECT OR CONSEQUENTIAL DAMAGES OR ANY DAMAGES WHATSOEVER RESULTING FROM LOSS OF USE, DATA OR PROFITS, WHETHER IN AN ACTION OF CONTRACT, NEGLIGENCE OR OTHER TORTIOUS ACTION, ARISING OUT OF OR IN CONNECTION WITH THE USE OR PERFORMANCE OF INFORMATION AVAILABLE FROM THE SERVICES. THE DOCUMENTS AND RELATED GRAPHICS CONTAINED HEREIN COULD INCLUDE TECHNICAL INACCURACIES OR TYPOGRAPHICAL ERRORS. CHANGES ARE PERIODICALLY ADDED TO THE INFORMATION HEREIN. MICROSOFT AND/OR ITS RESPECTIVE SUPPLIERS MAY MAKE IMPROVEMENTS AND/OR CHANGES IN THE PRODUCT(S) AND/OR THE PROGRAM(S) DESCRIBED HEREIN AT ANY TIME. PARTIAL SCREEN SHOTS MAY BE VIEWED IN FULL WITHIN THE SOFTWARE VERSION SPECIFIED.

Minitab © 2013. Portions of information contained in this publication/book are printed with permission of Minitab Inc. All such material remains the exclusive property and copyright of Minitab Inc. All rights reserved.

#### Library of Congress Cataloging-in-Publication Data

Levine, David M., 1946—  
Business statistics : a first course / David M. Levine, Baruch College City University, Kathryn A. Szabat, La Salle University, David F. Stephan, Instructional Technology.—7th edition.  
pages cm  
ISBN 978-0-321-97901-8  
1. Commercial statistics. 2. Industrial management—Statistical methods. I. Szabat, Kathryn A. II. Stephan, David. III. Title.  
HF1017.B382 2016  
519.5—dc23

2014019092

1 2 3 4 5 6 7 8 9 10—CRK—18 17 16 15 14

**PEARSON**

[www.pearsonhighered.com](http://www.pearsonhighered.com)

ISBN-10: 0-321-97901-X  
ISBN-13: 978-0-321-97901-8

*To our spouses and children,  
Marilyn, Sharyn, Mary, and Mark*

---

*and to our parents, in loving memory,  
Lee, Reuben, Mary, William, Ruth, and Francis*

# About the Authors



*Kathryn Szabat, David Levine, and David Stephan*

David M. Levine, Kathryn A. Szabat, and David F. Stephan are all experienced business school educators committed to innovation and improving instruction in business statistics and related subjects.

**David Levine**, Professor Emeritus of Statistics and CIS at Baruch College, CUNY, is a nationally recognized innovator in statistics education for more than three decades. Levine has coauthored 14 books, including several business statistics textbooks; textbooks and professional titles that explain and explore quality management and the Six Sigma approach; and, with David Stephan, a trade paperback that explains statistical concepts to a general audience. Levine has presented or chaired numerous sessions about business education at leading conferences conducted by the Decision Sciences Institute (DSI) and the American Statistical Association, and he and his co-

authors have been active participants in the annual DSI Making Statistics More Effective in Schools and Business (MSMESB) mini-conference. During his many years teaching at Baruch College, Levine was recognized for his contributions to teaching and curriculum development with the College's highest distinguished teaching honor. He earned B.B.A. and M.B.A. degrees from CCNY, and a Ph.D. in industrial engineering and operations research from New York University.

As Associate Professor and Chair of Business Systems and Analytics at La Salle University, **Kathryn Szabat** has transformed several business school majors into one interdisciplinary major that better supports careers in new and emerging disciplines of data analysis including analytics. Szabat strives to inspire, stimulate, challenge, and motivate students through innovation and curricular enhancements, and shares her coauthors' commitment to teaching excellence and the continual improvement of statistics presentations. Beyond the classroom she has provided statistical advice to numerous business, nonbusiness, and academic communities, with particular interest in the areas of education, medicine, and nonprofit capacity building. Her research activities have led to journal publications, chapters in scholarly books, and conference presentations. Szabat is a member of the American Statistical Association (ASA), DSI, Institute for Operation Research and Management Sciences (INFORMS), and DSI MSMESB. She received a B.S. from SUNY-Albany, an M.S. in statistics from the Wharton School of the University of Pennsylvania, and a Ph.D. degree in statistics, with a cognate in operations research, from the Wharton School of the University of Pennsylvania.

Advances in computing have always shaped **David Stephan's** professional life. As an undergraduate, he helped professors use statistics software that was considered advanced even though it could compute *only* several things discussed in Chapter 3, thereby gaining an early appreciation for the benefits of using software to solve problems (and perhaps positively influencing his grades). An early advocate of using computers to support instruction, he developed a prototype of a mainframe-based system that anticipated features found today in Pearson's MathXL and served as special assistant for computing to the Dean and Provost at Baruch College. In his many years teaching at Baruch, Stephan implemented the first computer-based *classroom*, helped redevelop the CIS curriculum, and, as part of a FIPSE project team, designed and implemented a multimedia learning environment. He was also nominated for teaching honors. Stephan has presented at the SEDSI conference and the DSI MSMESB mini-conferences, sometimes with his coauthors. Stephan earned a B.A. from Franklin & Marshall College and an M.S. from Baruch College, CUNY, and he studied instructional technology at Teachers College, Columbia University.

For all three coauthors, continuous improvement is a natural outcome of their curiosity about the world. Their varied backgrounds and many years of teaching experience have come together to shape this book in ways discussed in the Preface. To learn more about the coauthors, visit **authors.davidlevinestatics.com**.

# Brief Contents

---

Preface xv

Getting Started: Important Things to Learn First 1

- 1 Defining and Collecting Data 10
- 2 Organizing and Visualizing Variables 31
- 3 Numerical Descriptive Measures 97
- 4 Basic Probability 142
- 5 Discrete Probability Distributions 176
- 6 The Normal Distribution 200
- 7 Sampling Distributions 226
- 8 Confidence Interval Estimation 248
- 9 Fundamentals of Hypothesis Testing: One-Sample Tests 284
- 10 Two-Sample Tests and One-Way ANOVA 323
- 11 Chi-Square Tests 387
- 12 Simple Linear Regression 414
- 13 Multiple Regression 464
- 14 Statistical Applications in Quality Management (*online*) 14-1

Appendices A–G 496

Self-Test Solutions and Answers to Selected Even-Numbered Problems 542

Index 567



# Contents

## Preface xv

## Getting Started: Important Things to Learn First 1

### USING STATISTICS: “You Cannot Escape from Data” 1

- GS.1 Statistics: A Way of Thinking 2
- GS.2 Data: What Is It? 2
  - Statistics 3
- GS.3 The Changing Face of Statistics 4
  - Business Analytics 4
  - “Big Data” 4
  - Integral Role of Software in Statistics 5
- GS.4 Statistics: An Important Part of Your Business 5
  - Education 5
  - Making Best Use of This Book 5
  - Making Best Use of the Software Guides 6

### REFERENCES 7

### KEY TERMS 7

### EXCEL GUIDE 8

- EG1. Getting Started with Microsoft Excel 8
- EG2. Entering Data 8

### MINITAB GUIDE 9

- MG.1 Getting Started with Minitab 9
- MG.2 Entering Data 9

## 1 Defining and Collecting Data 10

### USING STATISTICS: Beginning of the End ... Or the End of the Beginning? 10

- 1.1 Defining Variables 11
  - Classifying Variables by Type 11
- 1.2 Collecting Data 13
  - Data Sources 13
  - Populations and Samples 14
  - Structured Versus Unstructured Data 14
  - Electronic Formats and Encodings 15
  - Data Cleaning 15
  - Recoding Variables 15
- 1.3 Types of Sampling Methods 16
  - Simple Random Sample 17
  - Systematic Sample 18
  - Stratified Sample 18
  - Cluster Sample 18
- 1.4 Types of Survey Errors 19
  - Coverage Error 20

- Nonresponse Error 20
- Sampling Error 20
- Measurement Error 20
- Ethical Issues About Surveys 21

### THINK ABOUT THIS: New Media Surveys/Old Sampling Problems 21

### USING STATISTICS: Beginning of the End ... Revisited 22

### SUMMARY 23

### REFERENCES 23

### KEY TERMS 23

### CHECKING YOUR UNDERSTANDING 24

### CHAPTER REVIEW PROBLEMS 24

### CASES FOR CHAPTER 1 25

- Managing Ashland MultiComm Services 25
- CardioGood Fitness 25
- Clear Mountain State Student Surveys 26
- Learning with the Digital Cases 26

### CHAPTER 1 EXCEL GUIDE 28

- EG1.1 Defining Variables 28
- EG1.2 Collecting Data 28
- EG1.3 Types of Sampling Methods 28

### CHAPTER 1 MINITAB GUIDE 29

- MG1.1 Defining Variables 29
- MG1.2 Collecting Data 29
- MG1.3 Types of Sampling Methods 30

## 2 Organizing and Visualizing Variables 31

### USING STATISTICS: The Choice Is Yours 31

- 2.1 Organizing Categorical Variables 33
  - The Summary Table 33
  - The Contingency Table 33
- 2.2 Organizing Numerical Variables 37
  - The Ordered Array 37
  - The Frequency Distribution 38
  - Classes and Excel Bins 40
  - The Relative Frequency Distribution and the Percentage Distribution 40
  - The Cumulative Distribution 42
  - Stacked and Unstacked Data 44
- 2.3 Visualizing Categorical Variables 46
  - The Bar Chart 46
  - The Pie Chart 47
  - The Pareto Chart 48
  - The Side-by-Side Bar Chart 50
- 2.4 Visualizing Numerical Variables 52
  - The Stem-and-Leaf Display 52
  - The Histogram 54
  - The Percentage Polygon 55
  - The Cumulative Percentage Polygon (Ogive) 56

- 2.5 Visualizing Two Numerical Variables 60
  - The Scatter Plot 60
  - The Time-Series Plot 61
- 2.6 Organizing and Visualizing a Set of Variables 63
  - Multidimensional Contingency Tables 64
  - Data Discovery 65
- 2.7 The Challenge in Organizing and Visualizing Variables 67
  - Obscuring Data 67
  - Creating False Impressions 68
  - Chartjunk 68
  - Best Practices for Constructing Visualizations 70

### USING STATISTICS: The Choice *Is Yours*, Revisited 71

#### SUMMARY 72

#### REFERENCES 72

#### KEY EQUATIONS 73

#### KEY TERMS 73

#### CHECKING YOUR UNDERSTANDING 73

#### CHAPTER REVIEW PROBLEMS 74

### CASES FOR CHAPTER 2 78

- Managing Ashland MultiComm Services 78
- Digital Case 79
- CardioGood Fitness 79
- The Choice *Is Yours* Follow-Up 79
- Clear Mountain State Student Surveys 79

### CHAPTER 2 EXCEL GUIDE 80

- EG2.1 Organizing Categorical Variables 80
- EG2.2 Organizing Numerical Variables 82
- EG2.3 Visualizing Categorical Variables 84
- EG2.4 Visualizing Numerical Variables 86
- EG2.5 Visualizing Two Numerical Variables 89
- EG2.6 Organizing and Visualizing a Set of Variables 89

### CHAPTER 2 MINITAB GUIDE 91

- MG2.1 Organizing Categorical Variables 91
- MG2.2 Organizing Numerical Variables 91
- MG2.3 Visualizing Categorical Variables 92
- MG2.4 Visualizing Numerical Variables 93
- MG2.5 Visualizing Two Numerical Variables 95
- MG2.6 Organizing and Visualizing a Set of Variables 96

## 3 Numerical Descriptive Measures

97

### USING STATISTICS: More Descriptive Choices 97

- 3.1 Central Tendency 98
  - The Mean 98
  - The Median 100
  - The Mode 101
- 3.2 Variation and Shape 102
  - The Range 102
  - The Variance and the Standard Deviation 103
  - The Coefficient of Variation 107
  - Z Scores 108
  - Shape: Skewness 109
  - Shape: Kurtosis 110

- 3.3 Exploring Numerical Data 113
  - Quartiles 113
  - The Interquartile Range 115
  - The Five-Number Summary 116
  - The Boxplot 117
- 3.4 Numerical Descriptive Measures for a Population 120
  - The Population Mean 120
  - The Population Variance and Standard Deviation 121
  - The Empirical Rule 122
  - The Chebyshev Rule 123
- 3.5 The Covariance and the Coefficient of Correlation 124
  - The Covariance 125
  - The Coefficient of Correlation 126

### 3.6 Descriptive Statistics: Pitfalls and Ethical Issues 130

### USING STATISTICS: More Descriptive Choices, Revisited 130

#### SUMMARY 131

#### REFERENCES 131

#### KEY EQUATIONS 131

#### KEY TERMS 132

#### CHECKING YOUR UNDERSTANDING 132

#### CHAPTER REVIEW PROBLEMS 133

### CASES FOR CHAPTER 3 136

- Managing Ashland MultiComm Services 136
- Digital Case 136
- CardioGood Fitness 136
- More Descriptive Choices Follow-up 136
- Clear Mountain State Student Surveys 136

### CHAPTER 3 EXCEL GUIDE 137

- EG3.1 Central Tendency 137
- EG3.2 Variation and Shape 137
- EG3.3 Exploring Numerical Data 138
- EG3.4 Numerical Descriptive Measures for a Population 139
- EG3.5 The Covariance and the Coefficient of Correlation 139

### CHAPTER 3 MINITAB GUIDE 140

- MG3.1 Central Tendency 140
- MG3.2 Variation and Shape 140
- MG3.3 Exploring Numerical Data 140
- MG3.4 Numerical Descriptive Measures for a Population 141
- MG3.5 The Covariance and the Coefficient of Correlation 141

## 4 Basic Probability

142

### USING STATISTICS: Possibilities at M&R Electronics World 142

- 4.1 Basic Probability Concepts 143
  - Events and Sample Spaces 144
  - Contingency Tables and Venn Diagrams 146
  - Simple Probability 146
  - Joint Probability 147
  - Marginal Probability 148
  - General Addition Rule 149
- 4.2 Conditional Probability 152
  - Computing Conditional Probabilities 152
  - Decision Trees 154
  - Independence 156
  - Multiplication Rules 157
  - Marginal Probability Using the General Multiplication Rule 158

4.3 Bayes' Theorem 160

**THINK ABOUT THIS** Divine Providence and Spam 163

4.4 Counting Rules 165

4.5 Ethical Issues and Probability 168

**USING STATISTICS:** Possibilities at M&R Electronics World, Revisited 169

**SUMMARY** 169

**REFERENCES** 169

**KEY EQUATIONS** 170

**KEY TERMS** 170

**CHECKING YOUR UNDERSTANDING** 171

**CHAPTER REVIEW PROBLEMS** 171

**CASES FOR CHAPTER 4** 173

Digital Case 173

CardioGood Fitness 173

The Choice *Is* Yours Follow-Up 173

Clear Mountain State Student Surveys 173

**CHAPTER 4 EXCEL GUIDE** 174

EG4.1 Basic Probability Concepts 174

EG4.2 Conditional Probability 174

EG4.3 Bayes' Theorem 174

EG4.4 Counting Rules 174

**CHAPTER 4 MINITAB GUIDE** 175

MG4.1 Basic Probability Concepts 175

MG4.2 Conditional Probability 175

MG4.3 Bayes' Theorem 175

MG4.4 Counting Rules 175

## 5 Discrete Probability Distributions 176

**USING STATISTICS:** Events of Interest at Ricknel Home Centers 176

5.1 The Probability Distribution for a Discrete Variable 177

Expected Value of a Discrete Variable 177

Variance and Standard Deviation of a Discrete Variable 178

5.2 Binomial Distribution 181

5.3 Poisson Distribution 188

**USING STATISTICS:** Events of Interest at Ricknel Home Centers, Revisited 192

**SUMMARY** 192

**REFERENCES** 192

**KEY EQUATIONS** 192

**KEY TERMS** 193

**CHECKING YOUR UNDERSTANDING** 193

**CHAPTER REVIEW PROBLEMS** 193

**CASES FOR CHAPTER 5** 195

Managing Ashland MultiComm Services 195

Digital Case 196

**CHAPTER 5 EXCEL GUIDE** 197

EG5.1 The Probability Distribution for a Discrete Variable 197

EG5.2 Binomial Distribution 197

EG5.3 Poisson Distribution 197

**CHAPTER 5 MINITAB GUIDE** 198

MG5.1 The Probability Distribution for a Discrete Variable 198

MG5.2 Binomial Distribution 198

MG5.3 Poisson Distribution 198

## 6 The Normal Distribution 200

**USING STATISTICS:** Normal Downloading at MyTVLab 200

6.1 Continuous Probability Distributions 201

6.2 The Normal Distribution 201

Computing Normal Probabilities 203

Finding *X* Values 208

**VISUAL EXPLORATIONS:** Exploring the Normal Distribution 212

**THINK ABOUT THIS:** What Is Normal? 212

6.3 Evaluating Normality 214

Comparing Data Characteristics to Theoretical Properties 214

Constructing the Normal Probability Plot 216

**USING STATISTICS:** Normal Downloading at MyTVLab, Revisited 218

**SUMMARY** 219

**REFERENCES** 219

**KEY EQUATIONS** 219

**KEY TERMS** 219

**CHECKING YOUR UNDERSTANDING** 220

**CHAPTER REVIEW PROBLEMS** 220

**CASES FOR CHAPTER 6** 221

Managing Ashland MultiComm Services 221

Digital Case 222

CardioGood Fitness 222

More Descriptive Choices Follow-up 222

Clear Mountain State Student Surveys 222

**CHAPTER 6 EXCEL GUIDE** 223

EG6.1 Continuous Probability Distributions 223

EG6.2 The Normal Distribution 223

EG6.3 Evaluating Normality 223

**CHAPTER 6 MINITAB GUIDE** 224

MG6.1 Continuous Probability Distributions 224

MG6.2 The Normal Distribution 224

MG6.3 Evaluating Normality 224

## 7 Sampling Distributions 226

**USING STATISTICS:** Sampling Oxford Cereals 226

7.1 Sampling Distributions 227

7.2 Sampling Distribution of the Mean 227

The Unbiased Property of the Sample Mean 227

Standard Error of the Mean 229

Sampling from Normally Distributed Populations 230

Sampling from Non-normally Distributed Populations—The Central Limit Theorem 233

**VISUAL EXPLORATIONS:** Exploring Sampling Distributions 237

7.3 Sampling Distribution of the Proportion 238

**USING STATISTICS:** Sampling Oxford Cereals, Revisited 242  
**SUMMARY** 242  
**REFERENCES** 242  
**KEY EQUATIONS** 242  
**KEY TERMS** 243  
**CHECKING YOUR UNDERSTANDING** 243  
**CHAPTER REVIEW PROBLEMS** 243

### CASES FOR CHAPTER 7 245

Managing Ashland MultiComm Services 245  
 Digital Case 245

### CHAPTER 7 EXCEL GUIDE 246

EG7.1 Sampling Distributions 246  
 EG7.2 Sampling Distribution of the Mean 246  
 EG7.3 Sampling Distribution of the Proportion 246

### CHAPTER 7 MINITAB GUIDE 247

MG7.1 Sampling Distributions 247  
 MG7.2 Sampling Distribution of the Mean 247  
 MG7.3 Sampling Distribution of the Proportion 247

## 8 Confidence Interval Estimation 248

**USING STATISTICS:** Getting Estimates at Ricknel Home Centers 248

- 8.1 Confidence Interval Estimate for the Mean ( $\sigma$  Known) 249  
 Can You Ever Know the Population Standard Deviation? 254
- 8.2 Confidence Interval Estimate for the Mean ( $\sigma$  Unknown) 255  
 Student's  $t$  Distribution 255  
 Properties of the  $t$  Distribution 256  
 The Concept of Degrees of Freedom 257  
 The Confidence Interval Statement 258
- 8.3 Confidence Interval Estimate for the Proportion 263
- 8.4 Determining Sample Size 266  
 Sample Size Determination for the Mean 266  
 Sample Size Determination for the Proportion 268
- 8.5 Confidence Interval Estimation and Ethical Issues 271
- 8.6 Bootstrapping (*online*) 272

**USING STATISTICS:** Getting Estimates at Ricknel Home Centers, Revisited 272

**SUMMARY** 272

**REFERENCES** 273

**KEY EQUATIONS** 273

**KEY TERMS** 273

**CHECKING YOUR UNDERSTANDING** 273

**CHAPTER REVIEW PROBLEMS** 274

### CASES FOR CHAPTER 8 277

Managing Ashland MultiComm Services 277  
 Digital Case 278  
 Sure Value Convenience Stores 278  
 CardioGood Fitness 279  
 More Descriptive Choices Follow-Up 279  
 Clear Mountain State Student Surveys 279

### CHAPTER 8 EXCEL GUIDE 280

EG8.1 Confidence Interval Estimate for the Mean ( $\sigma$  Known) 280  
 EG8.2 Confidence Interval Estimate for the Mean ( $\sigma$  Unknown) 280

EG8.3 Confidence Interval Estimate for the Proportion 281  
 EG8.4 Determining Sample Size 281

### CHAPTER 8 MINITAB GUIDE 282

MG8.1 Confidence Interval Estimate for the Mean ( $\sigma$  Known) 282  
 MG8.2 Confidence Interval Estimate for the Mean ( $\sigma$  Unknown) 282  
 MG8.3 Confidence Interval Estimate for the Proportion 282  
 MG8.4 Determining Sample Size 283

## 9 Fundamentals of Hypothesis Testing: One-Sample Tests 284

**USING STATISTICS:** Significant Testing at Oxford Cereals 284

- 9.1 Fundamentals of Hypothesis-Testing Methodology 285  
 The Null and Alternative Hypotheses 285  
 The Critical Value of the Test Statistic 286  
 Regions of Rejection and Nonrejection 287  
 Risks in Decision Making Using Hypothesis Testing 287  
 Z Test for the Mean ( $\sigma$  Known) 290  
 Hypothesis Testing Using the Critical Value Approach 290  
 Hypothesis Testing Using the  $p$ -Value Approach 293  
 A Connection Between Confidence Interval Estimation and Hypothesis Testing 295  
 Can You Ever Know the Population Standard Deviation? 296
- 9.2  $t$  Test of Hypothesis for the Mean ( $\sigma$  Unknown) 297  
 The Critical Value Approach 298  
 The  $p$ -Value Approach 300  
 Checking the Normality Assumption 300
- 9.3 One-Tail Tests 304  
 The Critical Value Approach 304  
 The  $p$ -Value Approach 305
- 9.4 Z Test of Hypothesis for the Proportion 308  
 The Critical Value Approach 309  
 The  $p$ -Value Approach 310
- 9.5 Potential Hypothesis-Testing Pitfalls and Ethical Issues 312  
 Statistical Significance Versus Practical Significance 312  
 Statistical *Insignificance* Versus Importance 313  
 Reporting of Findings 313  
 Ethical Issues 313

**USING STATISTICS:** Significant Testing at Oxford Cereals, Revisited 314

**SUMMARY** 314

**REFERENCES** 314

**KEY EQUATIONS** 315

**KEY TERMS** 315

**CHECKING YOUR UNDERSTANDING** 315

**CHAPTER REVIEW PROBLEMS** 315

### CASES FOR CHAPTER 9 317

Managing Ashland MultiComm Services 317  
 Digital Case 318  
 Sure Value Convenience Stores 318

### CHAPTER 9 EXCEL GUIDE 319

EG9.1 Fundamentals of Hypothesis-Testing Methodology 319  
 EG9.2  $t$  Test of Hypothesis for the Mean ( $\sigma$  Unknown) 319

- EG9.3 One-Tail Tests 320
- EG9.4 Z Test of Hypothesis for the Proportion 320

**CHAPTER 9 MINITAB GUIDE 321**

- MG9.1 Fundamentals of Hypothesis-Testing Methodology 321
- MG9.2 *t* Test of Hypothesis for the Mean ( $\sigma$  Unknown) 321
- MG9.3 One-Tail Tests 321
- MG9.4 Z Test of Hypothesis for the Proportion 322

## 10 Two-Sample Tests and One-Way ANOVA 323

**USING STATISTICS:** For North Fork, Are There Different Means to the Ends? 323

- 10.1 Comparing the Means of Two Independent Populations 324
  - Pooled-Variance *t* Test for the Difference Between Two Means 324
  - Confidence Interval Estimate for the Difference Between Two Means 329
  - t* Test for the Difference Between Two Means, Assuming Unequal Variances 330
  - Do People Really Do This? 330
- 10.2 Comparing the Means of Two Related Populations 333
  - Paired *t* Test 334
  - Confidence Interval Estimate for the Mean Difference 339
- 10.3 Comparing the Proportions of Two Independent Populations 341
  - Z Test for the Difference Between Two Proportions 341
  - Confidence Interval Estimate for the Difference Between Two Proportions 345
- 10.4 *F* Test for the Ratio of Two Variances 347
- 10.5 One-Way ANOVA 352
  - F* Test for Differences Among More Than Two Means 355
  - One-Way ANOVA *F* Test Assumptions 359
  - Levene Test for Homogeneity of Variance 360
  - Multiple Comparisons: The Tukey-Kramer Procedure 361
- 10.6 Effect Size (*online*) 366

**USING STATISTICS:** For North Fork, Are There Different Means to the Ends? Revisited 367

**SUMMARY 367**

**REFERENCES 368**

**KEY EQUATIONS 368**

**KEY TERMS 369**

**CHECKING YOUR UNDERSTANDING 370**

**CHAPTER REVIEW PROBLEMS 370**

**CASES FOR CHAPTER 10 372**

- Managing Ashland MultiComm Services 372
- Digital Case 373
- Sure Value Convenience Stores 373
- CardioGood Fitness 374
- More Descriptive Choices Follow-Up 374
- Clear Mountain State Student Surveys 374

**CHAPTER 10 EXCEL GUIDE 376**

- EG10.1 Comparing the Means of Two Independent Populations 376
- EG10.2 Comparing the Means of Two Related Populations 378
- EG10.3 Comparing the Proportions of Two Independent Populations 379

- EG10.4 *F* Test for the Ratio of Two Variances 379
- EG10.5 One-Way ANOVA 380

**CHAPTER 10 MINITAB GUIDE 383**

- MG10.1 Comparing the Means of Two Independent Populations 383
- MG10.2 Comparing the Means of Two Related Populations 383
- MG10.3 Comparing the Proportions of Two Independent Populations 384
- MG10.4 *F* Test for the Ratio of Two Variances 384
- MG10.5 One-Way ANOVA 385

## 11 Chi-Square Tests 387

**USING STATISTICS:** Avoiding Guesswork About Resort Guests 387

- 11.1 Chi-Square Test for the Difference Between Two Proportions 388
- 11.2 Chi-Square Test for Differences Among More Than Two Proportions 395
- 11.3 Chi-Square Test of Independence 400

**USING STATISTICS:** Avoiding Guesswork About Resort Guests, Revisited 405

**SUMMARY 406**

**REFERENCES 406**

**KEY EQUATIONS 407**

**KEY TERMS 407**

**CHECKING YOUR UNDERSTANDING 407**

**CHAPTER REVIEW PROBLEMS 407**

**CASES FOR CHAPTER 11 409**

- Managing Ashland MultiComm Services 409
- Digital Case 410
- CardioGood Fitness 410
- Clear Mountain State Student Surveys 411

**CHAPTER 11 EXCEL GUIDE 412**

- EG11.1 Chi-Square Test for the Difference Between Two Proportions 412
- EG11.2 Chi-Square Test for Differences Among More Than Two Proportions 412
- EG11.3 Chi-Square Test of Independence 412

**CHAPTER 11 MINITAB GUIDE 413**

- MG11.1 Chi-Square Test for the Difference Between Two Proportions 413
- MG11.2 Chi-Square Test for Differences Among More Than Two Proportions 413
- MG11.3 Chi-Square Test of Independence 413

## 12 Simple Linear Regression 414

**USING STATISTICS:** Knowing Customers at Sunflowers Apparel 414

- 12.1 Types of Regression Models 415
  - Simple Linear Regression Models 416
- 12.2 Determining the Simple Linear Regression Equation 417
  - The Least-Squares Method 417
  - Predictions in Regression Analysis: Interpolation Versus Extrapolation 420

Computing the  $Y$  Intercept,  $b_0$ , and the Slope,  $b_1$  420

**VISUAL EXPLORATIONS: Exploring Simple Linear Regression Coefficients** 423

12.3 Measures of Variation 425

Computing the Sum of Squares 425  
The Coefficient of Determination 426  
Standard Error of the Estimate 428

12.4 Assumptions of Regression 430

12.5 Residual Analysis 430

Evaluating the Assumptions 430

12.6 Measuring Autocorrelation: The Durbin-Watson Statistic 434

Residual Plots to Detect Autocorrelation 434  
The Durbin-Watson Statistic 435

12.7 Inferences About the Slope and Correlation Coefficient 438

$t$  Test for the Slope 438  
 $F$  Test for the Slope 440  
Confidence Interval Estimate for the Slope 441  
 $t$  Test for the Correlation Coefficient 442

12.8 Estimation of Mean Values and Prediction of Individual Values 445

The Confidence Interval Estimate for the Mean Response 445  
The Prediction Interval for an Individual Response 446

12.9 Potential Pitfalls in Regression 449

Six Steps for Avoiding the Potential Pitfalls 451

**USING STATISTICS: Knowing Customers at Sunflowers Apparel, Revisited** 451

**SUMMARY** 451

**REFERENCES** 452

**KEY EQUATIONS** 453

**KEY TERMS** 454

**CHECKING YOUR UNDERSTANDING** 454

**CHAPTER REVIEW PROBLEMS** 454

**CASES FOR CHAPTER 12** 458

Managing Ashland MultiComm Services 458  
Digital Case 458  
Bryne Packaging 458

**CHAPTER 12 EXCEL GUIDE** 460

EG12.1 Types of Regression Models 460  
EG12.2 Determining the Simple Linear Regression Equation 460  
EG12.3 Measures of Variation 461  
EG12.4 Assumptions of Regression 461  
EG12.5 Residual Analysis 461  
EG12.6 Measuring Autocorrelation: The Durbin-Watson Statistic 462  
EG12.7 Inferences About the Slope and Correlation Coefficient 462  
EG12.8 Estimation of Mean Values and Prediction of Individual Values 462

**CHAPTER 12 MINITAB GUIDE** 462

MG12.1 Types of Regression Models 462  
MG12.2 Determining the Simple Linear Regression Equation 462  
MG12.3 Measures of Variation 463  
MG12.4 Assumptions 463  
MG12.5 Residual Analysis 463  
MG12.6 Measuring Autocorrelation: The Durbin-Watson Statistic 463  
MG12.7 Inferences About the Slope and Correlation Coefficient 463  
MG12.8 Estimation of Mean Values and Prediction of Individual Values 463

## 13 Multiple Regression 464

**USING STATISTICS: The Multiple Effects of OmniPower Bars** 464

13.1 Developing a Multiple Regression Model 465

Interpreting the Regression Coefficients 466  
Predicting the Dependent Variable  $Y$  468

13.2  $r^2$ , Adjusted  $r^2$ , and the Overall  $F$  Test 470

Coefficient of Multiple Determination 470  
Adjusted  $r^2$  471  
Test for the Significance of the Overall Multiple Regression Model 472

13.3 Residual Analysis for the Multiple Regression Model 474

13.4 Inferences Concerning the Population Regression Coefficients 475

Tests of Hypothesis 475  
Confidence Interval Estimation 477

13.5 Using Dummy Variables and Interaction Terms in Regression Models 479

Dummy Variables 479  
Interactions 481

**USING STATISTICS: The Multiple Effects of Omnipower Bars, Revisited** 485

**SUMMARY** 485

**REFERENCES** 485

**KEY EQUATIONS** 487

**KEY TERMS** 487

**CHECKING YOUR UNDERSTANDING** 487

**CHAPTER REVIEW PROBLEMS** 487

**CASES FOR CHAPTER 13** 490

Managing Ashland MultiComm Services 490  
Digital Case 490

**CHAPTER 13 EXCEL GUIDE** 491

EG13.1 Developing a Multiple Regression Model 491  
EG13.2  $r^2$ , Adjusted  $r^2$ , and the Overall  $F$  Test 492  
EG13.3 Residual Analysis for the Multiple Regression Model 492  
EG13.4 Inferences Concerning the Population Regression Coefficients 493  
EG13.5 Using Dummy Variables and Interaction Terms in Regression Models 493

**CHAPTER 13 MINITAB GUIDE** 493

MG13.1 Developing a Multiple Regression Model 493  
MG13.2  $r^2$ , Adjusted  $r^2$ , and the Overall  $F$  Test 494  
MG13.3 Residual Analysis for the Multiple Regression Model 494  
MG13.4 Inferences Concerning the Population Regression Coefficients 494  
MG13.5 Using Dummy Variables and Interaction Terms in Regression Models 495

## 14 Statistical Applications in Quality Management (online) 14-1

**USING STATISTICS: Finding Quality at the Beachcomber** 14-1

14.1 The Theory of Control Charts 14-2

14.2 Control Chart for the Proportion: The  $p$  Chart 14-4

- 14.3 The Red Bead Experiment: Understanding Process Variability 14-10
- 14.4 Control Chart for an Area of Opportunity: The  $c$  Chart 14-12
- 14.5 Control Charts for the Range and the Mean 14-15
  - The  $R$  Chart 14-16
  - The  $\bar{X}$  Chart 14-18
- 14.6 Process Capability 14-21
  - Customer Satisfaction and Specification Limits 14-21
  - Capability Indices 14-23
  - $CPL$ ,  $CPU$ , and  $C_{pk}$  14-24
- 14.7 Total Quality Management 14-26
- 14.8 Six Sigma 14-28
  - The DMAIC Model 14-29
  - Roles in a Six Sigma Organization 14-30
  - Lean Six Sigma 14-30

**USING STATISTICS: Finding Quality at the Beachcomber, Revisited 14-31**

**SUMMARY 14-31**

**REFERENCES 14-32**

**KEY EQUATIONS 14-32**

**KEY TERMS 14-33**

**CHAPTER REVIEW PROBLEMS 14-34**

- The Harnswell Sewing Machine Company Case 14-36
- Managing Ashland Multicomm Services 14-38

**CHAPTER 14 EXCEL GUIDE 14-39**

- EG14.1 The Theory of Control Charts 14-39
- EG14.2 Control Chart for the Proportion: The  $p$  Chart 14-39
- EG14.3 The Red Bead Experiment: Understanding Process Variability 14-40
- EG14.4 Control Chart for an Area of Opportunity: The  $c$  Chart 14-40
- EG14.5 Control Charts for the Range and the Mean 14-41
- EG14.6 Process Capability 14-42

**CHAPTER 14 MINITAB GUIDE 14-42**

- MG14.1 The Theory of Control Charts 14-42
- MG14.2 Control Chart for the Proportion: The  $p$  Chart 14-42
- MG14.3 The Red Bead Experiment: Understanding Process Variability 14-42
- MG14.4 Control Chart for an Area of Opportunity: The  $c$  Chart 14-42
- MG14.5 Control Charts for the Range and the Mean 14-43
- MG14.6 Process Capability 14-44

**Appendices 496**

- A. Basic Math Concepts and Symbols 497
  - A.1 Rules for Arithmetic Operations 497
  - A.2 Rules for Algebra: Exponents and Square Roots 497
  - A.3 Rules for Logarithms 498
  - A.4 Summation Notation 499
  - A.5 Statistical Symbols 502
  - A.6 Greek Alphabet 502
- B. Important Excel and Minitab Skills 503
  - B.1 Basic Excel Operations 503
  - B.2 Formulas and Cell References 503
  - B.3 Entering Formulas into Worksheets 504
  - B.4 Pasting with Paste Special 505
  - B.5 Basic Worksheet Cell Formatting 505

- B.6 Chart Formatting 507
- B.7 Selecting Cell Ranges for Charts 508
- B.8 Deleting the “Extra” Histogram Bar 508
- B.9 Creating Histograms for Discrete Probability Distributions 508
- B.10 Basic Minitab Operations 509
- C. Online Resources 510
  - C.1 About the Online Resources for This Book 510
  - C.2 Accessing the Online Resources 510
  - C.3 Details of Downloadable Files 510
  - C.4 PHStat 515
- D. Configuring Microsoft Excel 516
  - D.1 Getting Microsoft Excel Ready for Use (ALL) 516
  - D.2 Getting PHStat Ready for Use (ALL) 517
  - D.3 Configuring Excel Security for Add-In Usage (WIN) 517
  - D.4 Opening PHStat (ALL) 518
  - D.5 Using a Visual Explorations Add-in Workbook (ALL) 518
  - D.6 Checking for the Presence of the Analysis ToolPak (ALL) 518
- E. Tables 519
  - E.1 Table of Random Numbers 519
  - E.2 The Cumulative Standardized Normal Distribution 521
  - E.3 Critical Values of  $t$  523
  - E.4 Critical Values of  $\chi^2$  525
  - E.5 Critical Values of  $F$  526
  - E.6 Critical Values of the Studentized Range,  $Q$  530
  - E.7 Critical Values,  $d_L$  and  $d_U$ , of the Durbin–Watson Statistic,  $D$  (Critical Values Are One-Sided) 532
  - E.8 Control Chart Factors 533
  - E.9 The Standardized Normal Distribution 534
- F. Useful Excel Knowledge 535
  - F.1 Useful Keyboard Shortcuts 535
  - F.2 Verifying Formulas and Worksheets 535
  - F.3 New Function Names 536
  - F.4 Understanding the Nonstatistical Functions 537
- G. Software FAQs 539
  - G.1 PHStat FAQs 539
  - G.2 Microsoft Excel FAQs 540
  - G.3 FAQs for New Users of Microsoft Excel 2013 540
  - G.4 Minitab FAQs 541

**Self-Test Solutions and Answers to Selected Even-Numbered Problems 542**

**Index 567**

# Preface

---

The world of business statistics has grown larger, expanding into and combining with other disciplines. And, in a reprise of something that occurred a generation ago, new fields of study, this time with names such as informatics, data analytics, and decision science, have emerged.

This time of change makes what is taught in business statistics and how it is taught all the more critical. We, the coauthors, think about these changes as we seek ways to continuously improve the teaching of business statistics. We actively participate in Decision Sciences Institute (DSI), American Statistical Association (ASA), and Making Statistics More Effective in Schools and Business (MSMESB) conferences. We use the ASA's Guidelines for Assessment and Instruction (GAISE) reports and combine them with our experiences teaching business statistics to a diverse student body at several universities. We also benefit from the interests and efforts of our past coauthors, Mark Berenson and Timothy Krehbiel.

## Our Educational Philosophy

When writing for introductory business statistics students, we are guided by these principles:

**Help students see the relevance of statistics to their own careers by providing examples drawn from the functional areas in which they may be specializing.** Students need to learn statistics in the context of the functional areas of business. We present each statistics topic in the context of areas such as accounting, finance, management, and marketing and explain the application of specific methods to business activities.

**Emphasize interpretation and analysis of statistical results over calculation.** We emphasize the interpretation of results, the evaluation of the assumptions, and the discussion of what should be done if the assumptions are violated. We believe that these activities are more important and will serve students better in the future than focusing on tedious hand calculations.

**Give students ample practice in understanding how to apply statistics to business.** We believe that both classroom examples and homework exercises should involve actual or realistic data, using small and large sets of data, to the extent possible.

**Familiarize students with the use of spreadsheet and statistical software.** We integrate spreadsheet and statistical software into all statistics topics to illustrate how this software assists business decision making. (Using software in this way also supports our second point about emphasizing interpretation over calculation).

**Provide clear instructions to students for using spreadsheet and statistical software.** We believe that providing such instructions facilitates learning and helps prevent minimizing the chance that learning software to the level necessary will distract from the learning of statistical concepts.

## What's New and Innovative in This Edition?

This seventh edition of *Business Statistics: A First Course* contains these new and innovative features.

**Getting Started: Important Things to Learn First** Created to help students get a jumpstart on the course, lessen any fear about learning statistics, and provide coverage of those things that would be helpful to know even before the first class of the term. "Getting Started" has been developed to be posted online or otherwise distributed before the first class section begins and is available for download as explained in Appendix C. Instructors teaching online or hybrid course sections may find this to be a particularly valuable tool to help organize the students in their section.

**Student Tips** In-margin notes that reinforce hard-to-master concepts and provide quick study tips for mastering important details.

**Discussion of Business Analytics** "Getting Started: Important Things to Learn First" quickly defines *business analytics* and *big data* and notes how these things are changing the face of statistics.

**PHStat version 4** For Microsoft Excel users, this successor to the PHStat2 statistics add-in contains several new and enhanced procedures, is simpler to set up and run, and is compatible with both Microsoft Windows and (Mac) OS X Excel versions.



**Additional Chapter Short Takes** Online PDF documents (available for download as explained in Appendix C) that supply additional insights or explanations to important statistical concepts or details about the results presented in this book.

## Revised and Enhanced Content

This seventh edition of *Business Statistics: A First Course* contains the following revised and enhanced content.

**New Continuing End-of-Chapter Cases** This edition features several new end-of-chapter cases. New and recurring throughout the book is a case that concerns the analysis of sales and marketing data for home fitness equipment (CardioGood Fitness), a case that concerns pricing decisions made by a retailer (Sure Value Convenience Stores), and the More Descriptive Choices Follow-Up case, which extends the use of the retirement funds sample first introduced in Chapter 2. Also recurring is the Clear Mountain State Student Surveys case, which uses data collected from surveys of undergraduate and graduate students to practice and reinforce statistical methods learned in various chapters. This case replaces end-of-chapter questions related to the student survey database in the previous edition. In addition, there is a new case in simple linear regression (Brynne Packaging).

**Many New Applied Examples and Problems** Many of the applied examples throughout this book use new problems or revised data. Approximately 43% of the problems are new to this edition. The end-of-section and end-of-chapter problem sets contain many new problems that use data from *The Wall Street Journal*, *USA Today*, and other sources.

**Revised Using Statistics Scenarios** Five chapters have new or revised Using Statistics scenarios.

**Revised Making Best Use of This Book section** Included as part of Section GS.4 of “Getting Started: Important Things to Learn First,” this section presents an overview of this book and checklist that helps students prepare for using Microsoft Excel or Minitab with this book.

**Revised Software Appendices** These appendices review the foundational skills for using Microsoft Excel and Minitab, review the latest technical information, and, for Excel users, cover optional but useful skills for working with Excel.

## Distinctive Features

This seventh edition of *Business Statistics: A First Course* continues the use of the following distinctive features.

**Using Statistics Business Scenarios** Each chapter begins with a Using Statistics example that shows how statistics is used in the functional areas of business—accounting, finance, information systems, management, and marketing. Each scenario is used throughout the chapter to provide an applied context for the concepts. The chapter concludes with a Using Statistics, Revisited section that reinforces the statistical methods and applications discussed in each chapter.

**Emphasis on Data Analysis and Interpretation of Excel and Minitab Results** Our focus emphasizes analyzing data by interpreting results while reducing emphasis on doing calculations. For example, in the coverage of tables and charts in Chapter 2, we help students interpret various charts and explain when to use each chart discussed. Our coverage of hypothesis testing in Chapters 9 through 11 and regression and multiple regression in Chapters 12 and 13 include extensive software results so that the  $p$ -value approach can be emphasized.

**Pedagogical Aids** We use an active writing style, boxed numbered equations, set-off examples that reinforce learning concepts, student tips, problems divided into “Learning the Basics” and “Applying the Concepts,” key equations, and key terms.

**Digital Cases** In the Digital Cases, available for download as explained in Appendix C, learners must examine interactive PDF documents to sift through various claims and information to discover the data most relevant to a business case scenario. Learners then determine whether the conclusions and claims are supported by the data. In doing so, learners discover and learn how to identify common misuses of statistical information. (Instructional tips for using the Digital Cases and solutions to the Digital Cases are included in the Instructor’s Solutions Manual.)

**Answers** Most answers to the even-numbered exercises are included at the end of the book.

**Flexibility Using Excel** For almost every statistical method discussed, students can use *In-Depth Excel* instructions to directly work with worksheet solution details *or* they can use either the *PHStat* instructions *or* the *Analysis ToolPak* instructions to automate the creation of those worksheet solutions.

**PHStat** PHStat is the Pearson Education statistics add-in that includes more than 60 procedures that create Excel worksheets and charts. Unlike other add-ins, PHStat results are real worksheets that contain real Excel calculations (called formulas in Excel). You can examine the contents of worksheet solutions to learn the appropriate functions and calculations necessary to apply a particular statistical method. With most of these worksheet solutions, you can change worksheet data and immediately see how those changes affect the results.

Descriptive Statistics: boxplot, descriptive summary, dot scale diagram, frequency distribution, histogram and polygons, Pareto diagram, scatter plot, stem-and-leaf display, one-way tables and charts, and two-way tables and charts

Probability and probability distributions: simple and joint probabilities, normal probability plot, and binomial, and Poisson probability distributions

Sampling: sampling distributions simulation

Confidence interval estimation: for the mean, sigma unknown; for the mean, sigma known; and for the proportion

Sample size determination: for the mean and the proportion

One-sample tests:  $Z$  test for the mean, sigma known;  $t$  test for the mean, sigma unknown; and  $Z$  test for the proportion

Two-sample tests (unsummarized data): pooled-variance  $t$  test, separate-variance  $t$  test, paired  $t$  test, and  $F$  test for differences in two variances

Two-sample tests (summarized data): pooled-variance  $t$  test, separate-variance  $t$  test, paired  $t$  test,  $Z$  test for the differences in two means,  $F$  test for differences in two variances, chi-square test for differences in two proportions, and  $Z$  test for the difference in two proportions

Multiple-sample tests: chi-square test, Levene test, one-way ANOVA, and Tukey-Kramer procedure

Regression: simple linear regression, and multiple regression

Data preparation: stack and unstack data

Control charts:  $p$  chart,  $c$  chart, and  $R$  and  $Xbar$  charts.

To learn more about PHStat, see Appendix C.

**Visual Explorations** The series of Excel workbooks that allow students to interactively explore important statistical concepts in the normal distribution, sampling distributions, and regression analysis. For the normal distribution, students see the effect of changes in the mean and standard deviation on the areas under the normal curve. For sampling distributions, students use simulation to explore the effect of sample size on a sampling distribution. For regression analysis, students fit a line of regression and observe how changes in the slope and intercept affect the goodness of fit. To learn more about Visual Explorations, see Appendix C.

## Chapter-by-Chapter Changes Made for This Edition

Besides the new and innovative content described in “What’s New and Innovative in This Edition?” the seventh edition of *Business Statistics: A First Course* contains the following specific changes to each chapter.

**Getting Started: Important Things to Learn First** This all-new chapter includes new material on business analytics and introduces the DCOVA framework and a basic vocabulary of statistics, both of which were introduced in Chapter 1 of the sixth edition.

- Chapter 1** Collecting data has been relocated to this chapter from Section 2.1. Sampling methods and types of survey errors have been relocated from Sections 7.1 and 7.2. There is a new subsection on data cleaning. The CardioGood Fitness and Clear Mountain State Surveys cases are included.
- Chapter 2** Section 2.1, “Data Collection,” has been moved to Chapter 1. The chapter uses a new data set that contains a sample of 316 mutual funds and a new set of restaurant cost data. The CardioGood Fitness, The Choice *Is* Yours Follow-up, and Clear Mountain State Surveys cases are included.
- Chapter 3** For many examples, this chapter uses the new mutual funds data set that is introduced in Chapter 2. There is increased coverage of skewness and kurtosis. There is a new example on computing descriptive measures from a population using “Dogs of the Dow.” The CardioGood Fitness, More Descriptive Choices Follow-up, and Clear Mountain State Surveys cases are included.
- Chapter 4** The chapter example has been updated. There are new problems throughout the chapter. The CardioGood Fitness, The Choice *Is* Yours Follow-up, and Clear Mountain State Surveys cases are included.
- Chapter 5** There are many new problems throughout the chapter. The notation used has been made more consistent.
- Chapter 6** This chapter has an updated Using Statistics scenario and some new problems. The CardioGood Fitness, More Descriptive Choices Follow-up, and Clear Mountain State Surveys cases are included.
- Chapter 7** Sections 7.1 and 7.2 have been moved to Chapter 1. An additional example of sampling distributions from a larger population has been included.
- Chapter 8** This chapter includes an updated Using Statistics scenario and new examples and exercises throughout the chapter. The Sure Value Convenience Stores, CardioGood Fitness, More Descriptive Choices Follow-up, and Clear Mountain State Surveys cases are included. There is an online section on bootstrapping.
- Chapter 9** This chapter includes additional coverage of the pitfalls of hypothesis testing. The Sure Value Convenience Stores case is included.
- Chapter 10** This chapter has an updated Using Statistics scenario, a new example on the paired *t*-test on textbook prices, a new example on the *Z*-test for the difference between two proportions, and a new one-way ANOVA example on mobile electronics sales at a general merchandiser. The Sure Value Convenience Stores, CardioGood Fitness, More Descriptive Choices Follow-up, and Clear Mountain State Surveys cases are included. There is a new online section on Effect Size.
- Chapter 11** The chapter includes many new problems. This chapter includes the Sure Value Convenience Stores, CardioGood Fitness, More Descriptive Choices Follow-up, and Clear Mountain State Surveys cases.
- Chapter 12** The Using Statistics scenario has been updated and changed, with new data used throughout the chapter. This chapter includes the Brynne Packaging case.
- Chapter 13** The chapter includes many new and revised problems.
- Chapter 14** The “Statistical Applications in Quality Management” chapter has been renumbered as Chapter 14 and is available for download as explained in Appendix C.

## Student and Instructor Resources

**Student Solutions Manual**, by Professor Pin Tian Ng of Northern Arizona University and accuracy checked by Annie Puciloski, provides detailed solutions to virtually all the even-numbered exercises and worked-out solutions to the self-test problems (ISBN-10: 0-321-99881-2; ISBN-13: 978-0-321-99881-1).

**Online resources** The complete set of online resources are discussed fully in Appendix C

For adopting instructors, the following resources are among those available at the Instructor's Resource Center, located at [www.pearsonhighered.com/irc](http://www.pearsonhighered.com/irc).

**Instructor's Solutions Manual**, by Professor Pin Tian Ng of Northern Arizona University and accuracy checked by Annie Puciloski, includes solutions for end-of-section and end-of-chapter problems, answers to case questions where applicable, and teaching tips for each chapter.

**Lecture PowerPoint Presentations**, by Professor Patrick Schur of Miami University and accuracy checked by David Levine and Kathryn Szabat, are available for each chapter. The PowerPoint slides provide an instructor with individual lecture outlines to accompany the text. The slides include many of the figures and tables from the text. Instructors can use these lecture notes as is or can easily modify the notes to reflect specific presentation needs.

**Test Bank**, by Professor Pin Tian Ng of Northern Arizona University, contains true/false, multiple-choice, fill-in, and problem-solving questions based on the definitions, concepts, and ideas developed in each chapter of the text.

**TestGen®** ([www.pearsoned.com/testgen](http://www.pearsoned.com/testgen)) enables instructors to build, edit, print, and administer tests using a computerized bank of questions developed to cover all the objectives of the text. TestGen is algorithmically based, allowing instructors to create multiple but equivalent versions of the same question or test with the click of a button. Instructors can also modify test bank questions or add new questions. The software and test bank are available for download from Pearson Education's online catalog.

## MathXL®

**MathXL® for Statistics Online Course (access code required)** MathXL® is the homework and assessment engine that runs MyStatLab. (MyStatLab is MathXL plus a learning management system.)

With *MathXL for Statistics*, instructors can:

- Create, edit, and assign online homework and tests using algorithmically generated exercises correlated at the objective level to the textbook.
- Create and assign their own online exercises and import TestGen tests for added flexibility.
- Maintain records of all student work, tracked in MathXL's online gradebook.

With *MathXL for Statistics*, students can:

- Take chapter tests in MathXL and receive personalized study plans and/or personalized homework assignments based on their test results.
- Use the study plan and/or the homework to link directly to tutorial exercises for the objectives they need to study.
- Access supplemental animations directly from selected exercises.
- Knowing that students often use external statistical software, we make it easy to copy our data sets, both from the eText and the MyStatLab questions, into StatCrunch™, Microsoft Excel, Minitab, and a variety of other software packages.

MathXL for Statistics is available to qualified adopters. For more information, visit [www.mathxl.com](http://www.mathxl.com) or contact your Pearson representative.

## MyStatLab™

**MyStatLab™ Online Course (access code required)** MyStatLab from Pearson is the world's leading online resource for teaching and learning statistics; integrating interactive homework, assessment, and media in a flexible, easy-to-use format. MyStatLab is a course management system that delivers **proven results** in helping individual students succeed.

- MyStatLab can be implemented successfully in any environment—lab-based, hybrid, fully online, traditional—and demonstrates the quantifiable difference that integrated usage has on student retention, subsequent success, and overall achievement.
- MyStatLab's comprehensive online gradebook automatically tracks students' results on tests, quizzes, homework, and in the study plan. Instructors can use the gradebook to provide positive feedback or intervene if students have trouble. Gradebook data can be easily exported to a variety of spreadsheet programs, such as Microsoft Excel.

MyStatLab provides **engaging experiences** that personalize, stimulate, and measure learning for each student. In addition to the resources below, each course includes a full interactive online version of the accompanying textbook.

- **Tutorial Exercises with Multimedia Learning Aids:** The homework and practice exercises in MyStatLab align with the exercises in the textbook, and most regenerate algorithmically to give students unlimited opportunity for practice and mastery. Exercises offer immediate helpful feedback, guided solutions, sample problems, animations, videos, statistical software tutorial videos and eText clips for extra help at point-of-use.
- **MyStatLab Accessibility:** MyStatLab is compatible with the JAWS screen reader, and enables multiple-choice and free-response problem-types to be read, and interacted with via keyboard controls and math notation input. MyStatLab also works with screen enlargers, including ZoomText, MAGic, and SuperNova. And all MyStatLab videos accompanying texts with copyright 2009 and later have closed captioning. More information on this functionality is available at <http://mymathlab.com/accessibility>.
- **StatTalk Videos:** Fun-loving statistician Andrew Vickers takes to the streets of Brooklyn, NY, to demonstrate important statistical concepts through interesting stories and real-life events. This series of 24 fun and engaging videos will help students actually understand statistical concepts. Available with an instructor's user guide and assessment questions.
- **Business Insight Videos:** 10 engaging videos show managers at top companies using statistics in their everyday work. Assignable questions encourage discussion.
- **Additional Question Libraries:** In addition to algorithmically regenerated questions that are aligned with your textbook, MyStatLab courses come with two additional question libraries:
  - **450 exercises in Getting Ready for Statistics** cover the developmental math topics students need for the course. These can be assigned as a prerequisite to other assignments, if desired.
  - **1000 exercises in the Conceptual Question Library** require students to apply their statistical understanding.
- **StatCrunch™:** MyStatLab integrates the web-based statistical software, StatCrunch, within the online assessment platform so that students can easily analyze data sets from exercises and the text. In addition, MyStatLab includes access to **www.StatCrunch.com**, a vibrant online community where users can access tens of thousands of shared data sets, create and conduct online surveys, perform complex analyses using the powerful statistical software, and generate compelling reports.
- **Statistical Software Support and Integration:** We make it easy to copy our data sets, both from the eText and the MyStatLab questions, into software such as StatCrunch, Minitab, Excel, and more. Students have access to a variety of support tools—Technology Tutorial Videos, Technology Study Cards, and Technology Manuals for select titles—to learn how to effectively use statistical software.

And, MyStatLab comes from an **experienced partner** with educational expertise and an eye on the future.

- Knowing that you are using a Pearson product means knowing that you are using quality content. That means that our eTexts are accurate and our assessment tools work. It means we are committed to making MyMathLab as accessible as possible.
- Whether you are just getting started with MyStatLab, or have a question along the way, we're here to help you learn about our technologies and how to incorporate them into your course.

To learn more about how MyStatLab combines proven learning applications with powerful assessment, visit **www.mystatlab.com** or contact your Pearson representative.

**StatCrunch™** StatCrunch is powerful web-based statistical software that allows users to perform complex analyses, share data sets, and generate compelling reports of their data. The vibrant online community offers tens of thousands of shared data sets for students to analyze.

Full access to StatCrunch is available with a MyStatLab kit, and StatCrunch is available by itself to qualified adopters. StatCrunch Mobile now available; just visit **www.statcrunch.com/mobile** from the browser on your smart phone or tablet. For more information, visit our website at **www.statcrunch.com**, or contact your Pearson representative.

We thank the RAND Corporation and the American Society for Testing and Materials for their kind permission to publish various tables in Appendix E, and to the American Statistical Association for its permission to publish diagrams from the *American Statistician*.

## A Note of Thanks

Creating a new edition of a textbook is a team effort, and we would like to thank our Pearson Education editorial, marketing, and production teammates: Suzanna Bainbridge, Chere Bemelmans, Sherry Berg, Erin Kelly, Deirdre Lynch, Christine Stavrou, Jean Choe, Marianne Stepanian, and Joe Vetere. We also thank our statistical reader and accuracy checker Annie Puciloski for her diligence in checking our work and Nancy Kincade of Lumina Datamatics. Finally, we would like to thank our families for their patience, understanding, love, and assistance in making this book a reality.

## Contact Us!

We invite you to email us at **authors@[davidlevinestatistics.com](mailto:authors@ davidlevinestatistics.com)** if you have a question or require clarification about the contents this book or if you have a suggestion for a future edition of this book. Please include “BSAFC7” in the subject line of your message. While we have strived to make this book as error-free as possible, we encourage you to also email us if you discover an error or have concern about the content in this book.

You can also visit us at **[davidlevinestatistics.com](http://davidlevinestatistics.com)**, where you will find additional information about us, this book, and our other textbooks and publications by the coauthors.

*David M. Levine, Kathryn A. Szabat, and David F. Stephan*

*This page intentionally left blank*

## GETTING STARTED

# Important Things to Learn First

### CONTENTS

- GS.1 Statistics: A Way of Thinking
- GS.2 Data: What Is It?
- GS.3 The Changing Face of Statistics
  - Business Analytics
  - “Big Data”
  - Integral Role of Software in Statistics
- GS.4 Statistics: An Important Part of Your Business Education
  - Making Best Use of This Book
  - Making Best Use of the Software Guides

### EXCEL GUIDE

- EG.1 Getting Started with Microsoft Excel
- EG.2 Entering Data

### MINITAB GUIDE

- MG.1 Getting Started with Minitab
- MG.2 Entering Data

### OBJECTIVES

- That the preponderance of data makes learning statistics critically important
- Statistics is a way of thinking that can lead to better decisions
- How applying the DCOVA framework for statistics can help solve business problems
- The significance of business analytics.
- The opportunity business analytics represent for business students
- How to prepare for using Microsoft Excel or Minitab with this book

### USING STATISTICS

## “You Cannot Escape from Data”

You hear the word *data* almost every day and may know that data are facts about the world. You might think about data as numbers, such as the poll results that show that 45% of the people polled believe the economy will improve during the next year. But data are more than just numerical facts. For example, every time you visit an online search engine, send or receive an email or text message, or post something to a social media site, you are creating and using data.

In this larger sense of *data*, you accept *as almost true* the premises of stories in which characters collect “lots of data” to uncover conspiracies, foretell disasters, or catch criminals. You might hear concerns about how a governmental agency might be collecting data to “spy” on you. You might even have heard how some businesses “mine” their data for profit. You may have realized that, in today’s world, **you cannot escape from data.**

Although you cannot escape from data, you might choose to avoid data. If you avoid data, you must blindly accept other people’s data summaries and that can expose you to fraud. (Recall financial scams that claimed great rewards that were totally fictitious.) If you avoid data, you must solely rely on “gut feelings” when making decisions—much less effective than using the rational processes you study in business courses. When you realize that avoiding data is not an option, you realize that knowing how to work with data effectively is an important skill. In identifying that skill, you have discovered that **you cannot escape learning statistics, the methods that allow you to work with data effectively.**



Angela Waye/Shutterstock



## GS.1 Statistics: A Way of Thinking

Statistics are the methods that allow you to work with data effectively. These methods represent a way of thinking that can help you make better decisions. If you ever created a chart to summarize data or calculated values such as averages to summarize data, you have used statistics. But there's even more to statistics than these commonly taught techniques, as a quick review of the detailed table of contents shows.

The statistics that you have learned at a lower grade level most likely required you to perform mathematical calculations. In contrast, businesses today rely on software to perform those calculations faster and more accurately than you could do by hand. In any case, computation by software forms only part of one task of many when applying statistics. To best understand that statistics is a way of thinking, you need a framework that organizes the set of tasks that form statistics. One such framework is the **DCOVA framework**.

### THE DCOVA FRAMEWORK

The tasks of DCOVA framework are:

- **Define** the data that you want to study to solve a problem or meet an objective.
- **Collect** the data from appropriate sources.
- **Organize** the data collected by developing tables.
- **Visualize** the data collected by developing charts.
- **Analyze** the data collected to reach conclusions and present those results.

The tasks **Define**, **Collect**, **Organize**, **Visualize**, and **Analyze** help you to apply statistics to business decision making. You must always do the first two tasks first to have meaningful results, but, in practice, the order of the other three can vary and sometimes are done concurrently. For example, certain ways of visualizing data help you to organize your data while performing preliminary analysis as well.

Using the DCOVA framework helps you to apply statistical methods to these four broad categories of business activities:

- Summarize and visualize business data
- Reach conclusions from those data
- Make reliable predictions about business activities
- Improve business processes

Throughout this book, and especially in the Using Statistics scenarios that begin the chapters, you will discover specific examples of how DCOVA helps you apply statistics. For example, in one chapter, you will learn how to demonstrate whether a marketing campaign has increased sales of a product, while in another you will learn how a television station can reduce unnecessary labor expenses.

## GS.2 Data: What Is It?

Defining data as just “facts about the world,” to quote the opening essay, can prove confusing as such facts could be singular, a value associated with something, or collective, a list of values associated with something. For example, “David Levine” is a singular fact, a coauthor of this book, whereas “David, Kathy, and David” is the *collective* list of authors of this book. Furthermore, if everything is data, how do you distinguish “David Levine” from “Business Statistics: A First Course,” two very different facts (coauthor and title) about this book. Statisticians avoid this confusion by using a more specific definition of data and by defining a second word, *variable*.

In statistics, **data** are “the values associated with a trait or property that help distinguish the occurrences of something.” For example, the names “David Levine” and “Kathryn Szabat” are data because they are both values that help distinguish one of the authors of this book from another. In this book, *data* is always plural to remind you that data are a collection, or set, of values. While one could say that a single value, such as “David Levine,” is a *datum*, the phrases *data point*, *observation*, *response*, and *single data value* are more typically encountered.

A trait or property of something with which values (data) are associated is called a **variable**. For example, you might define the variables “coauthor” and “title” if you were defining data about a set of textbooks.

Substituting the word characteristic for the phrase “trait or property” and using the phrase “an item or individual” instead of the vague “something” produces the definitions of *variable* and *data* that this book uses.

### Student Tip

Business convention places the data, or *set* of values, for a variable in a worksheet column. Because of this convention, people sometimes use the word *column* as a substitute for *variable*.

#### VARIABLE

A characteristic of an item or individual.

#### DATA

The set of individual values associated with a variable.

Think about characteristics that distinguish individuals in a human population. Name, height, weight, eye color, marital status, adjusted gross income, and place of residence are all characteristics of an individual. All of these traits are possible *variables* that describe people.

Defining a variable called author-name to be the first and last names of the authors of this text makes it clear that valid values would be “David Levine,” “Kathryn Szabat,” and “David Stephan” and not “Levine,” “Szabat,” and “Stephan.” Be careful of cultural or other assumptions in definitions—for example, is “last name” a family name, as is common usage in North America, or an individual’s own unique name, as is common usage in many Asian countries?

## Statistics

Having defined *data*, you can define the subject of this book, **statistics** as the methods that help transform data into useful information for decision makers. Statistics allows you to determine whether your data represent information that could be used in making better decisions. Therefore, statistics helps you determine whether differences in the numbers are meaningful in a significant way or are due to chance. To illustrate, consider the following news reports about various data findings:

- **“Acceptable Online Ad Length Before Seeing Free Content”** (*USA Today*, February 16, 2012, p. 1B) A survey of 1,179 adults 18 and over reported that 54% thought that 15 seconds was an acceptable online ad length before seeing free content.
- **“6 New Facts About Facebook.”** (Pew Research Center, [bit.ly/lkENZcA](http://bit.ly/lkENZcA), February 3, 2014) A survey reported that women were more likely than men to cite seeing photos or videos, sharing with many people at once, seeing entertaining or funny posts, learning about ways to help others, and receiving support from people in your network as reasons to use Facebook.
- **“Follow the Tweets”** (H. Rui, A. Whinston, and E. Winkler, *The Wall Street Journal*, November 30, 2009, p. R4) In this study, the authors found that the number of times a specific product was mentioned in comments in the Twitter social messaging service could be used to make accurate predictions of sales trends for that product.

Without statistics, you cannot determine whether the “numbers” in these stories represent useful information. Without statistics, you cannot validate claims such as the claim that the number of tweets can be used to predict the sales of certain products. And without statistics, you cannot see patterns that large amounts of data sometimes reveal.

When talking about statistics, you use the term **descriptive statistics** to refer to methods that primarily help summarize and present data. Counting physical objects in a kindergarten class may have been the first time you used a *descriptive* method. You use the term **inferential statistics** to refer to methods that use data collected from a small group to reach conclusions about a larger group. If you had formal statistics instruction in a lower grade, you were probably mostly taught descriptive methods, the focus of the early chapters of this book, and you may be unfamiliar with many of the inferential methods discussed in later chapters.

## GS.3 The Changing Face of Statistics

The data from which the Using Statistics scenario notes you cannot “escape” has encouraged the increasing use of statistical methods that either did not exist, were not practical to do, or were not widely known in the past. These methods and changes in information and communications technologies that you may have studied in another course have helped to extend the application of statistics in business and make statistical knowledge more critical to business success. This is the changing face of statistics.

### Business Analytics

Of all the recent changes that have made statistics more prominent and more important, the set of methods collectively known as **business analytics** best reflects this changing face of statistics. Business analytics combine traditional statistical methods with methods from management science and information systems to form an interdisciplinary tool that supports fact-based management decision making. Business analytics enables you to:

- Use statistical methods to analyze and explore data to uncover unforeseen relationships.
- Use management science methods to develop optimization models that support all levels of management, from strategic planning to daily operations.
- Use information systems methods to collect and process data sets of all sizes, including very large data sets that would otherwise be hard to examine efficiently.

Even if you have never heard of the term business analytics, you may be familiar with the application of these methods. Headlines about governmental agencies mining personal data to combat crime or terrorism, stories about how companies learn your secrets, including the example memorably summarized as “How Target Knows You’re Pregnant” (a bit of an overstatement), or even discussions about how social media or streaming media companies recommend choices to their users or sell advertisements to display to particular users, all reflect this changing face of statistics.

### “Big Data”

The data from which you cannot “escape” has taken new forms in recent years, including the form known as *big data*. **Big data** are the collections of data that cannot be easily browsed or analyzed using traditional methods.

Big data lacks a more precise operational definition, but using the term implies data that are being collected in huge volumes and at very fast rates (typically in near real-time) as well as data that takes a variety of forms other than the traditional structured forms such as data processing records, files, and tables. These attributes of “volume, velocity, and variety” (see reference 4) help distinguish “big data” from a set of data that happens to be “large” but that can be placed into a file that contains repeating records or rows that share the same arrangement or structure.

Big data presents opportunities to gain new management insights or extract value from the data resources of a business (see reference 7). Businesses gain these new insights or value through statistics, especially through the application of the newer methods of business analytics.

## Integral Role of Software in Statistics

Section GS.1 notes that businesses rely on software to perform statistical calculations faster and more accurately than you could do by hand. Consistent to this observation, this book emphasizes the interpretation of statistical results generated by software over the hand calculation of those results. The book uses both Microsoft Excel and Minitab to generate those results and show in a larger way how software is integral to applying statistical methods to business decision making.

Both Excel and Minitab use *worksheets* to store data for analysis. **Worksheets** are tabular arrangements of data, in which the intersections of rows and columns form **cells**, boxes into which you make entries. In Minitab and Excel, you use columns of cells to enter the data for variables, using one column for each variable. Typically to use a statistical method in either program, you select one or more columns of data (one or more variables) and then apply the appropriate program function. This means the examples and problems found in this book use traditional structured data and not collections of data that could be considered big data. Not to worry, learning with structured data will allow you to master statistical principles that you can apply later when using big data.

### Student Tip

The names of Excel and Minitab files that contain the data for examples and problems appear in this distinctive type face **Retirement Funds** throughout this book.

## GS.4 Statistics: An Important Part of Your Business Education

The changing face of statistics means that statistics has become a very important part of your business education. In the current data-driven environment of business, you need general analytical skills that allow you to manipulate data, interpret analytical results, and incorporate results in a variety of decision-making applications, such as accounting, finance, HR management, marketing, strategy and planning, and supply chain management.

The decisions you make will be increasingly based on data and not on gut or intuition supported by personal experience. Data-guided practice is proving to be successful; studies have shown an increase in productivity, innovation, and competition for organizations that embrace business analytics. The use of data and data analysis to drive business decisions cannot be ignored. Having a well-balanced mix of technical skills—such as statistics, modeling, and basic information technology skills—and managerial skills—such as business acumen, problem-solving skills, and communication skills—will best prepare you for today's, and tomorrow's, workplace (see reference 1).

Business students once considered statistics to be merely a required course that contained content unrelated to their own majors. If you opened this book and had similar thoughts, you were overlooking the changing face of statistics. Use this book to better understand the implications of this change as you learn to use the DCOVA framework to apply statistical methods to the four categories of business activities listed in Section GS.1.

### Student Tip

Don't worry if your course does not cover every section of every chapter. Introductory business statistics courses vary in terms of scope, length, and number of college credits earned. Your functional area of study or major may also affect what you learn.

## Making Best Use of This Book

This book uses the DCOVA framework to organize and present its statistical content. To make best use of this book, first make sure you understand the DCOVA framework (see page 2). With that knowledge, you can group the chapters of this book as follows:

- Chapter 1: the **Define** and **Collect** tasks, the mandatory starting tasks for applying a statistical method.
- Chapters 2 and 3: the **Organize** and **Visualize** tasks that help summarize and visualize business data (the first activity listed on page 2).
- Chapter 3 (again) and Chapters 4 through 11: the **Analyze** task methods that use sample data to help reach conclusions about populations (the second activity listed on page 2).
- Chapters 12 and 13: the **Analyze** task methods that help make reliable predictions (the third activity).
- Online Chapter 14: the **Analyze** task methods that help you improve business processes (the fourth activity).

To get the most from every chapter, first read the opening Using Statistics scenarios. Each chapter's scenario always describes a business situation in which the methods about to be discussed in the chapter could be used to help resolve issues or problems that the scenario described. Scenarios are the source of many of the in-chapter examples used to discuss statistical methods. At the end of each chapter, a "revisited" section reviews how the chapter's statistical methods would help solve the issues and problems raised initially in the opening scenario.

Each chapter fully integrates Microsoft Excel and Minitab illustrations with its examples, reflecting the integral role that software plays in applying statistical methods to business decision making. Each chapter concludes with software guides (discussed separately below) that contain how-to instructions for using Excel or Minitab for the statistical topic the chapter discusses.

Each chapter also ends with a summary and a list of key equations and key terms that help you review what you have learned. "Checking Your Understanding" questions test your understanding of basic concepts and "Chapter Review Problems" allow you to practice what you have learned.

As you read through a chapter you will find pointers to supplemental material available online, end-of-section questions and problems as well as these recurring features:

- **Student Tips** that help clarify and reinforce significant details about particular statistical concepts (such as the tip that occurs on this page).
- **Visual Explorations** that allow you to interactively explore statistical concepts in Microsoft Excel.
- **"Think About This" essays** that further explore statistical concepts.

You can enhance your analytic and communication skills by making best use of the many case studies found in this book. The continuing case *Managing Ashland MultiComm Services* appears in most chapters and asks you to use your analytic skills to help solve problems managers of a residential telecommunications provider face. Cases unique to a chapter or a subset of chapters provide report-writing practice and additional problem-solving opportunities. The unique "Digital Cases" additionally challenge you to use statistical principles to sort through claims found in various documents to uncover which claims are well supported and which ones are dubious, at best.

## Making Best Use of the Software Guides

To make best use of software guides, read the getting started information that appears later in this chapter for the program you will be using and complete the Table GS.1 checklist. The software how-to guides presume you already have awareness of basic computing concepts and skills such as mouse operations and interacting with windows and dialog boxes. Software guides use the following conventions in their instructions:

- Things to type and where to type them appear in boldface. (Enter **450** in **cell B5**.)
- Names of special keys are capitalized and in boldface. (Press **Enter**.)
- Targets of click or select operations appear in boldface. (Click **OK**. Select the **first 2-D Bar** gallery item.)
- When instructions require you to press more than one key at the same time, all keys are shown capitalized and in boldface and are joined together with the "+" symbol. (Press **Ctrl+C**. Press **Ctrl+Shift+Enter**. Press **Command+Enter**.)
- Consecutive menu or ribbon selections are shown capitalized, mixed case, and in boldface, joined together with the → symbol. (Select **File → New**. Select **Stat → Tables → Tally Individual Variables**.)
- Specific names of Excel and Minitab functions, worksheets, or files are shown capitalized, mixed case, and in boldface. (Open to the **DATA worksheet** of the **Retirement Funds workbook**.)
- Placeholder objects that express the general case of an instruction appear in italics and in boldface. Example: Use **AVERAGE(*cell range of variable*)** to compute the mean of a numerical variable.



### Student Tip

If you need to review these skills, read **Basic Computing Skills**, a PDF file that you can download using instructions found in Appendix C.

Starting with Chapter 1, the section numbers of the software guides reflect their in-chapter counterparts. For example, guide sections EG1.1 and MG1.1 contain the Excel and Minitab instructions for Section 1.1 “Defining Variables.”

### TABLE GS.1

Checklist for Using Microsoft Excel or Minitab with This Book

- Read Appendix C to learn about the online resources you need to make best use of this book.
- Download the online resources that you will need to use this book, using the instructions in Appendix C.
- Check for and apply updates to the software that you plan to use. (See the Appendix Section D.1 instructions).
- If you plan to use PHStat, the Visual Explorations add-in workbooks, or the Analysis ToolPak with Microsoft Windows Excel, read the special instructions in Appendix D.
- Read Appendix G to learn answers to frequently asked questions (FAQs).

## REFERENCES

1. Advani, D. “Preparing Students for the Jobs of the Future.” *University Business* (2011), [bit.ly/1gNLTJm](http://bit.ly/1gNLTJm).
2. Davenport, T., and J. Harris. *Competing on Analytics: The New Science of Winning*. Boston: Harvard Business School Press, 2007.
3. Davenport, T., J. Harris, and R. Morison. *Analytics at Work*. Boston: Harvard Business School Press, 2010.
4. Laney, D. *3D Data Management: Controlling Data Volume, Velocity, and Variety*. Stamford, CT: META Group. February 6, 2001.
5. Levine, D., and D. Stephan. “Teaching Introductory Business Statistics Using the DCOVA Framework.” *Decision Sciences Journal of Innovative Education* 9 (Sept. 2011): 393–398.
6. Liberatore, M., and W. Luo. “The Analytics Movement.” *Interfaces* 40 (2010): 313–324.
7. “What Is Big Data?” IBM Corporation, [www.ibm.com/big-data/us/en/](http://www.ibm.com/big-data/us/en/).

## KEY TERMS

big data	4	descriptive statistics	4	templates	8
cells	5	inferential statistics	4	variable	3
data	3	project	9	workbook	8
business analytics	4	statistical package	9	worksheets	5
DCOVA framework	2	statistics	3		

## EXCEL GUIDE

## EG.1 GETTING STARTED with MICROSOFT EXCEL

Microsoft Excel evolved from earlier applications that automated the preparation of accounting and financial worksheets. In Excel, worksheet cells can be individually formatted and contain either data values or programming-like statements called *formulas* (discussed fully in Appendix B). To make best use of Excel, businesses use worksheet solutions called **templates** that already contain formatted entries. Decision makers open such templates and make minor modifications, sometimes as simple as entering values into specific cells, to generate useful information.

Templates can be a single worksheet, but often are a set of worksheets that are stored in a **workbook**. This book uses a series of templates that collectively are called the Excel Guide workbooks. These workbooks typically contain one worksheet dedicated to computing and displaying the result, the worksheets that are pictured throughout this book, and at least one worksheet that stores the data being used by the results worksheet.

In this book, you can work with the templates in one of two ways. You can open the Excel Guide workbooks and make manual changes to its worksheets, similar to how an employee would open and use a business template. You can also use PHStat, the Pearson Education statistics add-in for Excel (discussed in Appendix C) that automates the retrieval and modification of these templates. Unless otherwise noted, using either method will result in results worksheets like the ones pictured in this book. If you choose to make manual changes to the Excel Guide workbooks, you will need to know how to edit formulas, alter worksheets, and correct charts, operations that are discussed in Appendix B. If you choose to use PHStat, you will need only the basic Microsoft Office skills of knowing how to enter data (discussed below), open and save files, print worksheets, and perform copy-and-paste and insert operations that are summarized at the start of Appendix B.

Occasionally, you will also find instructions for using the Data Analysis ToolPak, an add-in that comes with Microsoft Excel. Note that some templates have been designed to mimic the appearance of the worksheets created by ToolPak add-in procedures so that ToolPak users will see the same or similar results as template users. (However visually similar they may appear, ToolPak worksheets are formatted printed reports that are not templates and therefore cannot be reused with other data.)

The Excel Guide instructions work best with the current versions of Microsoft Windows Excel and (Mac) OS X Excel, including Excel 2011, Excel 2013, and Office 365 Excel. Versions occasionally vary and this book provides alternate instructions keyed to the version that varies when necessary. Starting with Excel 2010, Microsoft renamed, and in many cases revised, many of the statistical functions that formulas in this book's templates use. When a template uses one of these newer functions, an alternate template that uses the older function names will also be found in the workbook.

## EG.2 ENTERING DATA

To enter data into a specific cell, move the cell pointer to that cell by using the cursor keys, moving the mouse pointer, or completing the proper touch operation. As you type an entry, the entry appears in the formula bar area that is located over the top of the worksheet. You complete your entry by pressing **Tab** or **Enter** or by clicking the checkmark button in the formula bar.

All "Excel data files" and most Excel Guide workbooks contain a **DATA worksheet** similar to the example shown below. Consistent to the rules first stated in Section GS.3, DATA worksheets use columns of cells to enter the data for variables, using one column for each variable and use the cell in the first row of a column to enter the name of variable for that column.

Use the DATA worksheets as models for worksheets you prepare to store the data for your variables. As you create your own "data" worksheets, never skip a row when entering data for a variable and try to avoid using numbers as row 1 variable headings. (If you must use a number for a heading, precede the number with an apostrophe.) Also, pay attention to special instructions in this book that discuss the entry order and arrangement of the columns for your variables. For some statistical methods, entering variables in a column order that Excel does not expect will lead to incorrect results.

**Student Tip**  
All "Excel files" are *workbook* files, even those that contain a single worksheet, such as the Excel data files discussed in Appendix C. When instructions in this book use a workbook that contains two or more worksheets, the instructions identify the name of the worksheet that is the object of the instruction.

When a results worksheet or other template uses at least one of the newer functions, the template workbook includes a worksheet with the prefix *OLDER* that users of Excel 2007 should use. (PHStat automatically switches to the older names if you are using Excel 2007.)

	A	B	C	D	E	F	G	H	I	J	K	L
1	Fund Number	Market Cap	Type	Assets	Turnover Ratio	Beta	SD	Risk	1YrReturn%	3YrReturn%	5YrReturn%	10YrReturn%
2	RF001	Large	Growth	309.90	12.21	1.15	18.72	Low	28.99	24.26	11.06	8.97
3	RF002	Large	Growth	23.30	0.00	2.19	35.72	High	33.40	22.72	-4.89	0.02
4	RF003	Large	Growth	141.50	147.00	2.24	36.69	High	33.98	21.91	1.53	12.55
5	RF004	Large	Growth	118.50	5.00	2.24	36.63	High	33.78	21.89	1.57	12.69
6	RF005	Large	Growth	575.30	121.00	0.89	14.56	Low	21.62	16.47	9.40	10.30

# MINITAB GUIDE

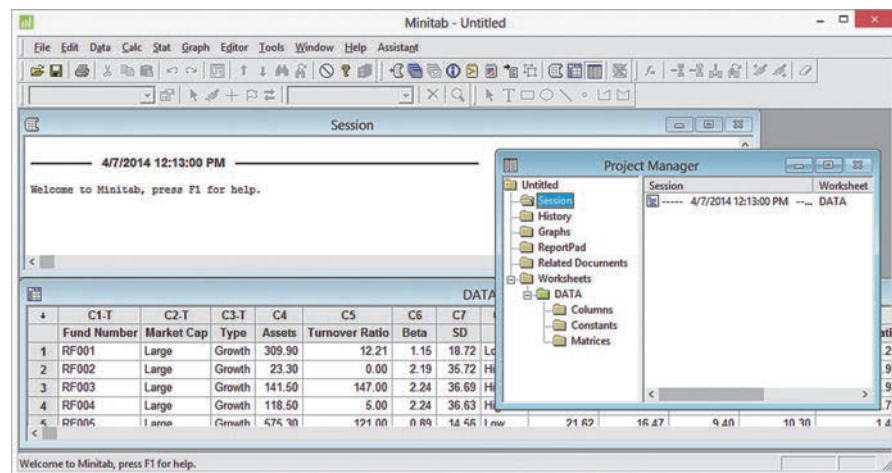
## MG.1 GETTING STARTED WITH MINITAB

Minitab is a **statistical package**, software developed specifically to perform a wide range of statistical analyses as accurately as possible. In Minitab, you enter data into a window that contains a worksheet, then select commands, and then see the results in other windows. The collection of all windows forms a **project** and you can save entire projects as **.mpj** project files or choose to save individual worksheets in **.mtw** worksheet files.

When you first open Minitab, you typically see a new project that contains a window with a blank worksheet and the Session window that records all commands you select and displays results. Pictured below is a project after a worksheet named DATA has been opened. Besides the slightly obscured DATA worksheet window and **Session** window, this figure also shows a **Project Manager** that summarizes the content of the current project. Note that all three windows appear *inside* the main Minitab window.

### Student Tip

You can arrange Minitab windows as you see fit and have more than one worksheet window open in a project. To view a window that may be obscured or hidden, select **Window** from the Minitab menu bar, and then select the name of the window you want to view.



To make effective use of Minitab, you should be familiar with how to open and save Minitab worksheet and project files as well as how to insert worksheets in a project, and how to print parts of a project. These skills are summarized in Appendix B. Minitab Guide instructions work best with the current commercial and student versions of Minitab and note differences when they occur.

## MG.2 ENTERING DATA

Minitab uses the standard business convention, expecting data for a variable to be entered into a column. In this book, data are entered in columns, left to right, starting with the first column. Column names take the form  $C_n$ , such that the first column is named C1, the second column is C2, and the tenth column is C10. Column names appear in the top border of a Minitab worksheet. Columns that contain non-numerical data have names that include “-T” (C1-T, C2-T, and C3-T in the DATA worksheet shown above). Columns that contain data that Minitab interprets as either dates or times have names that include “-D” (not seen in the DATA worksheet).

When entering data, you use the first, unnumbered and shaded row to enter variable names. You can then refer to the column by that name or its  $C_n$  name in Minitab procedures. If a variable name contains spaces or other special characters, such as **Market Cap**, Minitab will display that name in dialog boxes using a pair of single quotation marks (**'Market Cap'**). You must include those quotation marks any time you enter such a variable name in a dialog box.

To enter or edit data in a specific cell, either use the cursor keys to move the cell pointer to the cell or use your mouse to select the cell directly. Never skip a cell in numbered row when entering data because Minitab will interpret that skipped cell as a “missing value” (see Section 1.2).