



Basic Business Statistics

Concepts and Applications

14TH EDITION

mark BERENSON
david LEVINE
kathryn SZABAT
david STEPHAN



A ROADMAP FOR SELECTING A STATISTICAL METHOD

Data Analysis Task	For Numerical Variables	For Categorical Variables
Describing a group or several groups	<p>Ordered array, stem-and-leaf display, frequency distribution, relative frequency distribution, percentage distribution, cumulative percentage distribution, histogram, polygon, cumulative percentage polygon (Sections 2.2, 2.4)</p> <p>Mean, median, mode, geometric mean, quartiles, range, interquartile range, standard deviation, variance, coefficient of variation, skewness, kurtosis, boxplot, normal probability plot (Sections 3.1, 3.2, 3.3, 6.3)</p> <p>Index numbers (online Section 16.8)</p> <p>Dashboards (Section 17.2)</p>	<p>Summary table, bar chart, pie chart, doughnut chart, Pareto chart (Sections 2.1 and 2.3)</p>
Inference about one group	<p>Confidence interval estimate of the mean (Sections 8.1 and 8.2)</p> <p>t test for the mean (Section 9.2)</p> <p>Chi-square test for a variance or standard deviation (online Section 12.7)</p>	<p>Confidence interval estimate of the proportion (Section 8.3)</p> <p>Z test for the proportion (Section 9.4)</p>
Comparing two groups	<p>Tests for the difference in the means of two independent populations (Section 10.1)</p> <p>Wilcoxon rank sum test (Section 12.4)</p> <p>Paired t test (Section 10.2)</p> <p>F test for the difference between two variances (Section 10.4)</p> <p>Wilcoxon signed ranks test (online Section 12.8)</p>	<p>Z test for the difference between two proportions (Section 10.3)</p> <p>Chi-square test for the difference between two proportions (Section 12.1)</p> <p>McNemar test for two related samples (online Section 12.6)</p>
Comparing more than two groups	<p>One-way analysis of variance for comparing several means (Section 11.1)</p> <p>Kruskal-Wallis test (Section 12.5)</p> <p>Randomized block design (online Section 11.3)</p> <p>Two-way analysis of variance (Section 11.2)</p> <p>Friedman rank test (online Section 12.9)</p>	<p>Chi-square test for differences among more than two proportions (Section 12.2)</p>
Analyzing the relationship between two variables	<p>Scatter plot, time series plot (Section 2.5)</p> <p>Covariance, coefficient of correlation (Section 3.5)</p> <p>Simple linear regression (Chapter 13)</p> <p>t test of correlation (Section 13.7)</p> <p>Time-series forecasting (Chapter 16)</p> <p>Sparklines (Section 2.7)</p>	<p>Contingency table, side-by-side bar chart, PivotTables (Sections 2.1, 2.3, 2.6)</p> <p>Chi-square test of independence (Section 12.3)</p>
Analyzing the relationship between two or more variables	<p>Colored scatter plots, bubble chart, treemap (Section 2.7)</p> <p>Multiple regression (Chapters 14 and 15)</p> <p>Dynamic bubble charts (Section 17.2)</p> <p>Regression trees (Section 17.3)</p> <p>Cluster analysis (Section 17.5)</p> <p>Multidimensional scaling (Section 17.6)</p>	<p>Multidimensional contingency tables (Section 2.6)</p> <p>Drilldown and slicers (Section 2.7)</p> <p>Logistic regression (Section 14.7)</p> <p>Classification trees (Section 17.4)</p> <p>Multiple correspondence analysis (Section 17.6)</p>

This page intentionally left blank

Basic Business Statistics

Concepts and Applications

This page intentionally left blank

Basic Business Statistics Concepts and Applications

FOURTEENTH EDITION

Mark L. Berenson

Department of Information Management and Business Analytics
School of Business, Montclair State University

David M. Levine

Department of Information Systems and Statistics
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

Director, Portfolio Management: Marcia Horton
Editor in Chief: Deirdre Lynch
Courseware Portfolio Manager: Suzanna Bainbridge
Courseware Portfolio Management Assistant: Morgan Danna
Managing Producer: Karen Wernholm
Content Producer: Sherry Berg and Dana Bettez
Senior Producer: Stephanie Green
Associate Content Producer: Sneh Singh
Manager, Courseware QA: Mary Durnwald
Manager, Content Development: Robert Carroll
Product Marketing Manager: Kaylee Carlson
Product Marketing Assistant: Shannon McCormack
Field Marketing Manager: Thomas Hayward
Field Marketing Assistant: Derrica Moser
Senior Author Support/Technology Specialist: Joe Vetere
Manager, Rights and Permissions: Gina Cheselka
Manufacturing Buyer: Carol Melville, LSC Communications
Composition and Production Coordination: SPi Global
Text Design: Cenveo Publisher Services
Senior Designer and Cover Design: Barbara T. Atkinson
Cover Image: ArtisticPhoto/Shutterstock

Copyright © 2019, 2015, 2012 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/.

Attributions of third party content appear on page 805, which constitutes an extension of this copyright page.

PEARSON, ALWAYS LEARNING, MYLAB, MYLAB PLUS, MATHXL, LEARNING CATALYTICS, and TESTGEN are exclusive trademarks owned by Pearson Education, Inc. or its affiliates in the U.S. and/or other countries.

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® Windows®, and Microsoft office® 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.

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.

Library of Congress CIP data

Names: Berenson, Mark L., author. | Levine, David M., 1946– author. | Szabat, | Kathryn A., author.

Title: Basic business statistics : concepts and applications / Mark L.

Berenson, Department of Information Management and Business Analytics,
School of Business, Montclair State University, David M. Levine,
Department of Information Systems and Business Analytics, 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, NY.

Description: Fourteenth Edition. | New York : Pearson Education, [2017]

Revised edition of the authors' Basic business statistics, [2014]

Identifiers: LCCN 2017050511 | ISBN 9780134684840

Subjects: LCSH: Commercial statistics. | Statistics.

Classification: LCC HF1017 .B38 2017 | DDC 519.5—dc23

LC record available at <https://lccn.loc.gov/2017050511>

*To our spouses and children,
Rhoda, Marilyn, Mary, Kathy, Lori, Sharyn, and Mark*

*and to our parents, in loving memory,
Nat, Ethel, Lee, Reuben, Mary, William, Ruth and Francis J.*

About the Authors



Kathryn Szabat, David Levine, Mark Berenson,
and David Stephan

Mark L. Berenson, 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.

Mark L. Berenson is Professor of Information Management and Business Analytics at Montclair State University and also Professor Emeritus of Information Systems and Statistics at Baruch College. He currently teaches graduate and undergraduate courses in statistics and in operations management in the School of Business and an undergraduate course in international justice and human rights that he co-developed in the College of Humanities and Social Sciences.

Berenson received a B.A. in economic statistics and an M.B.A. in business statistics from City College of New York and a Ph.D. in

business from the City University of New York. Berenson's research has been published in *Decision Sciences Journal of Innovative Education*, *Review of Business Research*, *The American Statistician*, *Communications in Statistics*, *Psychometrika*, *Educational and Psychological Measurement*, *Journal of Management Sciences and Applied Cybernetics*, *Research Quarterly*, *Stats Magazine*, *The New York Statistician*, *Journal of Health Administration Education*, *Journal of Behavioral Medicine*, and *Journal of Surgical Oncology*. His invited articles have appeared in *The Encyclopedia of Measurement & Statistics* and *Encyclopedia of Statistical Sciences*. He has coauthored numerous statistics texts published by Pearson.

Over the years, Berenson has received several awards for teaching and for innovative contributions to statistics education. In 2005, he was the first recipient of the Catherine A. Becker Service for Educational Excellence Award at Montclair State University and, in 2012, he was the recipient of the Khubani/Telebrands Faculty Research Fellowship in the School of Business.

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 coauthors have been active participants in the annual DSI Data, Analytics, and Statistics Instruction (DASI) 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 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 DASI. 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 main-frame-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 SEDSI and DSI DASI (formerly 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 completed the instructional technology graduate program at Teachers College, Columbia University.

For all four 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.

This page intentionally left blank

Brief Contents

Preface xxiv

First Things First 1

1 Defining and Collecting Data 16

2 Organizing and Visualizing Variables 41

3 Numerical Descriptive Measures 120

4 Basic Probability 168

5 Discrete Probability Distributions 199

6 The Normal Distribution and Other Continuous Distributions 223

7 Sampling Distributions 252

8 Confidence Interval Estimation 275

9 Fundamentals of Hypothesis Testing: One-Sample Tests 311

10 Two-Sample Tests 351

11 Analysis of Variance 398

12 Chi-Square and Nonparametric Tests 440

13 Simple Linear Regression 484

14 Introduction to Multiple Regression 536

15 Multiple Regression Model Building 592

16 Time-Series Forecasting 629

17 Business Analytics 678

18 Getting Ready to Analyze Data in the Future 704

19 Statistical Applications in Quality Management (*online*) 19-1

20 Decision Making (*online*) 20-1

Appendices A-H 711

Self-Test Solutions and Answers to Selected Even-Numbered Problems 761

Index 793

Credits 805

This page intentionally left blank

Contents

Preface xxiv

First Things First 1

USING STATISTICS: “The Price of Admission” 1

FTF.1 Think Differently About Statistics 2

Statistics: A Way of Thinking 2

Statistics: An Important Part of Your Business Education 3

FTF.2 Business Analytics: The Changing Face of Statistics 4

“Big Data” 4

FTF.3 Starting Point for Learning Statistics 5

Statistic 5

Can Statistics (p !., statistic) Lie? 6

FTF.4 Starting Point for Using Software 6

Using Software Properly 8

REFERENCES 9

KEY TERMS 9

EXCEL GUIDE 10

EG.1 Getting Started with Excel 10

EG.2 Entering Data 10

EG.3 Open or Save a Workbook 10

EG.4 Working with a Workbook 11

EG.5 Print a Worksheet 11

EG.6 Reviewing Worksheets 11

EG.7 If You Use the *Workbook* Instructions 11

JMP GUIDE 12

JG.1 Getting Started with JMP 12

JG.2 Entering Data 13

JG.3 Create New Project or Data Table 13

JG.4 Open or Save Files 13

JG.5 Print Data Tables or Report Windows 13

JG.6 JMP Script Files 13

MINITAB GUIDE 14

MG.1 Getting Started with Minitab 14

MG.2 Entering Data 14

MG.3 Open or Save Files 14

MG.4 Insert or Copy Worksheets 15

MG.5 Print Worksheets 15

1 Defining and Collecting Data 16

USING STATISTICS: Defining Moments 16

1.1 Defining Variables 17

Classifying Variables by Type 17

Measurement Scales 18

1.2 Collecting Data 19

Populations and Samples 20

Data Sources 20

1.3 Types of Sampling Methods 21

Simple Random Sample 22

Systematic Sample 22

Stratified Sample 23

Cluster Sample 23

1.4 Data Cleaning 24

Invalid Variable Values 25

Coding Errors 25

Data Integration Errors 25

Missing Values 26

Algorithmic Cleaning of Extreme Numerical Values 26

1.5 Other Data Preprocessing Tasks 26

Data Formatting 26

Stacking and Unstacking Data 27

Recoding Variables 27

1.6 Types of Survey Errors 28

Coverage Error 28

Nonresponse Error 28

Sampling Error 28

Measurement Error 29

Ethical Issues About Surveys 29

CONSIDER THIS: New Media Surveys/Old Survey Errors 29

USING STATISTICS: Defining Moments, Revisited 31

SUMMARY 31

REFERENCES 31

KEY TERMS 31

CHECKING YOUR UNDERSTANDING 32

CHAPTER REVIEW PROBLEMS 32

CASES FOR CHAPTER 1 33

Managing Ashland MultiComm Services 33

CardioGood Fitness 33

Clear Mountain State Student Survey 34

Learning with the Digital Cases 34

CHAPTER 1 EXCEL GUIDE 35

EG1.1 Defining Variables 35

EG1.2 Collecting Data 35

EG1.3 Types of Sampling Methods 35

EG1.4 Data Cleaning 36

EG1.5 Other Data Preprocessing 36

CHAPTER 1 JMP GUIDE 37

JG1.1 Defining Variables 37

JG1.2 Collecting Data 37

JG1.3 Types of Sampling Methods 37

JG1.4 Data Cleaning 38

JG1.5 Other Preprocessing Tasks 39

CHAPTER 1 MINITAB GUIDE 39

- MG1.1 Defining Variables 39
- MG1.2 Collecting Data 39
- MG1.3 Types of Sampling Methods 39
- MG1.4 Data Cleaning 40
- MG1.5 Other Preprocessing Tasks 40

2 Organizing and Visualizing Variables 41

USING STATISTICS: “The Choice *Is* Yours” 41

- 2.1** Organizing Categorical Variables 42
 - The Summary Table 42
 - The Contingency Table 43
- 2.2** Organizing Numerical Variables 46
 - The Frequency Distribution 47
 - Classes and Excel Bins 49
 - The Relative Frequency Distribution and the Percentage Distribution 49
 - The Cumulative Distribution 51
- 2.3** Visualizing Categorical Variables 54
 - The Bar Chart 54
 - The Pie Chart and the Doughnut Chart 55
 - The Pareto Chart 56
 - Visualizing Two Categorical Variables 58
- 2.4** Visualizing Numerical Variables 61
 - The Stem-and-Leaf Display 61
 - The Histogram 61
 - The Percentage Polygon 63
 - The Cumulative Percentage Polygon (Ogive) 64
- 2.5** Visualizing Two Numerical Variables 67
 - The Scatter Plot 67
 - The Time-Series Plot 68
- 2.6** Organizing a Mix of Variables 70
 - Drill-down 71
- 2.7** Visualizing a Mix of Variables 72
 - Colored Scatter Plot 72
 - Bubble Charts 73
 - PivotChart (Excel) 73
 - Treemap (Excel, JMP) 73
 - Sparklines (Excel) 74
- 2.8** Filtering and Querying Data 75
 - Excel Slicers 75
- 2.9** Pitfalls in Organizing and Visualizing Variables 77
 - Obscuring Data 77
 - Creating False Impressions 78
 - Chartjunk 79

EXHIBIT: Best Practices for Creating Visual Summaries 80**USING STATISTICS: “The Choice *Is* Yours,” Revisited 81****SUMMARY 81****REFERENCES 82****KEY EQUATIONS 82****KEY TERMS 83****CHECKING YOUR UNDERSTANDING 83****CHAPTER REVIEW PROBLEMS 83****CASES FOR CHAPTER 2 88**

- Managing Ashland MultiComm Services 88
- Digital Case 88
- CardioGood Fitness 89
- The Choice *Is* Yours Follow-Up 89
- Clear Mountain State Student Survey 89

CHAPTER 2 EXCEL GUIDE 90

- EG2.1 Organizing Categorical Variables 90
- EG2.2 Organizing Numerical Variables 92
 - Charts Group Reference 94
- EG2.3 Visualizing Categorical Variables 94
- EG2.4 Visualizing Numerical Variables 96
- EG2.5 Visualizing Two Numerical Variables 99
- EG2.6 Organizing a Mix of Variables 100
- EG2.7 Visualizing a Mix of Variables 101
- EG2.8 Filtering and Querying Data 102

CHAPTER 2 JMP GUIDE 102

- JG2 JMP Choices for Creating Summaries 102
- JG2.1 Organizing Categorical Variables 103
- JG2.2 Organizing Numerical Variables 104
- JG2.3 Visualizing Categorical Variables 106
- JG2.4 Visualizing Numerical Variables 107
- JG2.5 Visualizing Two Numerical Variables 109
- JG2.6 Organizing a Mix of Variables 110
- JG2.7 Visualizing a Mix of Variables 110
- JG2.8 Filtering and Querying Data 111
 - JMP Guide Gallery 112

CHAPTER 2 MINITAB GUIDE 113

- MG2.1 Organizing Categorical Variables 113
- MG2.2 Organizing Numerical Variables 113
- MG2.3 Visualizing Categorical Variables 113
- MG2.4 Visualizing Numerical Variables 115
- MG2.5 Visualizing Two Numerical Variables 117
- MG2.6 Organizing a Mix of Variables 118
- MG2.7 Visualizing a Mix of Variables 118
- MG2.8 Filtering and Querying Data 119

3 Numerical Descriptive Measures 120

USING STATISTICS: More Descriptive Choices 120

- 3.1** Measures of Central Tendency 121
 - The Mean 121
 - The Median 123
 - The Mode 124
 - The Geometric Mean 125
 - 3.2** Measures of Variation and Shape 126
 - The Range 126
 - The Variance and the Standard Deviation 127
 - The Coefficient of Variation 130
 - Z Scores 130
 - Shape: Skewness 132
 - Shape: Kurtosis 132
 - 3.3** Exploring Numerical Variables 137
 - Quartiles 137
- EXHIBIT:** Rules for Calculating the Quartiles from a Set of Ranked Values 137

The Interquartile Range 139
 The Five-Number Summary 139
 The Boxplot 141

- 3.4** Numerical Descriptive Measures for a Population 143
 The Population Mean 144
 The Population Variance and Standard Deviation 144
 The Empirical Rule 145
 Chebyshev's Theorem 146

- 3.5** The Covariance and the Coefficient of Correlation 148
 The Covariance 148
 The Coefficient of Correlation 148

- 3.6** Descriptive Statistics: Pitfalls and Ethical Issues 152

USING STATISTICS: More Descriptive Choices, Revisited 153

SUMMARY 153

REFERENCES 154

KEY EQUATIONS 154

KEY TERMS 154

CHECKING YOUR UNDERSTANDING 155

CHAPTER REVIEW PROBLEMS 155

CASES FOR CHAPTER 3 158

Managing Ashland MultiComm Services 158
 Digital Case 158
 CardioGood Fitness 158
 More Descriptive Choices Follow-up 159
 Clear Mountain State Student Survey 159

CHAPTER 3 EXCEL GUIDE 160

EG3.1 Measures of Central Tendency 160
 EG3.2 Measures of Variation and Shape 161
 EG3.3 Exploring Numerical Variables 161
 EG3.4 Numerical Descriptive Measures for a Population 162
 EG3.5 The Covariance and the Coefficient of Correlation 162

CHAPTER 3 JMP GUIDE 163

JG3.1 Measures of Central Tendency 163
 JG3.2 Measures of Variation and Shape 163
 JG3.3 Exploring Numerical Variables 164
 JG3.4 Numerical Descriptive Measures for a Population 164
 JG3.5 The Covariance and the Coefficient of Correlation 164

CHAPTER 3 MINITAB GUIDE 165

MG3.1 Measures of Central Tendency 165
 MG3.2 Measures of Variation and Shape 166
 MG3.3 Exploring Numerical Variables 166
 MG3.4 Numerical Descriptive Measures for a Population 167
 MG3.5 The Covariance and the Coefficient of Correlation 167

4 Basic Probability 168

USING STATISTICS: Possibilities at M&R Electronics World 168

- 4.1** Basic Probability Concepts 169
 Events and Sample Spaces 169
 Types of Probability 170
 Summarizing Sample Spaces 171
 Simple Probability 172
 Joint Probability 173
 Marginal Probability 174
 General Addition Rule 174

- 4.2** Conditional Probability 178
 Computing Conditional Probabilities 178
 Decision Trees 179
 Independence 181
 Multiplication Rules 182
 Marginal Probability Using the General Multiplication Rule 183

- 4.3** Ethical Issues and Probability 185

- 4.4** Bayes' Theorem 186

CONSIDER THIS: Divine Providence and Spam 188

- 4.5** Counting Rules 189

USING STATISTICS: Possibilities at M&R Electronics World, Revisited 192

SUMMARY 193

REFERENCES 193

KEY EQUATIONS 193

KEY TERMS 194

CHECKING YOUR UNDERSTANDING 194

CHAPTER REVIEW PROBLEMS 194

CASES FOR CHAPTER 4 196

Digital Case 196
 CardioGood Fitness 196
 The Choice Is Yours Follow-Up 196
 Clear Mountain State Student Survey 196

CHAPTER 4 EXCEL GUIDE 197

EG4.1 Basic Probability Concepts 197
 EG4.4 Bayes' Theorem 197
 EG4.5 Counting Rules 197

CHAPTER 4 JMP

JG4.4 Bayes' Theorem 198

CHAPTER 4 MINITAB GUIDE 198

MG4.5 Counting Rules 198

5 Discrete Probability Distributions 199

USING STATISTICS: Events of Interest at Ricknel Home Centers 199

- 5.1** The Probability Distribution for a Discrete Variable 200
 Expected Value of a Discrete Variable 200
 Variance and Standard Deviation of a Discrete Variable 201
- 5.2** Binomial Distribution 204
EXHIBIT: Properties of the Binomial Distribution 204
 Histograms for Discrete Variables 207
 Summary Measures for the Binomial Distribution 208
- 5.3** Poisson Distribution 211
- 5.4** Covariance of a Probability Distribution and Its Application in Finance 214
- 5.5** Hypergeometric Distribution (*online*) 214
- 5.6** Using the Poisson Distribution to Approximate the Binomial Distribution (*online*) 214

USING STATISTICS: Events of Interest..., Revisited 215

SUMMARY 215

REFERENCES 215

KEY EQUATIONS 215

KEY TERMS 216

CHECKING YOUR UNDERSTANDING 216

CHAPTER REVIEW PROBLEMS 216

CASES FOR CHAPTER 5 218

- Managing Ashland MultiComm Services 218
- Digital Case 218

CHAPTER 5 EXCEL GUIDE 219

- EG5.1 The Probability Distribution for a Discrete Variable 219
- EG5.2 Binomial Distribution 219
- EG5.3 Poisson Distribution 219

CHAPTER 5 JMP GUIDE 220

- JG5.1 The Probability Distribution for a Discrete Variable 220
- JG5.2 Binomial Distribution 220
- JG5.3 Poisson Distribution 221

CHAPTER 5 MINITAB GUIDE 221

- MG5.1 The Probability Distribution for a Discrete Variable 221
- MG5.2 Binomial Distribution 222
- MG5.3 Poisson Distribution 222

6 The Normal Distribution and Other Continuous Distributions 223

USING STATISTICS: Normal Load Times at MyTVLab 223

6.1 Continuous Probability Distributions 224

6.2 The Normal Distribution 224

- EXHIBIT:** Normal Distribution Important Theoretical Properties 225
 - Role of the Mean and the Standard Deviation 226
 - Calculating Normal Probabilities 227
- VISUAL EXPLORATIONS:** Exploring the Normal Distribution 231
 - Finding X Values 232

CONSIDER THIS: What Is Normal? 235

6.3 Evaluating Normality 237

- Comparing Data Characteristics to Theoretical Properties 237
- Constructing the Normal Probability Plot 238

6.4 The Uniform Distribution 241

6.5 The Exponential Distribution (*online*) 2436.6 The Normal Approximation to the Binomial Distribution (*online*) 243
USING STATISTICS: Normal Load Times..., Revisited 243

SUMMARY 243

REFERENCES 244

KEY EQUATIONS 244

KEY TERMS 244

CHECKING YOUR UNDERSTANDING 245

CHAPTER REVIEW PROBLEMS 245

CASES FOR CHAPTER 6 246

- Managing Ashland MultiComm Services 246
- CardioGood Fitness 247
- More Descriptive Choices Follow-up 247
- Clear Mountain State Student Survey 247
- Digital Case 247

CHAPTER 6 EXCEL GUIDE 248

- EG6.2 The Normal Distribution 248
- EG6.3 Evaluating Normality 248

CHAPTER 6 JMP GUIDE 249

- JG6.2 The Normal Distribution 249
- JG6.3 Evaluating Normality 249

CHAPTER 6 MINITAB GUIDE 250

- MG6.2 The Normal Distribution 250
- MG6.3 Evaluating Normality 250

7 Sampling Distributions 252

USING STATISTICS: Sampling Oxford Cereals 252

7.1 Sampling Distributions 253

7.2 Sampling Distribution of the Mean 253

- The Unbiased Property of the Sample Mean 253
- Standard Error of the Mean 255
- Sampling from Normally Distributed Populations 256
- Sampling from Non-normally Distributed Populations—The Central Limit Theorem 259

EXHIBIT: Normality and the Sampling Distribution of the Mean 260

- VISUAL EXPLORATIONS:** Exploring Sampling Distributions 263

7.3 Sampling Distribution of the Proportion 264

7.4 Sampling from Finite Populations (*online*) 267
USING STATISTICS: Sampling Oxford Cereals, Revisited 267

SUMMARY 268

REFERENCES 268

KEY EQUATIONS 268

KEY TERMS 268

CHECKING YOUR UNDERSTANDING 269

CHAPTER REVIEW PROBLEMS 269

CASES FOR CHAPTER 7 270

- Managing Ashland MultiComm Services 270
- Digital Case 271

CHAPTER 7 EXCEL GUIDE 272

- EG7.2 Sampling Distribution of the Mean 272

CHAPTER 7 JMP GUIDE 273

- JG7.2 Sampling Distribution of the Mean 273

CHAPTER 7 MINITAB GUIDE 274

- MG7.2 Sampling Distribution of the Mean 274

8 Confidence Interval Estimation 275

USING STATISTICS: Getting Estimates at Ricknel Home Centers 275

- 8.1 Confidence Interval Estimate for the Mean (σ Known) 276
 - Sampling Error 277
 - Can You Ever Know the Population Standard Deviation? 280
- 8.2 Confidence Interval Estimate for the Mean (σ Unknown) 281
 - Student's t Distribution 282
 - The Concept of Degrees of Freedom 282
 - Properties of the t Distribution 282
 - The Confidence Interval Statement 284
- 8.3 Confidence Interval Estimate for the Proportion 289
- 8.4 Determining Sample Size 292
 - Sample Size Determination for the Mean 292
 - Sample Size Determination for the Proportion 294
- 8.5 Confidence Interval Estimation and Ethical Issues 297
- 8.6 Application of Confidence Interval Estimation in Auditing (*online*) 297
- 8.7 Estimation and Sample Size Estimation for Finite Populations (*online*) 298
- 8.8 Bootstrapping (*online*) 298

USING STATISTICS: Getting Estimates . . . , Revisited 298

SUMMARY 298

REFERENCES 299

KEY EQUATIONS 299

KEY TERMS 299

CHECKING YOUR UNDERSTANDING 299

CHAPTER REVIEW PROBLEMS 300

CASES FOR CHAPTER 8 302

- Managing Ashland MultiComm Services 302
- Digital Case 303
- Sure Value Convenience Stores 304
- CardioGood Fitness 304
- More Descriptive Choices Follow-Up 304
- Clear Mountain State Student Survey 304

CHAPTER 8 EXCEL GUIDE 305

- EG8.1 Confidence Interval Estimate for the Mean (σ Known) 305
- EG8.2 Confidence Interval Estimate for the Mean (σ Unknown) 305
- EG8.3 Confidence Interval Estimate for the Proportion 306
- EG8.4 Determining Sample Size 306

CHAPTER 8 JMP GUIDE 307

- JG8.1 Confidence Interval Estimate for the Mean (σ Known) 307
- JG8.2 Confidence Interval Estimate for the Mean (σ Unknown) 307
- JG8.3 Confidence Interval Estimate for the Proportion 308
- JG8.4 Determining Sample Size 308

CHAPTER 8 MINITAB GUIDE 309

- MG8.1 Confidence Interval Estimate for the Mean (σ Known) 309
- MG8.2 Confidence Interval Estimate for the Mean (σ Unknown) 309

MG8.3 Confidence Interval Estimate for the Proportion 310

MG8.4 Determining Sample Size 310

9 Fundamentals of Hypothesis Testing: One-Sample Tests 311

USING STATISTICS: Significant Testing at Oxford Cereals 311

- 9.1 Fundamentals of Hypothesis Testing 312
 - EXHIBIT:** Fundamental Hypothesis Testing Concepts 313
 - The Critical Value of the Test Statistic 313
 - Regions of Rejection and Nonrejection 314
 - Risks in Decision Making Using Hypothesis Testing 314
 - Z Test for the Mean (σ Known) 316
 - Hypothesis Testing Using the Critical Value Approach 317
 - EXHIBIT:** The Critical Value Approach to Hypothesis Testing 318
 - Hypothesis Testing Using the p -Value Approach 320
 - EXHIBIT:** The p -Value Approach to Hypothesis Testing 321
 - A Connection Between Confidence Interval Estimation and Hypothesis Testing 322
 - Can You Ever Know the Population Standard Deviation? 323
 - 9.2 t Test of Hypothesis for the Mean (σ Unknown) 324
 - The Critical Value Approach 325
 - p -Value Approach 326
 - Checking the Normality Assumption 327
 - 9.3 One-Tail Tests 330
 - The Critical Value Approach 330
 - The p -Value Approach 331
 - EXHIBIT:** The Null and Alternative Hypotheses in One-Tail Tests 333
 - 9.4 Z Test of Hypothesis for the Proportion 334
 - The Critical Value Approach 335
 - The p -Value Approach 336
 - 9.5 Potential Hypothesis-Testing Pitfalls and Ethical Issues 338
 - EXHIBIT:** Questions for the Planning Stage of Hypothesis Testing 338
 - Statistical Significance Versus Practical Significance 338
 - Statistical *Insignificance* Versus Importance 339
 - Reporting of Findings 339
 - Ethical Issues 339
 - 9.6 Power of the Test (*online*) 339
- USING STATISTICS:** Significant Testing . . . Revisited 340
- SUMMARY** 340
- REFERENCES** 340
- KEY EQUATIONS** 341
- KEY TERMS** 341
- CHECKING YOUR UNDERSTANDING** 341
- CHAPTER REVIEW PROBLEMS** 341
- CASES FOR CHAPTER 9** 343
- Managing Ashland MultiComm Services 343
 - Digital Case 343
 - Sure Value Convenience Stores 344

CHAPTER 9 EXCEL GUIDE 345

- EG9.1 Fundamentals of Hypothesis Testing 345
- EG9.2 t Test Of Hypothesis for The Mean (σ Unknown) 345
- EG9.3 One-Tail Tests 346
- EG9.4 Z Test Of Hypothesis For The Proportion 346

CHAPTER 9 JMP GUIDE 347

- JG9.1 Fundamentals of Hypothesis Testing 347
- JG9.2 t Test of Hypothesis for the Mean (σ Unknown) 347
- JG9.3 One-Tail Tests 348
- JG9.4 Z Test Of Hypothesis For The Proportion 348

CHAPTER 9 MINITAB GUIDE 348

- MG9.1 Fundamentals of Hypothesis Testing 348
- MG9.2 t Test Of Hypothesis for The Mean (σ Unknown) 349
- MG9.3 One-Tail Tests 349
- MG9.4 Z Test Of Hypothesis for The Proportion 349

10 Two-Sample Tests 351**USING STATISTICS:** Differing Means for Selling Streaming Media Players at Arlingtons? 351**10.1** Comparing the Means of Two Independent Populations 352

- Pooled-Variance t Test for the Difference Between Two Means Assuming Equal Variances 352
- Evaluating the Normality Assumption 355
- Confidence Interval Estimate for the Difference Between Two Means 357
- Separate-Variance t Test for the Difference Between Two Means, Assuming Unequal Variances 358

CONSIDER THIS: Do People Really Do This? 359**10.2** Comparing the Means of Two Related Populations 361

- Paired t Test 362
- Confidence Interval Estimate for the Mean Difference 367

10.3 Comparing the Proportions of Two Independent Populations 369

- Z Test for the Difference Between Two Proportions 369
- Confidence Interval Estimate for the Difference Between Two Proportions 374

10.4 F Test for the Ratio of Two Variances 376**10.5** Effect Size (*online*) 380**USING STATISTICS:** Differing Means for Selling ... , Revisited 381**SUMMARY 381****REFERENCES 382****KEY EQUATIONS 382****KEY TERMS 383****CHECKING YOUR UNDERSTANDING 383****CHAPTER REVIEW PROBLEMS 383****CASES FOR CHAPTER 10 385**

- Managing Ashland MultiComm Services 385
- Digital Case 386
- Sure Value Convenience Stores 386
- CardioGood Fitness 386
- More Descriptive Choices Follow-Up 386
- Clear Mountain State Student Survey 387

CHAPTER 10 EXCEL GUIDE 388

- EG10.1 Comparing The Means of Two Independent Populations 388
- EG10.2 Comparing The Means of Two Related Populations 390
- EG10.3 Comparing The Proportions of Two Independent Populations 391
- EG10.4 F Test For The Ratio of Two Variances 392

CHAPTER 10 JMP GUIDE 393

- JG10.1 Comparing The Means of Two Independent Populations 393
- JG10.2 Comparing The Means of Two Related Populations 394
- JG10.3 Comparing The Proportions of Two Independent Populations 394
- JG10.4 F Test For The Ratio of Two Variances 394

CHAPTER 10 MINITAB GUIDE 395

- MG10.1 Comparing The Means of Two Independent Populations 395
- MG10.2 Comparing The Means of Two Related Populations 396
- MG10.3 Comparing The Proportions of Two Independent Populations 396
- MG10.4 F Test For The Ratio of Two Variances 397

11 Analysis of Variance 398**USING STATISTICS:** The Means to Find Differences at Arlingtons 398**11.1** The Completely Randomized Design: One-Way ANOVA 399

- Analyzing Variation in One-Way ANOVA 400
- F Test for Differences Among More Than Two Means 402
- One-Way ANOVA F Test Assumptions 407
- Levene Test for Homogeneity of Variance 407
- Multiple Comparisons: The Tukey-Kramer Procedure 409
- The Analysis of Means (ANOM) 411

11.2 The Factorial Design: Two-Way ANOVA 414

- Factor and Interaction Effects 415
- Testing for Factor and Interaction Effects 416
- Multiple Comparisons: The Tukey Procedure 420
- Visualizing Interaction Effects: The Cell Means Plot 421
- Interpreting Interaction Effects 422

11.3 The Randomized Block Design (*online*) 426**11.4** Fixed Effects, Random Effects, and Mixed Effects Models (*online*) 426**USING STATISTICS:** The Means to Find Differences at Arlingtons Revisited 426**SUMMARY 426****REFERENCES 427****KEY EQUATIONS 427****KEY TERMS 428****CHECKING YOUR UNDERSTANDING 428****CHAPTER REVIEW PROBLEMS 428****CASES FOR CHAPTER 11 430**

- Managing Ashland MultiComm Services 430
- Digital Case 431
- Sure Value Convenience Stores 431
- CardioGood Fitness 431
- More Descriptive Choices Follow-Up 431
- Clear Mountain State Student Survey 431

CHAPTER 11 EXCEL GUIDE 432

- EG11.1 The Completely Randomized Design: One-Way ANOVA 432
- EG11.2 The Factorial Design: Two-Way Anova 434

CHAPTER 11 JMP GUIDE 435

- JG11.1 The Completely Randomized Design: One-Way ANOVA 435
- JG11.2 The Factorial Design: Two-Way Anova 436

CHAPTER 11 MINITAB GUIDE 437

- MG11.1 The Completely Randomized Design: One-Way ANOVA 437
- MG11.2 The Factorial Design: Two-Way Anova 438

12 Chi-Square and Nonparametric Tests 440

USING STATISTICS: Avoiding Guesswork About Resort Guests 440

- 12.1** Chi-Square Test for the Difference Between Two Proportions 441
- 12.2** Chi-Square Test for Differences Among More Than Two Proportions 448
 - The Marascuilo Procedure 451
 - The Analysis of Proportions (ANOP) 453
- 12.3** Chi-Square Test of Independence 454
- 12.4** Wilcoxon Rank Sum Test for Two Independent Populations 460
- 12.5** Kruskal-Wallis Rank Test for the One-Way ANOVA 466
 - Assumptions of the Kruskal-Wallis Rank Test 469
- 12.6** McNemar Test for the Difference Between Two Proportions (Related Samples) (*online*) 470
- 12.7** Chi-Square Test for the Variance or Standard Deviation (*online*) 470
- 12.8** Wilcoxon Signed Ranks Test for Two Related Populations (*online*) 471
- 12.9** Friedman Rank Test for the Randomized Block Design (*online*) 471

USING STATISTICS: Avoiding Guesswork..., Revisited 471

SUMMARY 471**REFERENCES 472****KEY EQUATIONS 472****KEY TERMS 473****CHECKING YOUR UNDERSTANDING 473****CHAPTER REVIEW PROBLEMS 473****CASES FOR CHAPTER 12 475**

- Managing Ashland MultiComm Services 475
- Digital Case 476
- Sure Value Convenience Stores 476
- CardioGood Fitness 476
- More Descriptive Choices Follow-Up 477
- Clear Mountain State Student Survey 477

CHAPTER 12 EXCEL GUIDE 478

- EG12.1 Chi-Square Test for the Difference Between Two Proportions 478
- EG12.2 Chi-Square Test for Differences Among More Than Two Proportions 478

- EG12.3 Chi-Square Test of Independence 479
- EG12.4 Wilcoxon Rank Sum Test: A Nonparametric Method for Two Independent Populations 479
- EG12.5 Kruskal-Wallis Rank Test: A Nonparametric Method for the One-Way ANOVA 480

CHAPTER 12 JMP GUIDE 481

- JG12.1 Chi-Square Test for the Difference Between Two Proportions 481
- JG12.2 Chi-Square Test for Difference Among More Than Two Proportions 481
- JG12.3 Chi-Square Test Of Independence 481
- JG12.4 Wilcoxon Rank Sum Test For Two Independent Populations 481
- JG12.5 Kruskal-Wallis Rank Test For The One-Way Anova 482

CHAPTER 12 MINITAB GUIDE 482

- MG12.1 Chi-Square Test for The Difference Between Two Proportions 482
- MG12.2 Chi-Square Test for Differences Among More Than Two Proportions 483
- MG12.3 Chi-Square Test of Independence 483
- MG12.4 Wilcoxon Rank Sum Test: A Nonparametric Method for Two Independent Populations 483
- MG12.5 Kruskal-Wallis Rank Test: A Nonparametric Method for The One-Way ANOVA 483

13 Simple Linear Regression 484

USING STATISTICS: Knowing Customers at Sunflowers Apparel 484

- Preliminary Analysis 485

- 13.1** Simple Linear Regression Models 486
- 13.2** Determining the Simple Linear Regression Equation 487
 - The Least-Squares Method 487
 - Predictions in Regression Analysis: Interpolation Versus Extrapolation 490
 - Computing the Y Intercept, b_0 and the Slope, b_1 491
 - VISUAL EXPLORATIONS: Exploring Simple Linear Regression Coefficients 493
- 13.3** Measures of Variation 495
 - Computing the Sum of Squares 495
 - The Coefficient of Determination 496
 - Standard Error of the Estimate 498
- 13.4** Assumptions of Regression 500
- 13.5** Residual Analysis 500
 - Evaluating the Assumptions 500
- 13.6** Measuring Autocorrelation: The Durbin-Watson Statistic 504
 - Residual Plots to Detect Autocorrelation 504
 - The Durbin-Watson Statistic 505
- 13.7** Inferences About the Slope and Correlation Coefficient 508
 - t Test for the Slope 508
 - F Test for the Slope 509
 - Confidence Interval Estimate for the Slope 511
 - t Test for the Correlation Coefficient 511

13.8	Estimation of Mean Values and Prediction of Individual Values 514
	The Confidence Interval Estimate for the Mean Response 515
	The Prediction Interval for an Individual Response 516
13.9	Potential Pitfalls in Regression 518
	EXHIBIT: Seven Steps for Avoiding the Potential Pitfalls 518
	USING STATISTICS: Knowing Customers..., Revisited 520
	SUMMARY 521
	REFERENCES 522
	KEY EQUATIONS 522
	KEY TERMS 523
	CHECKING YOUR UNDERSTANDING 523
	CHAPTER REVIEW PROBLEMS 524
	CASES FOR CHAPTER 13 527
	Managing Ashland MultiComm Services 527
	Digital Case 527
	Brynn Packaging 528
	CHAPTER 13 EXCEL GUIDE 529
	EG13.2 Determining the Simple Linear Regression Equation 529
	EG13.3 Measures of Variation 530
	EG13.4 Assumptions of Regression 530
	EG13.5 Residual Analysis 530
	EG13.6 Measuring Autocorrelation: the Durbin-Watson Statistic 531
	EG13.7 Inferences about the Slope and Correlation Coefficient 531
	EG13.8 Estimation of Mean Values and Prediction of Individual Values 531
	CHAPTER 13 JMP GUIDE 532
	JG13.2 Determining The Simple Linear Regression Equation 532
	JG13.3 Measures Of Variation 532
	JG13.4 Assumptions Of Regression 532
	JG13.5 Residual Analysis 532
	JG13.6 Measuring Autocorrelation: The Durbin-Watson Statistic 532
	JG13.7 Inferences About The Slope And Correlation Coefficient 532
	JG13.8 Estimation Of Mean Values And Prediction Of Individual Values 533
	CHAPTER 13 MINITAB GUIDE 534
	MG13.2 Determining The Simple Linear Regression Equation 534
	MG13.3 Measures Of Variation 535
	MG13.4 Assumptions Of Regression 535
	MG13.5 Residual Analysis 535
	MG13.6 Measuring Autocorrelation: The Durbin-Watson Statistic 535
	MG13.7 Inferences about The Slope and Correlation Coefficient 535
	MG13.8 Estimation of Mean Values and Prediction of Individual Values 535

14 Introduction to Multiple Regression 536

USING STATISTICS: The Multiple Effects of OmniPower Bars 536

14.1	Developing a Multiple Regression Model 537
	Interpreting the Regression Coefficients 538
	Predicting the Dependent Variable Y 540
14.2	r^2 , Adjusted r^2 , and the Overall F Test 542
	Coefficient of Multiple Determination 542
	Adjusted r^2 543
	Test for the Significance of the Overall Multiple Regression Model 543

14.3	Multiple Regression Residual Analysis 546
14.4	Inferences About the Population Regression Coefficients 547
	Tests of Hypothesis 548
	Confidence Interval Estimation 549
14.5	Testing Portions of the Multiple Regression Model 551
	Coefficients of Partial Determination 555
14.6	Using Dummy Variables and Interaction Terms 557
	Interactions 560
14.7	Logistic Regression 569
14.8	Influence Analysis (<i>online</i>) 575
	USING STATISTICS: The Multiple Effects..., Revisited 575
	SUMMARY 575
	REFERENCES 577
	KEY EQUATIONS 577
	KEY TERMS 578
	CHECKING YOUR UNDERSTANDING 578
	CHAPTER REVIEW PROBLEMS 578
	CASES FOR CHAPTER 14 581
	Managing Ashland MultiComm Services 581
	Digital Case 581

CHAPTER 14 EXCEL GUIDE 582

EG14.1	Developing a Multiple Regression Model 582
EG14.2	r^2 , Adjusted r^2 , and the Overall F Test 583
EG14.3	Multiple Regression Residual Analysis 583
EG14.4	Inferences about the Population Regression Coefficients 584
EG14.5	Testing Portions of the Multiple Regression Model 584
EG14.6	Using Dummy Variables and Interaction Terms 584
EG14.7	Logistic Regression 585

CHAPTER 14 JMP GUIDE 585

JG14.1	Developing a Multiple Regression Model 585
JG14.2	r^2 , Adjusted r^2 , and the Overall F Test Measures of Variation 586
JG14.3	Multiple Regression Residual Analysis 586
JG14.4	Inferences About the Population 586
JG14.5	Testing Portions of the Multiple Regression Model 587
JG14.6	Using Dummy Variables and Interaction Terms 587
JG14.7	Logistic Regression 587

CHAPTER 14 MINITAB GUIDE 588

MG14.1	Developing a Multiple Regression Model 588
MG14.2	r^2 , Adjusted r^2 , and the Overall F Test 589
MG14.3	Multiple Regression Residual Analysis 589
MG14.4	Inferences About the Population Regression Coefficients 589
MG14.5	Testing Portions of the Multiple Regression Model 589
MG14.6	Using Dummy Variables and Interaction Terms in Regression Models 589
MG14.7	Logistic Regression 590
MG14.8	Influence Analysis 591

15 Multiple Regression Model Building 592

USING STATISTICS: Valuing Parsimony at WSTA-TV 592

15.1	Quadratic Regression Model 593
	Finding the Regression Coefficients and Predicting Y 594
	Testing for the Significance of the Quadratic Model 596

- Testing the Quadratic Effect 597
- The Coefficient of Multiple Determination 599
- 15.2** Using Transformations in Regression Models 601
 - The Square-Root Transformation 601
 - The Log Transformation 603
- 15.3** Collinearity 605
- 15.4** Model Building 607
 - EXHIBIT:** Successful Model Building 607
 - The Stepwise Regression Approach to Model Building 609
 - The Best Subsets Approach to Model Building 610
 - Model Validation 613
- 15.5** Pitfalls in Multiple Regression and Ethical Issues 615
 - Pitfalls in Multiple Regression 615
 - Ethical Issues 616
- USING STATISTICS:** Valuing Parsimony..., Revisited 616
- SUMMARY** 617
- REFERENCES** 618
- KEY EQUATIONS** 618
- KEY TERMS** 618
- CHECKING YOUR UNDERSTANDING** 618
- CHAPTER REVIEW PROBLEMS** 618
- CASES FOR CHAPTER 15** 620
 - The Mountain States Potato Company 620
 - Sure Value Convenience Stores 621
 - Digital Case 621
 - The Craybill Instrumentation Company Case 621
 - More Descriptive Choices Follow-Up 622
- CHAPATER 15 EXCEL GUIDE** 623
 - EG15.1 The Quadratic Regression Model 623
 - EG15.2 Using Transformations in Regression Models 623
 - EG15.3 Collinearity 624
 - EG15.4 Model Building 624
- CHAPATER 15 JMP GUIDE** 625
 - JG15.1 The Quadratic Regression Model 625
 - JG15.2 Using Transformations in Regression Models 625
 - JG15.3 Collinearity 625
 - JG15.4 Model Building 625
- CHAPATER 15 MINITAB GUIDE** 626
 - MG15.1 The Quadratic Regression Model 626
 - MG15.2 Using Transformations in Regression Models 627
 - MG15.3 Collinearity 627
 - MG15.4 Model Building 627

16 Time-Series Forecasting 629

- USING STATISTICS:** Is the ByYourDoor Service Trending? 629
- 16.1** Time Series Component Factors 630
- 16.2** Smoothing an Annual Time Series 632
 - Moving Averages 633
 - Exponential Smoothing 635
- 16.3** Least-Squares Trend Fitting and Forecasting 637
 - The Linear Trend Model 637
 - The Quadratic Trend Model 639
 - The Exponential Trend Model 640
 - Model Selection Using First, Second, and Percentage Differences 642

- EXHIBIT:** Model Selection Using First, Second, and Percentage Differences 642
- 16.4** Autoregressive Modeling for Trend Fitting and Forecasting 647
 - Selecting an Appropriate Autoregressive Model 648
 - Determining the Appropriateness of a Selected Model 649
 - EXHIBIT:** Autoregressive Modeling Steps 651
- 16.5** Choosing an Appropriate Forecasting Model 655
 - Residual Analysis 655
 - The Magnitude of the Residuals Through Squared or Absolute Differences 656
 - The Principle of Parsimony 656
 - A Comparison of Four Forecasting Methods 657
- 16.6** Time-Series Forecasting of Seasonal Data 659
 - Least-Squares Forecasting with Monthly or Quarterly Data 659
- 16.7** Index Numbers (*online*) 665
- CONSIDER THIS:** Let the Model User Beware 665
- USING STATISTICS:** Is the ByYourDoor..., Revisited 665
- SUMMARY** 665
- REFERENCES** 666
- KEY EQUATIONS** 666
- KEY TERMS** 667
- CHECKING YOUR UNDERSTANDING** 668
- CHAPTER REVIEW PROBLEMS** 668
- CASES FOR CHAPTER 16** 669
 - Managing Ashland MultiComm Services 669
 - Digital Case 669
- CHAPTER 16 EXCEL GUIDE** 670
 - EG16.2 Smoothing an Annual Time Series 670
 - EG16.3 Least-Squares Trend Fitting and Forecasting 671
 - EG16.4 Autoregressive Modeling for Trend Fitting and Forecasting 671
 - EG16.5 Choosing An Appropriate Forecasting Model 672
 - EG16.6 Time-Series Forecasting Of Seasonal Data 672
- CHAPTER 16 JMP GUIDE** 673
 - JG16.2 Smoothing an Annual Time Series 673
 - JG16.3 Least-Squares Trend Fitting and Forecasting 674
 - JG16.4 Autoregressive Modeling for Trend Fitting and Forecasting 674
 - JG16.5 Choosing an Appropriate Forecasting Model 675
 - JG16.6 Time-Series Forecasting of Seasonal Data 675
- CHAPTER 16 MINITAB GUIDE** 675
 - MG16.2 Smoothing an Annual Time Series 675
 - MG16.3 Least-Squares Trend Fitting and Forecasting 676
 - MG16.4 Autoregressive Modeling for Trend Fitting and Forecasting 677
 - MG16.5 Choosing an Appropriate Forecasting Model 677
 - MG16.6 Time-Series Forecasting of Seasonal Data 677

17 Business Analytics 678

- USING STATISTICS:** Back to Arlingtons for the Future 678
- 17.1** Business Analytics Categories 679
 - Inferential Statistics and Predictive Analytics 680
 - Supervised and Unsupervised Methods 680
- CONSIDER THIS:** What's My Major if I Want to be a Data Miner? 681

- 17.2** Descriptive Analytics 682
 - Dashboards 682
 - Data Dimensionality and Descriptive Analytics 683
- 17.3** Predictive Analytics for Prediction 684
- 17.4** Predictive Analytics for Classification 687
- 17.5** Predictive Analytics for Clustering 688
- 17.6** Predictive Analytics for Association 691
 - Multidimensional scaling (MDS) 692
- 17.7** Text Analytics 693
- 17.8** Prescriptive Analytics 694
- USING STATISTICS:** Back to Arlingtons..., Revisited 695

REFERENCES 695**KEY EQUATIONS 696****KEY TERMS 696****CHECKING YOUR UNDERSTANDING 696****CHAPTER REVIEW PROBLEMS 696****CASES FOR CHAPTER 17 698**

- The Mountain States Potato Company 698
- The Craybill Instrumentation Company 698

CHAPTER 17 SOFTWARE GUIDE 699

- Introduction 699
- SG17.2 Descriptive Analytics 699
- SG17.3 Predictive Analytics for Prediction 701
- SG17.4 Predictive Analytics for Classification 701
- SG17.5 Predictive Analytics for Clustering 702
- SG17.6 Predictive Analytics for Association 702

18 Getting Ready to Analyze Data in the Future 704

USING STATISTICS: Mounting Future Analyses 704

- 18.1** Analyzing Numerical Variables 705
 - EXHIBIT:** Questions to Ask When Analyzing Numerical Variables 705
 - Describe the Characteristics of a Numerical Variable? 705
 - Reach Conclusions About the Population Mean or the Standard Deviation? 705
 - Determine Whether the Mean and/or Standard Deviation Differs Depending on the Group? 706
 - Determine Which Factors Affect the Value of a Variable? 706
 - Predict the Value of a Variable Based on the Values of Other Variables? 707
 - Classify or Associate Items 707
 - Determine Whether the Values of a Variable Are Stable Over Time? 707
- 18.2** Analyzing Categorical Variables 707
 - EXHIBIT:** Questions to Ask When Analyzing Categorical Variables 707
 - Describe the Proportion of Items of Interest in Each Category? 707
 - Reach Conclusions About the Proportion of Items of Interest? 708
 - Determine Whether the Proportion of Items of Interest Differs Depending on the Group? 708

- Predict the Proportion of Items of Interest Based on the Values of Other Variables? 708
- Classify or Associate Items 708
- Determine Whether the Proportion of Items of Interest Is Stable Over Time? 708

USING STATISTICS: The Future to Be Visited 709**CHAPTER REVIEW PROBLEMS 709**

19 Statistical Applications in Quality Management (online) 19-1

USING STATISTICS: Finding Quality at the Beachcomber 19-1

- 19.1** The Theory of Control Charts 19-2
- 19.2** Control Chart for the Proportion: The p Chart 19-4
- 19.3** The Red Bead Experiment: Understanding Process Variability 19-10
- 19.4** Control Chart for an Area of Opportunity: The c Chart 19-11
- 19.5** Control Charts for the Range and the Mean 19-15
 - The R Chart 19-15
 - The \bar{X} Chart 19-18
- 19.6** Process Capability 19-21
 - Customer Satisfaction and Specification Limits 19-21
 - Capability Indices 19-23
 - CPL , CPU , and C_{pk} 19-24
- 19.7** Total Quality Management 19-26
- 19.8** Six Sigma 19-27
 - The DMAIC Model 19-28
 - Roles in a Six Sigma Organization 19-29
 - Lean Six Sigma 19-29

USING STATISTICS: Finding Quality at the Beachcomber, Revisited 19-30**SUMMARY 19-30****REFERENCES 19-31****KEY EQUATIONS 19-31****KEY TERMS 19-32****CHAPTER REVIEW PROBLEMS 19-32****CASES FOR CHAPTER 19 19-34**

- The Harnswell Sewing Machine Company Case 19-34
- Managing Ashland Multicomm Services 19-37

CHAPTER 19 EXCEL GUIDE 19-38

- EG19.2 Control Chart for the Proportion: The p Chart 19-38
- EG19.4 Control Chart for an Area of Opportunity: The c Chart 19-39
- EG19.5 Control Charts for the Range and the Mean 19-40
- EG19.6 Process Capability 19-41

CHAPTER 19 JMP GUIDE 19-41

- JG19.2 Control Chart for the Proportion: The p Chart 19-41
- JG19.4 Control Chart for an Area of Opportunity: The c Chart 19-41
- JG19.5 Control Charts for the Range and the Mean 19-42
- JG19.6 Process Capability 19-42

CHAPTER 19 MINITAB GUIDE 19-42

- MG19.2 Control Chart for the Proportion: The p Chart 19-42
- MG19.4 Control Chart for an Area of Opportunity:
The c Chart 19-43
- MG19.5 Control Charts for the Range and the Mean 19-43
- MG19.6 Process Capability 19-43

20 Decision Making (online) 20-1**USING STATISTICS: Reliable Decision Making 20-1****20.1** Payoff Tables and Decision Trees 20-2**20.2** Criteria for Decision Making 20-6

- Maximax Payoff 20-6
- Maximin Payoff 20-7
- Expected Monetary Value 20-7
- Expected Opportunity Loss 20-9
- Return-to-Risk Ratio 20-11

20.3 Decision Making with Sample Information 20-16**20.4** Utility 20-21**CONSIDER THIS: Risky Business 20-22****USING STATISTICS: Reliable Decision-Making,
Revisited 20-22****SUMMARY 20-23****REFERENCES 20-23****KEY EQUATIONS 20-23****KEY TERMS 20-23****CHAPTER REVIEW PROBLEMS 20-23****CASES FOR CHAPTER 20 20-26**

- Digital Case 20-26

CHAPTER 20 EXCEL GUIDE 20-27

- EG20.1 Payoff Tables and Decision Trees 20-27
- EG20.2 Criteria for Decision Making 20-27

Appendices 711**A. Basic Math Concepts and Symbols 712**

- A.1 Operators 712
- A.2 Rules for Arithmetic Operations 712
- A.3 Rules for Algebra: Exponents and Square Roots 712
- A.4 Rules for Logarithms 713
- A.5 Summation Notation 714
- A.6 Greek Alphabet 717

B. Important Software Skills and Concepts 718

- B.1 Identifying the Software Version 718
- B.2 Formulas 718
- B.3 Excel Cell References 720
- B.4 Excel Worksheet Formatting 721
- B.5E Excel Chart Formatting 722
- B.5J JMP Chart Formatting 723
- B.5M Minitab Chart Formatting 724

- B.6 Creating Histograms for Discrete Probability
Distributions (Excel) 724

- B.7 Deleting the “Extra” Histogram Bar (Excel) 725

C. Online Resources 726

- C.1 About the Online Resources for This Book 726
- C.2 Data Files 726
- C.3 Files Integrated With Microsoft Excel 733
- C.4 Supplemental Files 733

D. Configuring Software 734

- D.1 Microsoft Excel Configuration 734
- D.2 JMP Configuration 736
- D.3 Minitab Configuration 736

E. Table 737

- E.1 Table of Random Numbers 737
- E.2 The Cumulative Standardized Normal
Distribution 739
- E.3 Critical Values of t 741
- E.4 Critical Values of χ^2 743
- E.5 Critical Values of F 744
- E.6 Lower and Upper Critical Values, T_1 , of the Wilcoxon
Rank Sum Test 748
- E.7 Critical Values of the Studentized Range, Q 749
- E.8 Critical Values, d_L and d_U , of the Durbin–Watson
Statistic, D (Critical Values Are One–Sided) 751
- E.9 Control Chart Factors 752
- E.10 The Standardized Normal Distribution 753

F. Useful Knowledge 754

- F.1 Keyboard Shortcuts 754
- F.2 Understanding the Nonstatistical Functions 754

G. Software FAQs 756

- G.1 Microsoft Excel FAQs 756
- G.2 PHStat FAQs 756
- G.3 JMP FAQs 757
- G.4 Minitab FAQs 757

H. All About PHStat 758

- H.1 What is PHStat? 758
- H.2 Obtaining and Setting Up PHStat 759
- H.3 Using PHStat 759
- H.4 PHStat Procedures, by Category 760

**Self-Test Solutions and Answers to
Selected Even-Numbered Problems 761****Index 793****Credits 805**

Preface

As business statistics evolves and becomes an increasingly important part of one's business education, how business statistics gets taught and what gets taught becomes all the more important.

We, the authors, think about these issues 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 Data, Analytics, and Statistics Instruction and Business (DASI) 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.

When writing for introductory business statistics students, five principles guide us.

Help students see the relevance of statistics to their own careers by using examples from the functional areas that may become their areas of specialization. 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 to students' futures and will serve them better than focusing on tedious manual 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 data analysis software. We integrate using Microsoft Excel, JMP, and Minitab into all statistics topics to illustrate how software can assist the business decision making process. (Using software in this way also supports our second point about emphasizing interpretation over calculation).

Provide clear instructions to students that facilitate their use of data analysis software. We believe that providing such instructions assists learning and minimizes the chance that the software will distract from the learning of statistical concepts.

What's New in This Edition?

This fourteenth edition of *Basic Business Statistics* features many passages rewritten in a more concise style that emphasize definitions as the foundation for understanding statistical concepts. In addition to changes that readers of past editions have come to expect, such as new examples and Using Statistics case scenarios and an extensive number of new end-of-section or end-of-chapter problems, the edition debuts:

- **A First Things First Chapter** that builds on the previous edition's novel Important Things to Learn First Chapter by using real-world examples to illustrate how developments such as the increasing use of business analytics and "big data" have made knowing and understanding statistics that much more critical. This chapter is available as complimentary online download, allowing students to get a head start on learning.

- **JMP Guides** that provide detailed, hands-on instructions for using JMP to illustrate the concepts that this book teaches. JMP provides a starting point for continuing studies in business statistics and business analytics and features visualizations that are easy to construct and that summarize data in innovative ways. The JMP Guides join the Excel and Minitab Guides, themselves updated to reflect the most recent editions of those programs.
- **Tabular Summaries** that state hypothesis test and regression example results along with the conclusions that those results support now appear in Chapters 9 through 15.
- **An All-New Business Analytics Chapter (Chapter 17)** that makes extensive use of JMP and Minitab to illustrate predictive analytics for prediction, classification, clustering, and association as well as explaining what text analytics does and how descriptive and prescriptive analytics relate to predictive analytics. This chapter benefits from the insights the coauthors have gained from teaching and lecturing on business analytics as well as research the coauthors have done for a companion title on business analytics forthcoming for Fall 2018.

Continuing Features that Readers Have Come to Expect

This edition of *Basic Business Statistics* continues to incorporate a number of distinctive features that has led to its wide adoption over the previous editions. Table 1 summarizes these carry-over features:

TABLE 1 Distinctive Features Continued in the Fourteenth Edition

Feature	Details
Using Statistics Business Scenarios	A Using Statistics scenario that highlights how statistics is used in a business functional area begins each chapter. Each scenario provides an applied context for learning in its chapter. End-of-chapter “Revisited” sections reinforces the statistical methods that a chapter discusses and apply those methods to the questions raised in the scenario. <i>In this edition, seven chapters have new or revised Using Statistics scenarios.</i>
Emphasis on Data Analysis and Interpretation of Results	<i>Basic Business Statistics</i> was among the first business statistics textbooks to focus on interpretation of the results of a statistical method and not on the mathematics of a method. This tradition continues, now supplemented by JMP results complimenting the Excel and Minitab results of recent prior editions.
Software Integration	Software instructions in this book feature chapter examples and were personally written by the authors, who collectively have over one hundred years experience teaching the application of software to business. Software usage also features templates and applications developed by the authors that minimize the frustration of using software while maximizing statistical learning
Opportunities for Additional Learning	Student Tips, LearnMore bubbles, and Consider This features extend student-paced learning by reinforcing important points or examining side issues or answering questions that arise while studying business statistics such as “What is so ‘normal’ about the normal distribution?”
Highly Tailorable Context	With an extensive library of separate online topics, sections, and even two full chapters, instructors can combine these materials and the opportunities for additional learning to meet their curricular needs.
Software Flexibility	With modularized software instructions, instructors and students can switch among Excel, Excel with PHStat, JMP, and Minitab as they use this book, taking advantage of the strengths of each program to enhance learning.

TABLE 1 Distinctive Features Continued in the Fourteenth Edition (*continued*)

Feature	Details
End-of-Section and End-of-Chapter Reinforcements	“Exhibits” summarize key processes throughout the book. “Key Terms” provides an index to the definitions of the important vocabulary of a chapter. “Learning the Basics” questions test the basic concepts of a chapter. “Applying the Concepts” problems test the learner’s ability to apply those problems to business problems. For the more quantitatively-minded, “Key Equations” list the boxed number equations that appear in a chapter.
Innovative Cases	End-of-chapter cases include a case that continues through many chapters as well as “Digital Cases” that require students to examine business documents and other information sources to sift through various claims and discover the data most relevant to a business case problem as well as common misuses of statistical information. (Instructional tips for these cases and solutions to the Digital Cases are included in the Instructor’s Solutions Manual.)
Answers to Even-Numbered Problems	An appendix provides additional self-study opportunities by provides answers to the “Self-Test” problems and most of the even-numbered problems in this book.
Unique Excel Integration	Many textbooks feature Microsoft Excel, but <i>Basic Business Statistics</i> comes from the authors who originated both the Excel Guide workbooks that illustrate model solutions, developed Visual Explorations that demonstrate selected basic concepts, and designed and implemented PHStat, the Pearson statistical add-in for Excel that places the focus on statistical learning. (See Appendix H for a complete summary of PHStat.)

Chapter-by-Chapter Changes Made for This Edition

Because the authors believe in continuous quality improvement, *every* chapter of *Basic Business Statistics* contains changes to enhance, update, or just freshen this book. Table 2 provides a chapter-by-chapter summary of these changes.

TABLE 2
Chapter-by-Chapter
Change Matrix

Chapter	Using Statistics Changed	JMP Guide	Problems Changed	Selected Chapter Changes
FTF	•	•	n.a.	Think Differently About Statistics Starting Point for Learning Statistics
1	•	•	40%	Data Cleaning Other Data Preprocessing Tasks
2		•	60%	Organizing a Mix of Variables Visualizing A Mix of Variables Filtering and Querying Data Reorganized categorical variables discussion. Expanded data visualization discussion. New samples of 379 retirement funds and 100 restaurant meal costs for examples.
3		•	50%	New samples of 379 retirement funds and 100 restaurant meal costs for examples. Updated NBA team values data set.

Chapter	Using Statistics Changed	JMP Guide	Problems Changed	Selected Chapter Changes
4		•	43%	Basic Probability Concepts rewritten. Bayes' theorem example moved online
5		•	60%	Section 5.1 and Binomial Distribution revised. Covariance of a Probability Distribution and The Hypergeometric Distribution moved online.
6	•	•	33%	Normal Distribution rewritten. The Exponential Distribution moved online.
7		•	47%	Sampling Distribution of the Proportion rewritten.
8		•	40%	Confidence Interval Estimate for the Mean revised. Revised "Managing Ashland MultiComm Services" continuing case.
9		•	20%	Chapter introduction revised. Section 9.1 rewritten. New Section 9.4 example.
10	•	•	45%	New Effect Size (online). Using Statistics scenario linked to Chapter 11 and 17. New paired <i>t</i> test and the difference between two proportions examples.
11	•	•	20%	New Using Statistics scenario data. The Randomized Block Design moved online.
12		•	42%	Extensive use of new tabular summaries. Revised "Managing Ashland MultiComm Services" continuing case.
13		•	46%	Chapter introduction revised. Section 13.2 revised.
14		•	30%	Section 14.1 revised. Section 14.3 reorganized and revised. New dummy variable example. Influence Analysis moved online.
15		•	37%	Using Transformations in Regression Models rewritten and expanded. Model Building rewritten
16	•	•	67%	Chapter introduction reorganized and revised. All-new chapter examples.
17	•		42%	All-new chapter. Predictive analytics discussion expanded. Uses JMP and Minitab extensively.
18			47%	

Serious About Writing Improvements

Ever review a textbook that reads the same as an edition from years ago? Or read a preface that claims writing improvements but offers no evidence? Among the writing improvements in this edition of *Basic Business Statistics*, the authors have turned to tabular summaries to guide readers to reaching conclusions and making decisions based on statistical information. The authors believe that this writing improvement, which appears in Chapters 9 through 15, not only adds clarity to the purpose of the statistical method being discussed but better illustrates the role of statistics in business decision-making processes. Judge for yourself using the sample from Chapter 10 Example 10.1.

Previously, part of the solution to Example 10.1 was presented as:

You do not reject the null hypothesis because $t_{STAT} = -1.6341 > -1.7341$. The p -value (as computed in Figure 10.5) is 0.0598. This p -value indicates that the probability that $t_{STAT} < -1.6341$ is equal to 0.0598. In other words, if the population means are equal, the probability that the sample mean delivery time for the local pizza restaurant is at least 2.18 minutes faster than the national chain is 0.0598. Because the p -value is greater than $\alpha = 0.05$, there is insufficient evidence to reject the null hypothesis. Based on these results, there is insufficient evidence for the local pizza restaurant to make the advertising claim that it has a faster delivery time.

In this edition, we present the equivalent solution (on page 357):

Table 10.4 summarizes the results of the pooled-variance t test for the pizza delivery data using the calculation above (*not shown in this sample*) and Figure 10.5 results. Based on the conclusions, local branch of the national chain and a local pizza restaurant have similar delivery times. Therefore, as part of the last step of the DCOVA framework, you and your friends exclude delivery time as a decision criteria when choosing from which store to order pizza.

TABLE 10.4

Pooled-variance t test summary for the delivery times for the two pizza restaurants

Result	Conclusions
The $t_{STAT} = -1.6341$ is greater than -1.7341 .	1. Do not reject the null hypothesis H_0 .
The t test p -value = 0.0598 is greater than the level of significance, $\alpha = 0.05$.	2. Conclude that insufficient evidence exists that the mean delivery time is lower for the local restaurant than for the branch of the national chain.
	3. There is a probability of 0.0598 that $t_{STAT} < -1.6341$.

A Note of Thanks

Creating a new edition of a textbook is a team effort, and we thank our Pearson Education editorial, marketing, and production teammates: Suzanna Bainbridge, Dana Bettez, Kaylee Carlson, Thomas Hayward, Deirdre Lynch, Stephanie Green, and Morgan Danna. Special thanks to the recently-retired Sherry Berg for her design and production oversight in helping to get this edition underway. (Her contributions will be missed!) And we would be remiss not to note the continuing work of Joe Vetere to prepare our screen shot illustrations and the efforts of Julie Kidd of SPi Global to ensure that this edition meets the highest standard of book production quality that is possible.

We also thank Alan Chesen of Wright State University for his diligence in being the accuracy checker for this edition and thank the following people whose comments helped us improve this edition: Mohammad Ahmadi, University of Tennessee-Chattanooga; Sung Ahn,

Washington State University; Kelly Alvey, Old Dominion University; Al Batten, University of Colorado-Colorado Springs; Alan Chesen, Wright State University; Gail Hafer, St. Louis Community College-Meramec; Chun Jin, Central Connecticut State University; Benjamin Lev, Drexel University; Lilian Prince, Kent State University; Bharatendra Rai, University of Massachusetts Dartmouth; Ahmad Vakil, St. John's University (NYC); and Shiro Withanachchi, Queens College (CUNY).

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*. Finally, we would like to thank our families for their patience, understanding, love, and assistance in making this book a reality.

Contact Us!

Please email us at authors@davidlevinestatistics.com or tweet us @[BusStatBooks](https://twitter.com/BusStatBooks) with your questions about the contents of this book. Please include the hashtag #BBS14 in your tweet or in the subject line of your email. We also welcome suggestions you may have for a future edition of this book. And while we have strived to make this book as error-free as possible, we also appreciate those who share with us any perceived problems or errors that they encounter.

If you need assistance using software, please contact your academic support person or Pearson Support at support.pearson.com/getsupport/. They have the resources to resolve and walk you through a solution to many technical issues in a way we do not.

As you use this book, be sure to make use of the "Resources for Success" that Pearson Education supplies for this book (described on the following pages). We also invite you to visit bbs14.davidlevinestatistics.com (bit.ly/2xwQoBT), where we may post additional information or new content as necessary.

Mark L. Berenson
David M. Levine
Kathryn A. Szabat
David F. Stephan

Get the Most Out of MyLab Statistics

MyLab™ Statistics is the leading online homework, tutorial, and assessment program for teaching and learning statistics, built around Pearson's best-selling content. MyLab Statistics helps students and instructors improve results; it provides engaging experiences and personalized learning for each student so learning can happen in any environment. Plus, it offers flexible and time-saving course management features to allow instructors to easily manage their classes while remaining in complete control, regardless of course format.

Preparedness

One of the biggest challenges in many mathematics and statistics courses is making sure students are adequately prepared with the prerequisite skills needed to successfully complete their course work. Pearson offers a variety of content and course options to support students with just-in-time remediation and key-concept review.

- Build homework assignments, quizzes, and tests to support your course learning outcomes. From *Getting Ready* (GR) questions to the *Conceptual Question Library* (CQL), we have your assessment needs covered from the mechanics to the critical understanding of Statistics. The exercise libraries include technology-led instruction, including new Excel-based exercises, and learning aids to reinforce your students' success.
- Using proven, field-tested technology, auto-graded Excel Projects allow instructors to seamlessly integrate Microsoft® Excel® content into their course without having to manually grade spreadsheets. Students have the opportunity to practice important statistical skills in Excel, helping them to master key concepts and gain proficiency with the program.

Used by more than 37 million students worldwide, MyLab Statistics delivers consistent, measurable gains in student learning outcomes, retention, and subsequent course success.

Resources for Success

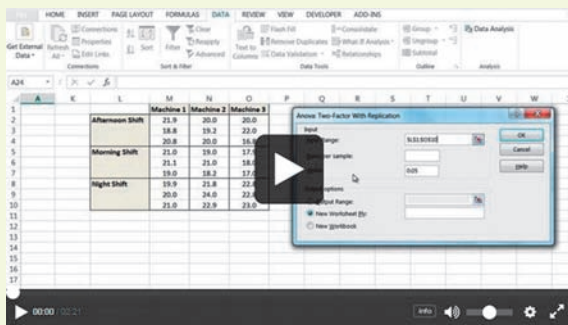
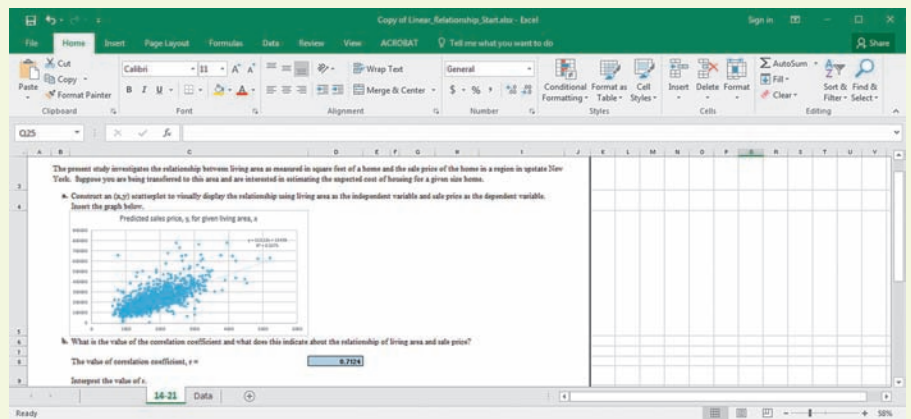
MyLab™ Statistics Online Course for Basic Business Statistics by Berenson/Levine/Szabat/Stephan (requires access code for use)

MyLab™ Statistics is available to accompany Pearson’s market leading text offerings. To give students a consistent tone, voice, and teaching method each text’s flavor and approach is tightly integrated throughout the accompanying MyLab Statistics course, making learning the material as seamless as possible.

New! Auto-Graded Excel Grader Projects

Using proven, field-tested technology, auto-graded Excel Projects allow instructors to seamlessly integrate Microsoft® Excel® content into their course without having to manually grade spreadsheets.

Students have the opportunity to practice important statistical skills in Excel, helping them to master key concepts and gain proficiency with the program.

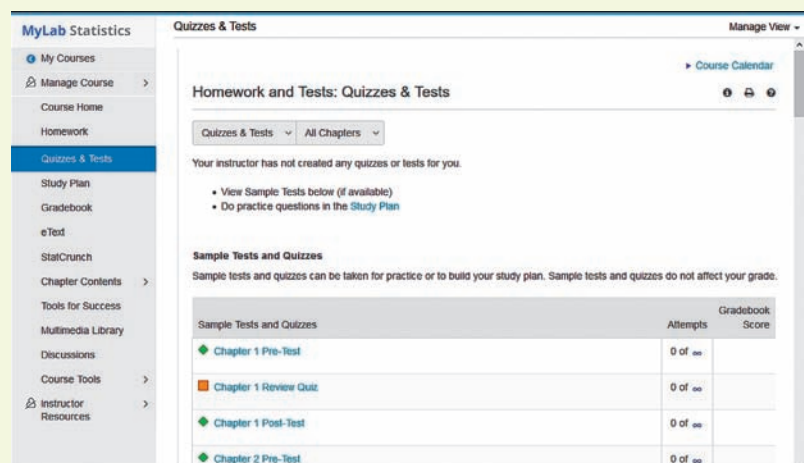


Tutorials and Study Cards for Statistical Software

Tutorials provide brief video walkthroughs and step-by-step instructional study cards on common statistical procedures such as confidence interval estimation, ANOVA, regression, and hypothesis testing. Tutorials and study cards are supplied for Excel 2013 and 2016, Excel with PHStat, JMP, and Minitab.

Diverse Question Libraries

Build homework assignments, quizzes, and tests to support your course learning outcomes. From Getting Ready (GR) questions to the Conceptual Question Library (CQL), we have your assessment needs covered from the mechanics to the critical understanding of Statistics. The exercise libraries include technology-led instruction, including new Excel-based exercises, and learning aids to reinforce your students’ success.



Resources for Success

Instructor Resources

Instructor's Solutions Manual, by Alan Chesen, Wright State University, presents solutions for end-of-section and end-of-chapter problems and answers to case questions, and provides teaching tips for each chapter. The Instructor's Solutions Manual is available for download at www.Pearson.com or in MyLab Statistics.

Lecture PowerPoint Presentations, by Patrick Schur, Miami University (Ohio), are available for each chapter. These presentations provide instructors with individual lecture notes to accompany the text. The slides include many of the figures and tables from the textbook. Instructors can use these lecture notes as is or customize them in Microsoft PowerPoint. The PowerPoint presentations are available for download at www.Pearson.com or in MyLab Statistics.

Test Bank, 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. The Test Bank is available for download at www.Pearson.com or in MyLab Statistics.

TestGen® (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.

Student Resources

Student's Solutions Manual, by Alan Chesen, Wright State University, provides detailed solutions to virtually all the even-numbered exercises and worked-out solutions to the self-test problems. (ISBN-13: 978-0-13-468504-5)

Online resources complement and extend the study of business statistics and support the content of this book. These resources include **data files** for in-chapter examples and problems, **templates and model solutions**, and **optional topics and chapters**. (See Appendix C for a complete description of the online resources.)

PHStat helps create Excel worksheet solutions to statistical problems. PHStat uses Excel building blocks to create worksheet solutions. These worksheet solutions illustrate Excel techniques and students can examine them to gain new Excel skills. Additionally, many solutions are what-if templates in which the effects of changing data on the results can be explored. Such templates are fully reusable on any computer on which Excel has been installed. PHStat requires an access code and separate download for use. PHStat access codes can be bundled with this textbook using ISBN-13: 978-0-13-468497-0.

Minitab® More than 4,000 colleges and universities worldwide use Minitab software to help students learn quickly and to provide them with a skill-set that's in demand in today's data-driven workforce. Minitab® includes a comprehensive collection of statistical tools to teach beginning through advanced courses. Bundling Minitab software ensures students have the software they need for the duration of their course work. (ISBN-10: 0-13-445640-8; ISBN-13: 978-0-13-445640-9)

JMP® Student Edition software is statistical discovery software from SAS Institute Inc., the leader in business analytics software and services. JMP® Student Edition is a streamlined version of JMP that provides all the statistics and graphics covered in introductory and intermediate statistics courses. Available for bundling with this textbook. (ISBN-10: 0-13-467979-2; ISBN-13: 978-0-13-467979-2)

First Things First



▼ USING STATISTICS “The Price of Admission”

It's the year 1900 and you are a promoter of theatrical productions, in the business of selling seats for individual performances. Using your knowledge and experience, you establish a selling price for the performances, a price you hope represents a good trade-off between maximizing revenues and avoiding driving away demand for your seats. You print up tickets and flyers, place advertisements in local media, and see what happens. After the event, you review your results and consider if you made a wise trade-off.

Tickets sold very quickly? Next time perhaps you can charge more. The event failed to sell out? Perhaps next time you could charge less or take out more advertisements to drive demand. If you lived over 100 years ago, that's about all you could do.

Jump ahead about 70 years. You're still a promoter but now using a computer system that allows your customers to buy tickets over the phone. You can get summary reports of advance sales for future events and adjust your advertising on radio and on TV and, perhaps, add or subtract performance dates using the information in those reports.

Jump ahead to today. You're still a promoter but you now have a fully computerized sales system that allows you to constantly adjust the price of tickets. You also can manage many more categories of tickets than just the near-stage and far-stage categories you might have used many years ago. You no longer have to wait until after an event to make decisions about changing your sales program. Through your sales system you have gained insights about your customers such as where they live, what other tickets they buy, and their appropriate demographic traits. Because you know more about your customers, you can make your advertising and publicity more efficient by aiming your messages at the types of people more likely to buy your tickets. By using social media networks and other online media, you can also learn almost immediately who is noticing and responding to your advertising messages. You might even run experiments online presenting your advertising in two different ways and seeing which way sells better.

Your current self has capabilities that allow you to be a more effective promoter than any older version of yourself. Just how much better? Turn the page.

CONTENTS

“The Price of Admission”

- FTF.1** Think Differently About Statistics
- FTF.2** Business Analytics: The Changing Face of Statistics
- FTF.3** Starting Point for Learning Statistics
- FTF.4** Starting Point for Using Software

EXCEL GUIDE

JMP GUIDE

MINITAB GUIDE

OBJECTIVES

- Statistics is a way of thinking that can lead to better decision making
- Statistics requires analytics skills and is an important part of your business education
- Recent developments such as the use of business analytics and “big data” have made knowing statistics even more critical
- The DCOVA framework guides your application of statistics
- The opportunity business analytics represents for business students

Now Appearing on Broadway ... and Everywhere Else

In early 2014, Disney Theatrical Productions woke up the rest of Broadway when reports revealed that its 17-year-old production of *The Lion King* had been the top-grossing Broadway show in 2013. How could such a long-running show, whose most expensive ticket was less than half the most expensive ticket on Broadway, earn so much while being so old? Over time, grosses for a show decline and, sure enough, weekly grosses for *The Lion King* had dropped about 25% by the year 2009. But, for 2013, grosses were up 67% from 2009 and weekly grosses for 2013 typically exceeded the grosses of opening weeks in 1997, adjusted for inflation!

Heavier advertising and some changes in ticket pricing helped, but the major reason for this change was something else: combining business acumen with the systematic application of *business statistics and analytics* to the problem of selling tickets. As a producer of the newest musical at the time said, “We make educated predictions on price. Disney, on the other hand, has turned this into a science” (see reference 3).

Disney had followed the plan of action that this book presents. It had collected its daily and weekly results, and summarized them, using techniques this book introduces in the next three chapters. Disney then analyzed those results by performing experiments and tests on the data collected (using techniques that later chapters introduce). In turn, those analyses were applied to a new interactive seating map that allowed customers to buy tickets for specific seats and permitted Disney to adjust the pricing of each seat for each performance. The whole system was constantly reviewed and refined, using the semiautomated methods to which Chapter 17 will introduce you. The end result was a system that outperformed the ticket-selling methods others used.

student TIP

From other business courses, you may recognize that Disney’s system uses dynamic pricing.

FTF.1 Think Differently About Statistics

The “Using Statistics” scenario suggests, and the Disney example illustrates, that modern-day information technology has allowed businesses to apply statistics in ways that could not be done years ago. This scenario and example reflect how this book teaches you about statistics. In these first two pages, you may notice

- the lack of calculation details and “math.”
- the emphasis on enhancing business methods and management decision making.
- that none of this seems like the content of a middle school or high school statistics class you may have taken.

You may have had some prior knowledge or instruction in *mathematical statistics*. This book discusses *business statistics*. While the boundary between the two can be blurry, business statistics emphasizes business problem solving and shows a preference for using software to perform calculations.

One similarity that you might notice between these first two pages and any prior instruction is *data*. **Data** are the facts about the world that one seeks to study and explore. Some data are unsummarized, such as the facts about a single ticket-selling transaction, whereas other facts, such as weekly ticket grosses, are **summarized**, derived from a set of unsummarized data. While you may think of data as being numbers, such as the cost of a ticket or the percentage that weekly grosses have increased in a year, do not overlook that data can be non-numerical as well, such as ticket-buyer’s name, seat location, or method of payment.

Statistics: A Way of Thinking

Statistics are the methods that allow you to work with data effectively. Business statistics focuses on interpreting the results of applying those methods. You interpret those results to help you enhance business processes and make better decisions. Specifically, business statistics provides you with a formal basis to summarize and visualize business data, reach conclusions about that data, make reliable predictions about business activities, and improve business processes.

You must apply this way of thinking correctly. Any “bad” things you may have heard about statistics, including the famous quote “there are lies, damned lies, and statistics” made famous by Mark Twain, speak to the errors that people make when either misusing statistical methods or mistaking statistics as a substitution for, and not an enhancement of, a decision-making process. (Disney Theatrical Productions’ success was based on *combining* statistics with business acumen, not *replacing* that acumen.)

DCOVA Framework To minimize errors, you use a framework that organizes the set of tasks that you follow to apply statistics properly. The five tasks that comprise the **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.

You must always do the **Define** and **Collect** tasks before doing the other three. The order of the other three varies and sometimes all three are done concurrently. In this book, you will learn more about the **Define** and **Collect** tasks in Chapter 1 and then be introduced to the **Organize** and **Visualize** tasks in Chapter 2. Beginning with Chapter 3, you will learn methods that help complete the **Analyze** task. Throughout this book, you will see specific examples that apply the DCOVA framework to specific business problems and examples.

Analytical Skills More Important than Arithmetic Skills The business preference for using software to automate statistical calculations maximizes the importance of having analytical skills while it minimizes arithmetic skills. With software, you perform calculations faster and more accurately than if you did those calculations by hand, minimizing the need for advanced arithmetic skills. However, with software you can *also* generate inappropriate or meaningless results if you have not fully understood a business problem or goal under study or if you use that software without a proper understanding of statistics.

Therefore, using software to create results that help solve business problems or meet business goals is *always* intertwined with using a framework. And using software does not mean memorizing long lists of software commands or how-to operations, but knowing how to review, modify, and possibly create software solutions. If you can analyze what you need to do and have a general sense of what you need, you can always find instructions or illustrative sample solutions to guide you. (This book provides detailed instructions *as well as* sample solutions for every statistical activity discussed in end-of-chapter software guides and through the use of various downloadable files and sample solutions.)

If you were introduced to using software in an application development setting or an introductory information systems class, do not mistake building applications from scratch as being a necessary skill. A “smart” smartphone user knows how to use apps such as Facebook, Instagram, YouTube, Google Maps, and Gmail effectively to communicate or discover and use information and has no idea how to construct a social media network, create a mapping system, or write an email program. Your approach to using the software in this book should be the same as that smart user. Use your analytical skills to focus on being an effective user and to understand *conceptually* what a statistical method or the software that implements that method does.

Statistics: An Important Part of Your Business Education

Until you read these pages, you may have seen a course in business statistics solely as a required course with little relevance to your overall business education. In just two pages, you have learned that statistics is a way of thinking that can help enhance your effectiveness in business—that is, applying statistics correctly is a fundamental, global skill in your business education.

In the current data-driven environment of business, you need the general analytical skills that allow you to work with data and interpret analytical results regardless of the discipline in which you work. No longer is statistics only for accounting, economics, finance, or other disciplines that directly work with numerical data. As the Disney example illustrates, the decisions you make will be increasingly based on data and not on your gut or intuition supported by past experience. Having a well-balanced mix of statistics, modeling, and basic technical skills as well as managerial skills, such as business acumen and problem-solving and communication skills, will best prepare you for the workplace today ... *and* tomorrow (see reference 1).

FTF.2 Business Analytics: The Changing Face of Statistics

Of the recent changes that have made statistics an important part of your business education, the emergence of the set of methods collectively known as business analytics may be the most significant change of all. **Business analytics** combine traditional statistical methods with methods from management science and information systems to form an interdisciplinary tool that supports fact-based decision making. Business analytics include

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

In the Disney Theatrical Productions example, statistical methods helped determine pricing factors, information systems methods made the interactive seating map and pricing analysis possible, and management science methods helped adjust pricing rules to match Disney's goal of sustaining ticket sales into the future. Other businesses use analytics to send custom mailings to their customers, and businesses such as the travel review site tripadvisor.com use analytics to help optimally price advertising as well as generate information that makes a persuasive case for using that advertising.

Generally, studies have shown that businesses that actively use business analytics and combine that use with data-guided management see increases in productivity, innovation, and competition (see reference 1). Chapter 17 introduces you to the statistical methods typically used in business analytics and shows how these methods are related to statistical methods that the book discusses in earlier chapters.

“Big Data”

Big data are collections of data that cannot be easily browsed or analyzed using traditional methods. Big data implies data that are being collected in huge volumes, at very fast rates or velocities (typically in near real time), and in a variety of forms that can differ from the structured forms such as records stored in files or rows of data stored in worksheets that businesses use every day. These attributes of volume, velocity, and variety (see reference 5) distinguish big data from a “big” (large) set of data that contains numerous records or rows of similar data. When combined with business analytics and the basic statistical methods discussed in this book, big data presents opportunities to gain new management insights and extract value from the data resources of a business (see reference 8).

Unstructured data Big data may also include **unstructured data**, data that has an irregular pattern and contain values which are not comprehensible without additional automated or manual interpretation. Unstructured data takes many forms such as unstructured text, pictures, videos, and audio tracks, with unstructured text, such as social media comments, getting the most immediate attention today for its possible use in customer, branding, or marketing analyses.

Unstructured data can be adapted for use with a number of methods, such as regression, which this book illustrates with conventional, structured files and worksheets. Unstructured data may require one to perform data collection and preparation tasks beyond those tasks that Chapter 1 discusses. While those tasks are beyond the scope of this book, Chapter 17 does include a small example that uses unstructured text to illustrate some of these differences one would face using unstructured data.

FTF.3 Starting Point for Learning Statistics

Statistics has its own vocabulary and learning the precise meanings, or **operational definitions**, of several basic terms provides a start to understanding the statistical methods that this book discusses. For example, *in statistics*, a **variable** defines a characteristic, or property, of an item or individual that can vary among the occurrences of those items or individuals. For example, for the item “book,” variables would include the title and number of chapters, as these facts can vary from book to book. For a given book, these variables have a specific value. For *this* book, the value of the title variable would be “Basic Business Statistics,” and “20” would be the value for the number of chapters variable. Note that a statistical variable is not an algebraic variable, which serves as a stand-in to represent one value in an algebraic statement and could never take a non-numerical value such as “Basic Business Statistics.”

Using the definition of variable, data, in its statistical sense, can be defined as the set of values associated with one or more variables. In statistics, each value for a specific variable is a single fact, not a list of facts. For example, what would be the value of the variable author for this book? Without this rule, you might say that the single list “Berenson, Levine, Szabat, Stephan” is the value. However, applying this rule, one would say that the variable has four separate values: “Berenson”, “Levine”, “Stephan”, and “Szabat”. This distinction of using only *single-value data* has the practical benefit of simplifying the task of entering data for software analysis.

Using the definitions of data and variable, the definition of statistics can be restated as the methods that analyze the data of the variables of interest. The methods that primarily help summarize and present data comprise **descriptive statistics**. Methods that use data collected from a small group to reach conclusions about a larger group comprise **inferential statistics**. Chapters 2 and 3 introduce descriptive methods, many of which are applied to support the inferential methods that the rest of the book presents.

Statistic

The previous section uses *statistics* in the sense of a collective noun, a noun that is the name for a collection of things (methods in this case). The word statistics also serves as the plural form of the noun statistic, as in “one uses methods of descriptive statistics (collective noun) to generate descriptive statistics (plural of the singular noun).” In this sense, a **statistic** refers to a value that summarizes the data of a particular variable. (More about this in coming chapters.) In the Disney Theatrical Productions example, the statement “for 2013, weekly grosses were up 67% from 2009” cites a statistic that summarizes the variable weekly grosses using the 2013 data—all 52 values.

When someone warns you of a possible unfortunate outcome by saying, “Don’t be a statistic!” you can always reply, “I can’t be.” *You* always represent one value and a *statistic* always summarizes multiple values. For the statistic “87% of our employees suffer a workplace accident,” you, as an employee, will either have suffered or have not suffered a workplace accident. The “have” or “have not” value contributes to the statistic but cannot be the statistic. A statistic can facilitate preliminary decision making. For example, would you immediately accept a position at a company if you learned that 87% of their employees suffered a workplace accident? (Sounds like this might be a dangerous place to work and that further investigation is necessary.)

Can Statistics (*pl.*, *statistic*) Lie?

The famous quote “lies, damned lies, and statistics” actually refers to the plural form of *statistic* and does not refer to statistics, the field of study. Can any statistic “lie”? No, faulty or invalid statistics can only be produced through willful misuse of statistics or when DCOVA framework tasks are done incorrectly. For example, many statistical methods are valid only if the data being analyzed have certain properties. To the extent possible, you test the assertion that the data have those properties, which in statistics are called *assumptions*. When an assumption is *violated*, shown to be invalid for the data being analyzed, the methods that require that assumption should not be used.

For the inferential methods that this book discusses in later chapters, you must always look for logical causality. **Logical causality** means that you can plausibly claim something directly causes something else. For example, you wear black shoes today and note that the weather is sunny. The next day, you again wear black shoes and notice that the weather continues to be sunny. The third day, you change to brown shoes and note that the weather is rainy. The fourth day, you wear black shoes again and the weather is again sunny. These four days seem to suggest a strong pattern between your shoe color choice and the type of weather you experience. You begin to think if you wear brown shoes on the fifth day, the weather will be rainy. Then you realize that your shoes cannot plausibly influence weather patterns, that your shoe color choice cannot *logically cause* the weather. What you are seeing is mere coincidence. (On the fifth day, you do wear brown shoes and it happens to rain, but that is just another coincidence.)

You can easily spot the lack of logical causality when trying to correlate shoe color choice with the weather, but in other situations the lack of logical causality may not be so easily seen. Therefore, relying on such correlations by themselves is a fundamental misuse of statistics. When you look for patterns in the data being analyzed, you must *always* be thinking of logical causes. Otherwise, you are misrepresenting your results. Such misrepresentations sometimes cause people to wrongly conclude that all statistics are “lies.” Statistics (*pl.*, *statistic*) are not lies or “damned lies.” They play a significant role in *statistics*, the way of thinking that can enhance your decision making and increase your effectiveness in business.

FTF.4 Starting Point for Using Software

Because software plays an important role in the application of business statistics, this book uses Excel, JMP, and Minitab to help explain and illustrate statistical concepts and methods. All three programs require knowledge of basic user interface skills, operations, and vocabulary that Table FTF.1 summarizes.

TABLE FTF.1
Basic Computing
Knowledge

Skill or Operation	Specifics
Identify and use standard window objects	Title bar, minimize/resize/close buttons, scroll bars, mouse pointer, menu bars or ribbons, dialog box, window subdivisions such as areas, panes, or child windows
Identify and use common dialog box items	Command button, list box, drop-down list, edit box, option button, check box, tabs (tabbed panels)
Mouse operations	Click, called select in some list or menu contexts and check or clear in some check box contexts; double-click; right-click to make a shortcut menu appear; drag and drag-and-drop

If you found anything new to you in this table, download and review a complimentary copy of the online pamphlet *Basic Computing Skills* and make its study your starting point. (Appendix C discusses how and from where you download online materials.)

Otherwise, a starting point with software begins with review of basic data and document operations. Excel, JMP, and Minitab all use **worksheets** to display the contents of a data set and as the means to enter or edit data. (JMP calls its worksheets **data tables**.) Worksheets are tabular arrangements of data, in which the intersections of rows and columns form **cells**, boxes into which you make individual entries. One places the data for a variable into the cells of a column such that each column contains the data for a different variable, if more than one variable is under study. By convention, one uses the cell in the initial row to enter names of the variables (variable columns). JMP and Minitab provide a special unnumbered row for entering variable names; in Excel, one must use row 1 for this purpose, which can sometimes lead to inadvertent errors. Figure FTF.1 shows the similarities and this key difference among the worksheets of the three programs.

FIGURE FTF.1

Minitab, JMP, and Excel worksheets

student TIP

Many of the Excel solutions as well as selected JMP and Minitab solutions that this book presents exist as templates that simplify the production of results and serve as models for learning more about using formulas in the three programs.

student TIP

Appendix D provides some technical information for add-ins appropriate for use with this book.

Generally, entries in each cell are single data values that can be text or numbers. All three programs also permit **formulas**, instructions to process data, to compute cell values. Formulas can include **functions** that simplify certain arithmetic tasks or provide access to advanced processing or statistical features. Formulas play an important role in designing **templates**, *reusable* solutions that have been previously audited and verified. However, JMP and Minitab allow only *column* formulas that define calculations for all the cells in a column, whereas Excel allows only *cell* formulas that define calculations for individual cells.

All three programs save worksheet data and results as one file, called a **workbook** in Excel and a **project** in JMP and Minitab. JMP and Minitab also allows the saving of individual worksheets or results as separate files, whereas Excel always saves a workbook even if the workbook contains (only) one worksheet. Both JMP and Minitab can open the data worksheets of an Excel workbook, making the Excel workbook a universal format for sharing of files that contain only data, such as the set of data files for use with this book that Appendix C documents. Table FTF.2 summarizes some of the various file formats that the three programs use.

Appendix B discusses the basic document operations of opening, saving, and printing documents, the specifics of which slightly differ among the three programs and further explains file formats as necessary.

TABLE FTF.2

Excel, JMP, and Minitab file formats

File Type	Excel	JMP	Minitab
All-in-one-file	.xlsx (workbook)	.jmpprj (project)	.mpj (project)
Single worksheet	.xlsx (see discussion)	.jmp	.mtw
Results only	n.a.	.jrp (report), .jmpappsource (dashboard)	.mgf (graph)
Macro or add-in (simplifies user operations)	.xlsm, .xlam	.jsl, .jmpaddin	.mtb, .mac

student TIP

Check the student download web page for this book for more information about PHStat and JMP and Minitab macros and add-ins that may be available for download.

Using Software Properly

Learning to use software *properly* can be hard as software has limited ways to provide feedback for user actions that are invalid operations. In addition, no software will ever know if you are following proper procedures for using that software. The principles that Exhibit FTF.1 list will assist you and should govern your use of software with this book. These principles will minimize your chance of making errors and lessen the frustration that often occurs when these principles are unknown or overlooked by a user.

EXHIBIT FTF.1**Principles of Using Software Properly**

Ensure that software is properly updated. Many users that manage their own computers often overlook the importance of ensuring that all installed software is up to date.

Understand the basic operational tasks. Take the time to master the tasks of starting the software, loading and entering data, and how to select or choose commands in a general way.

Understand the statistical concepts that a software procedure uses. Not understanding those concepts can cause you to make wrong choices in the software and can make interpreting software results difficult.

Know how to review software use for errors. Review and verify that the proper data preparation procedures (see Chapter 1) have been applied to the data before analysis. Verify that you have selected the correct procedures, commands, and software options. For any information that you entered for results labeling purposes, verify that no typographical errors exist.

Seek reuse of preexisting solutions to solve new problems. Build solutions from scratch only as necessary, particularly if using Excel in which errors can be most easily made. Some solutions, and almost all Excel solutions that this book presents, exist as models or templates that can *and should* be reused because such reuse models best practice.

Understand how to organize and present information from the results that the software creates. Think about the best ways to arrange and label your data. Consider ways to enhance or reorganize results that will facilitate communication with others.

Use self-identifying names, especially for the files that you create and save. Naming files Document 1, Document 2, and so on, will not help you later when you seek to retrieve a file for review and study.

In addition, also look for ways in which you can simplify the user interface of the software you use. If using Excel with this book, consider using PHStat, supplied separately or as part of a bundle by Pearson. PHStat simplifies the user interface by providing a consistent dialog box driven interface that minimizes keystrokes and mouse selections. If using JMP and Minitab, look for macro and add-ins that simplify command sequences or automate repetitive activities.

Software instruction conventions and notation The instructions that appear in the end-of-chapter software guides and certain appendices use a set of conventions and notation that Table FTF.3 summarizes. These conventions provide a concise and clear way of expressing specific user activities.

TABLE FTF.3 Conventions That This Book Uses

Convention	Example
Names of special keys appear capitalized and in boldface	Press Enter . Press Command or Ctrl .
Key combinations appear in boldface, with key names linked using this symbol: +	Enter the formula and press Ctrl+Enter . Press Ctrl+C .
Menu or Ribbon selections appear in boldface and sequences of consecutive selections are shown using this symbol: →	Select File → New Select PHStat → Descriptive Statistics → Boxplot .
Target of mouse operations appear in boldface	Click OK . Select Attendance and then click the Y button .
Entries and the location of where entries are made appear in boldface	Enter 450 in cell B5 . Add Temperature to the Construct Model Effects list.
Variables in data files that the text names appear capitalized	This file contains the Fund Type, Assets, and Expense Ratio variables.
Placeholders that express a general case appear in italics and may also appear in boldface as part of a function definition	AVERAGE (<i>cell range of variable</i>) Replace <i>cell range of variable</i> with the cell range that contains the Asset variable.
Names of data files that sections or problems refer to explicitly appear in a special font, but names of files in instructions appear in boldface	Retirement Funds Open the Retirement Funds workbook .
When current versions of Excel and Minitab differ in their user interface, alternate instructions for older versions appear in a second color immediately following the primary instructions	In the Select Data Source display, click the icon inside the Horizontal (Category) axis labels box. Click Edit under the Horizontal (Categories) Axis Labels heading.

▼ REFERENCES

1. Advani, D. "Preparing Students for the Jobs of the Future." *University Business* (2011), bit.ly/1gNLTJm.
2. Davenport, T., J. Harris, and R. Morison. *Analytics at Work*. Boston: Harvard Business School Press, 2010.
3. Healy, P. "Ticker Pricing Puts 'Lion King' atop Broadway's Circle of Life." *New York Times, New York edition*, March 17, 2014, p. A1, and nyti.ms.1zDkzki.
4. JP Morgan Chase. "Report of JPMorgan Chase & Co. Management Task Force Regarding 2012 CIO Losses," bit.ly/1BnQZzY, as quoted in J. Ewok, "The Importance of Excel," *The Baseline Scenario*, bit.ly/1LPeQUy.
5. Laney, D. *3D Data Management: Controlling Data Volume, Velocity, and Variety*. Stamford, CT: META Group. February 6, 2001.
6. Levine, D., and D. Stephan. "Teaching Introductory Business Statistics Using the DCOVA Framework." *Decision Sciences Journal of Innovative Education* 9 (Sept. 2011): 393–398.
7. Liberatore, M., and W. Luo. "The Analytics Movement." *Interfaces* 40 (2010): 313–324.
8. "What Is Big Data?" IBM Corporation, www.ibm.com/big-data/us/en/.

▼ KEY TERMS

big data	4	formula	7	statistics	2
cells	7	function	7	summarized data	2
data	2	inferential statistics	5	template	7
data table	7	logical causality	6	unstructured data	4
business analytics	4	operational definition	5	variable	5
DCOVA framework	3	project (JMP, Minitab)	7	workbook	7
descriptive statistics	5	statistic	5	worksheet	7