

GLOBAL
EDITION



Statistics for Managers Using Microsoft[®] Excel

EIGHTH EDITION

David M. Levine • David F. Stephan • Kathryn A. Szabat



A ROADMAP FOR SELECTING A STATISTICAL METHOD

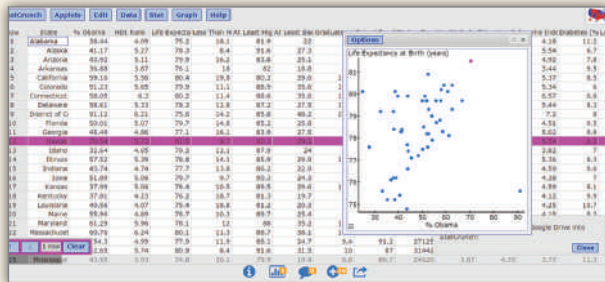
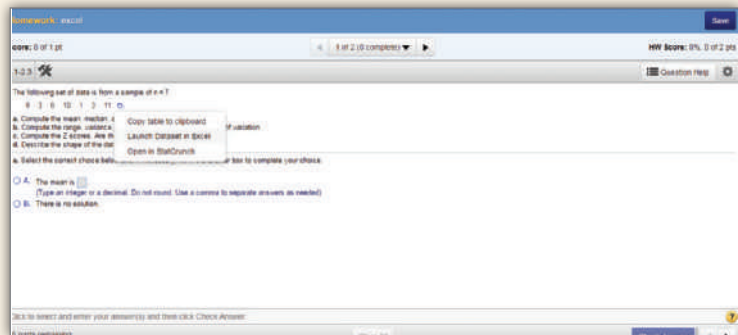
Data Analysis Task	For Numerical Variables	For Categorical Variables
Describing a group or several groups	<p>Ordered array, stem-and-leaf display, frequency distribution, relative frequency distribution, percentage distribution, cumulative percentage distribution, histogram, polygon, cumulative percentage polygon, sparklines, gauges, treemaps (Sections 2.2, 2.4, 2.6, 17.4)</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>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>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>Two-way analysis of variance (Section 11.2)</p> <p>Randomized block design (online Section 11.3)</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 2.6)</p>	<p>Contingency table, side-by-side bar chart, doughnut 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>Multiple regression (Chapters 14 and 15)</p> <p>Regression trees (Section 17.5)</p>	<p>Multidimensional contingency tables (Section 2.6)</p> <p>Drilldown and slicers (Section 2.6)</p> <p>Logistic regression (Section 14.7)</p> <p>Classification trees (Section 17.5)</p>

Available with MyStatLab™ for Your Business Statistics Courses

MyStatLab is the market-leading online learning management program for learning and teaching business statistics.

Statistical Software Support

Built-in tutorial videos and functionality make using the most popular software solutions seamless and intuitive. Tutorial videos, study cards, and manuals (for select titles) are available within MyStatLab and accessible at the point of use. Easily launch exercise and eText data sets into Excel or StatCrunch, or copy and paste into any other software program.



Leverage the Power of StatCrunch

MyStatLab leverages the power of StatCrunch –powerful, web-based statistical software. In addition, access to the full online community allows users to take advantage of a wide variety of resources and applications at www.statcrunch.com.

Bring Statistics to Life

Virtually flip coins, roll dice, draw cards, and interact with animations on your mobile device with the extensive menu of experiments and applets in StatCrunch. Offering a number of ways to practice resampling procedures, such as permutation tests and bootstrap confidence intervals, StatCrunch is a complete and modern solution.



Statistics for Managers Using Microsoft® Excel

8th Edition
Global Edition

David M. Levine

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

David F. Stephan

Two Bridges Instructional Technology

Kathryn A. Szabat

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



Pearson

Editorial Director: Chris Hoag
Editor in Chief: Deirdre Lynch
Acquisitions Editor: Suzanna Bainbridge
Senior Editorial Assistant: Justin Billing
Program Manager: Chere Bemelmans
Project Manager: Sherry Berg
Project Manager, Global Edition: Laura Perry
Associate Acquisitions Editor, Global Edition: Ananya Srivastava
Associate Project Editor, Global Edition: Paromita Banerjee
Manager, Media Production, Global Edition: M. Vikram Kumar
Manufacturing Controller, Production, Global Edition: Kay Holman
Program Management Team Lead: Karen Wernholm
Project Management Team Lead: Peter Silva
Media Producer: Jean Choe

TestGen Content Manager: John Flanagan
QA Manager, Assessment Content: Marty Wright
Senior MathXL Content Developer: Robert Carroll
Marketing Manager: Tiffany Bitzel
Field Marketing Manager: Adam Goldstein
Marketing Assistant: Jennifer Myers
Senior Author Support/Technology Specialist: Joe Vetere
Manager, Rights Management: Gina Cheselka
Senior Procurement Specialist: Carol Melville
Associate Director of Design: Andrea Nix
Program Design Lead and Cover Design: Barbara Atkinson
Text Design: Lumina Datamatics, Inc.
Production Coordination, Composition, and Illustrations: Lumina Datamatics
Cover Image: Shutterstock

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

MICROSOFT AND/OR ITS RESPECTIVE SUPPLIERS MAKE NO REPRESENTATIONS ABOUT THE SUITABILITY OF THE INFORMATION CONTAINED IN THE DOCUMENTS AND RELATED GRAPHICS PUBLISHED AS PART OF THE SERVICES FOR ANY PURPOSE. ALL SUCH DOCUMENTS AND RELATED GRAPHICS ARE PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND. MICROSOFT AND/OR ITS RESPECTIVE SUPPLIERS HEREBY DISCLAIM ALL WARRANTIES AND CONDITIONS WITH REGARD TO THIS INFORMATION, INCLUDING ALL WARRANTIES AND CONDITIONS OF MERCHANTABILITY, WHETHER EXPRESS, IMPLIED OR STATUTORY, FITNESS FOR A PARTICULAR PURPOSE, TITLE AND NON-INFRINGEMENT. IN NO EVENT SHALL MICROSOFT AND/OR ITS RESPECTIVE SUPPLIERS BE LIABLE FOR ANY SPECIAL, INDIRECT OR CONSEQUENTIAL DAMAGES OR ANY DAMAGES WHATSOEVER RESULTING FROM LOSS OF USE, DATA OR PROFITS, WHETHER IN AN ACTION OF CONTRACT, NEGLIGENCE OR OTHER TORTIOUS ACTION, ARISING OUT OF OR IN CONNECTION WITH THE USE OR PERFORMANCE OF INFORMATION AVAILABLE FROM THE SERVICES.

THE DOCUMENTS AND RELATED GRAPHICS CONTAINED HEREIN COULD INCLUDE TECHNICAL INACCURACIES OR TYPOGRAPHICAL ERRORS. CHANGES ARE PERIODICALLY ADDED TO THE INFORMATION HEREIN. MICROSOFT AND/OR ITS RESPECTIVE SUPPLIERS MAY MAKE IMPROVEMENTS AND/OR CHANGES IN THE PRODUCT(S) AND/OR THE PROGRAM(S) DESCRIBED HEREIN AT ANY TIME. PARTIAL SCREEN SHOTS MAY BE VIEWED IN FULL WITHIN THE SOFTWARE VERSION SPECIFIED.

Pearson Education Limited
Edinburgh Gate
Harlow
Essex CM20 2JE
England

and Associated Companies throughout the world

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

© Pearson Education Limited 2017

The rights of David M. Levine, David F. Stephan, and Kathryn A. Szabat to be identified as the authors of this work have been asserted by them in accordance with the Copyright, Designs and Patents Act 1988.

Authorized adaptation from the United States edition, entitled Statistics for Managers Using Microsoft Excel, 8th edition, ISBN 978-0-13-417305-4, by David M. Levine, David F. Stephan, and Kathryn A. Szabat, published by Pearson Education © 2017.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without either the prior written permission of the publisher or a license permitting restricted copying in the United Kingdom issued by the Copyright Licensing Agency Ltd, Saffron House, 6–10 Kirby Street, London EC1N 8TS.

All trademarks used herein are the property of their respective owners. The use of any trademark in this text does not vest in the author or publisher any trademark ownership rights in such trademarks, nor does the use of such trademarks imply any affiliation with or endorsement of this book by such owners.

ISBN 10: 1-292-15634-1
ISBN 13: 978-1-292-15634-7

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

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

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

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

About the Authors



Kathryn Szabat, David Levine, and David Stephan

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

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

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

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

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 MSMESB. She received a B.S. from SUNY-Albany, an M.S. in statistics from the Wharton School of the University of Pennsylvania, and a Ph.D. degree in statistics, with a cognate in operations research, from the Wharton School of the University of Pennsylvania.

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

Brief Contents

Preface 17

First Things First 25

- 1 Defining and Collecting Data 36
- 2 Organizing and Visualizing Variables 56
- 3 Numerical Descriptive Measures 119
- 4 Basic Probability 165
- 5 Discrete Probability Distributions 190
- 6 The Normal Distribution and Other Continuous Distributions 213
- 7 Sampling Distributions 240
- 8 Confidence Interval Estimation 261
- 9 Fundamentals of Hypothesis Testing: One-Sample Tests 294
- 10 Two-Sample Tests 331
- 11 Analysis of Variance 372
- 12 Chi-Square Tests and Nonparametric Tests 410
- 13 Simple Linear Regression 451
- 14 Introduction to Multiple Regression 499
- 15 Multiple Regression Model Building 545
- 16 Time-Series Forecasting 577
- 17 Getting Ready To Analyze Data In The Future 622
- 18 Statistical Applications in Quality Management (*online*) 18-1
- 19 Decision Making (*online*) 19-1

Appendices A–G 637

Self-Test Solutions and Answers to Selected Even-Numbered Problems 685

Index 714

Credits 721

Contents

Preface 17

First Things First 25

USING STATISTICS: “The Price of Admission” 25

Now Appearing on Broadway . . . and Everywhere Else 26

FTF.1 Think Differently About Statistics 26

Statistics: A Way of Thinking 26

Analytical Skills More Important than Arithmetic Skills 27

Statistics: An Important Part of Your Business Education 27

FTF.2 Business Analytics: The Changing Face of Statistics 28

“Big Data” 28

Structured Versus Unstructured Data 28

FTF.3 Getting Started Learning Statistics 29

Statistic 29

Can Statistics (*pl.*, Statistic) Lie? 30

FTF.4 Preparing to Use Microsoft Excel for Statistics 30

Reusability Through Recalculation 31

Practical Matters: Skills You Need 31

Ways of Working with Excel 31

Excel Guides 32

Which Excel Version to Use? 32

Conventions Used 32

REFERENCES 33

KEY TERMS 33

EXCEL GUIDE 34

EG.1 Entering Data 34

EG.2 Reviewing Worksheets 34

EG.3 If You Plan to Use the *Workbook* Instructions 35

1 Defining and Collecting Data 36

USING STATISTICS: Defining Moments 36

1.1 Defining Variables 37

Classifying Variables by Type 38

Measurement Scales 38

1.2 Collecting Data 39

Populations and Samples 40

Data Sources 40

1.3 Types of Sampling Methods 41

Simple Random Sample 42

Systematic Sample 42

Stratified Sample 43

Cluster Sample 43

1.4 Data Preparation 44

Data Cleaning 44

Data Formatting 45

Stacked and Unstacked Variables 45

Recoding Variables 46

1.5 Types of Survey Errors 47

Coverage Error 47

Nonresponse Error 47

Sampling Error 47

Measurement Error 48

Ethical Issues About Surveys 48

CONSIDER THIS: New Media Surveys/Old Survey Errors 48

USING STATISTICS: Defining Moments, Revisited 50

SUMMARY 50

REFERENCES 50

KEY TERMS 50

CHECKING YOUR UNDERSTANDING 51

CHAPTER REVIEW PROBLEMS 51

CASES FOR CHAPTER 1 52

Managing Ashland MultiComm Services 52

CardioGood Fitness 52

Clear Mountain State Student Survey 53

Learning with the Digital Cases 53

CHAPTER 1 EXCEL GUIDE 54

EG1.1 Defining Variables 54

EG1.2 Collecting Data 54

EG1.3 Types of Sampling Methods 55

EG1.4 Data Preparation 55

2 Organizing and Visualizing Variables 56

USING STATISTICS: “The Choice Is Yours” 56

2.1 Organizing Categorical Variables 57

The Summary Table 57

The Contingency Table 58

2.2 Organizing Numerical Variables 61

The Frequency Distribution 62

Classes and Excel Bins 64

The Relative Frequency Distribution and the Percentage Distribution 65

The Cumulative Distribution 67

2.3 Visualizing Categorical Variables 70

The Bar Chart 70

The Pie Chart and the Doughnut Chart 71

- The Pareto Chart 72
- Visualizing Two Categorical Variables 74
- 2.4** Visualizing Numerical Variables 76
 - The Stem-and-Leaf Display 77
 - The Histogram 78
 - The Percentage Polygon 79
 - The Cumulative Percentage Polygon (Ogive) 80
- 2.5** Visualizing Two Numerical Variables 83
 - The Scatter Plot 83
 - The Time-Series Plot 85
- 2.6** Organizing and Visualizing a Mix of Variables 87
 - Multidimensional Contingency Table 87
 - Adding a Numerical Variable to a Multidimensional Contingency Table 88
 - Drill Down 88
 - Excel Slicers 89
 - PivotChart 90
 - Sparklines 90
- 2.7** The Challenge in Organizing and Visualizing Variables 92
 - Obscuring Data 92
 - Creating False Impressions 93
 - Chartjunk 94
 - EXHIBIT: Best Practices for Creating Visualizations 96

USING STATISTICS: The Choice Is Yours, Revisited 97

SUMMARY 97

REFERENCES 98

KEY EQUATIONS 98

KEY TERMS 99

CHECKING YOUR UNDERSTANDING 99

CHAPTER REVIEW PROBLEMS 99

CASES FOR CHAPTER 2 104

- Managing Ashland MultiComm Services 104
- Digital Case 104
- CardioGood Fitness 105
- The Choice Is Yours Follow-Up 105
- Clear Mountain State Student Survey 105

CHAPTER 2 EXCEL GUIDE 106

- EG2.1 Organizing Categorical Variables 106
- EG2.2 Organizing Numerical Variables 108
- EG2.3 Visualizing Categorical Variables 110
- EG2.4 Visualizing Numerical Variables 112
- EG2.5 Visualizing Two Numerical Variables 116
- EG2.6 Organizing and Visualizing a Set of Variables 116

3 Numerical Descriptive Measures 119

USING STATISTICS: More Descriptive Choices 119

- 3.1** Central Tendency 120
 - The Mean 120
 - The Median 122
 - The Mode 123
 - The Geometric Mean 124
- 3.2** Variation and Shape 125
 - The Range 125

- The Variance and the Standard Deviation 126
- EXHIBIT: Manually Calculating the Sample Variance, S^2 , and Sample Standard Deviation, S 127
- The Coefficient of Variation 129
- Z Scores 130
- Shape: Skewness 132
- Shape: Kurtosis 132

3.3 Exploring Numerical Data 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 149

3.6 Statistics: Pitfalls and Ethical Issues 154

USING STATISTICS: More Descriptive Choices, Revisited 154

SUMMARY 154

REFERENCES 155

KEY EQUATIONS 155

KEY TERMS 156

CHECKING YOUR UNDERSTANDING 156

CHAPTER REVIEW PROBLEMS 157

CASES FOR CHAPTER 3 160

- Managing Ashland MultiComm Services 160
- Digital Case 160
- CardioGood Fitness 160
- More Descriptive Choices Follow-up 160
- Clear Mountain State Student Survey 160

CHAPTER 3 EXCEL GUIDE 161

- EG3.1 Central Tendency 161
- EG3.2 Variation and Shape 162
- EG3.3 Exploring Numerical Data 162
- EG3.4 Numerical Descriptive Measures for a Population 163
- EG3.5 The Covariance and the Coefficient of Correlation 163

4 Basic Probability 165

USING STATISTICS: Possibilities at M&R Electronics World 165

- 4.1** Basic Probability Concepts 166
 - Events and Sample Spaces 167
 - Contingency Tables 169
 - Simple Probability 169
 - Joint Probability 170
 - Marginal Probability 171
 - General Addition Rule 171

- 4.2** Conditional Probability 175
 Computing Conditional Probabilities 175
 Decision Trees 176
 Independence 178
 Multiplication Rules 179
 Marginal Probability Using the General Multiplication Rule 180

- 4.3** Ethical Issues and Probability 182

- 4.4** Bayes' Theorem 183

CONSIDER THIS: Divine Providence and Spam 183

- 4.5** Counting Rules 184

USING STATISTICS: Possibilities at M&R Electronics World, Revisited 185

SUMMARY 185

REFERENCES 185

KEY EQUATIONS 185

KEY TERMS 186

CHECKING YOUR UNDERSTANDING 186

CHAPTER REVIEW PROBLEMS 186

CASES FOR CHAPTER 4 188

- Digital Case 188
- CardioGood Fitness 188
- The Choice Is Yours Follow-Up 188
- Clear Mountain State Student Survey 188

CHAPTER 4 EXCEL GUIDE 189

- EG4.1 Basic Probability Concepts 189
- EG4.4 Bayes' Theorem 189

5 Discrete Probability Distributions 190

USING STATISTICS: Events of Interest at Ricknel Home Centers 190

- 5.1** The Probability Distribution for a Discrete Variable 191
 Expected Value of a Discrete Variable 191
 Variance and Standard Deviation of a Discrete Variable 192

- 5.2** Binomial Distribution 195

- 5.3** Poisson Distribution 202

- 5.4** Covariance of a Probability Distribution and its Application in Finance 205

- 5.5** Hypergeometric Distribution 206

USING STATISTICS: Events of Interest..., Revisited 206

SUMMARY 206

REFERENCES 206

KEY EQUATIONS 206

KEY TERMS 207

CHECKING YOUR UNDERSTANDING 207

CHAPTER REVIEW PROBLEMS 207

CASES FOR CHAPTER 5 209

- Managing Ashland MultiComm Services 209
- Digital Case 210

CHAPTER 5 EXCEL GUIDE 211

- EG5.1 The Probability Distribution for a Discrete Variable 211

EG5.2 Binomial Distribution 211

EG5.3 Poisson Distribution 212

6 The Normal Distribution and Other Continuous Distributions 213

USING STATISTICS: Normal Load Times at MyTVLab 213

- 6.1** Continuous Probability Distributions 214

- 6.2** The Normal Distribution 215

EXHIBIT: Normal Distribution Important Theoretical Properties 215

Computing Normal Probabilities 216

VISUAL EXPLORATIONS: Exploring the Normal Distribution 222

Finding X Values 222

CONSIDER THIS: What Is Normal? 226

- 6.3** Evaluating Normality 227

Comparing Data Characteristics to Theoretical Properties 228

Constructing the Normal Probability Plot 229

- 6.4** The Uniform Distribution 231

- 6.5** The Exponential Distribution 233

- 6.6** The Normal Approximation to the Binomial Distribution 233

USING STATISTICS: Normal Load Times..., Revisited 234

SUMMARY 234

REFERENCES 234

KEY EQUATIONS 235

KEY TERMS 235

CHECKING YOUR UNDERSTANDING 235

CHAPTER REVIEW PROBLEMS 235

CASES FOR CHAPTER 6 237

- Managing Ashland MultiComm Services 237
- CardioGood Fitness 237
- More Descriptive Choices Follow-up 237
- Clear Mountain State Student Survey 237
- Digital Case 237

CHAPTER 6 EXCEL GUIDE 238

- EG6.1 Continuous Probability Distributions 238
- EG6.2 The Normal Distribution 238
- EG6.3 Evaluating Normality 238

7 Sampling Distributions 240

USING STATISTICS: Sampling Oxford Cereals 240

- 7.1** Sampling Distributions 241

- 7.2** Sampling Distribution of the Mean 241

The Unbiased Property of the Sample Mean 241

Standard Error of the Mean 243

Sampling from Normally Distributed Populations 244

Sampling from Non-normally Distributed Populations—The Central Limit Theorem 247

EXHIBIT: Normality and the Sampling Distribution of the Mean 248

VISUAL EXPLORATIONS: Exploring Sampling Distributions 251

7.3 Sampling Distribution of the Proportion 252

USING STATISTICS: Sampling Oxford Cereals, Revisited 255

SUMMARY 256

REFERENCES 256

KEY EQUATIONS 256

KEY TERMS 256

CHECKING YOUR UNDERSTANDING 257

CHAPTER REVIEW PROBLEMS 257

CASES FOR CHAPTER 7 259

Managing Ashland Multicomm Services 259

Digital Case 259

CHAPTER 7 EXCEL GUIDE 260

EG7.2 Sampling Distribution of the Mean 260

8 Confidence Interval Estimation 261

USING STATISTICS: Getting Estimates at Ricknel Home Centers 261

8.1 Confidence Interval Estimate for the Mean (σ Known) 262
Can You Ever Know the Population Standard Deviation? 267

8.2 Confidence Interval Estimate for the Mean (σ Unknown) 268
Student's t Distribution 268
Properties of the t Distribution 269
The Concept of Degrees of Freedom 270
The Confidence Interval Statement 271

8.3 Confidence Interval Estimate for the Proportion 276

8.4 Determining Sample Size 279
Sample Size Determination for the Mean 279
Sample Size Determination for the Proportion 281

8.5 Confidence Interval Estimation and Ethical Issues 284

8.6 Application of Confidence Interval Estimation in Auditing 285

8.7 Estimation and Sample Size Estimation for Finite Populations 285

8.8 Bootstrapping 285

USING STATISTICS: Getting Estimates. . . , Revisited 285

SUMMARY 286

REFERENCES 286

KEY EQUATIONS 286

KEY TERMS 287

CHECKING YOUR UNDERSTANDING 287

CHAPTER REVIEW PROBLEMS 287

CASES FOR CHAPTER 8 290

Managing Ashland MultiComm Services 290

Digital Case 291

Sure Value Convenience Stores 291

CardioGood Fitness 291

More Descriptive Choices Follow-Up 291

Clear Mountain State Student Survey 291

CHAPTER 8 EXCEL GUIDE 292

EG8.1 Confidence Interval Estimate for the Mean (σ Known) 292

EG8.2 Confidence Interval Estimate for the Mean (σ Unknown) 292

EG8.3 Confidence Interval Estimate for the Proportion 293

EG8.4 Determining Sample Size 293

9 Fundamentals of Hypothesis Testing: One-Sample Tests 294

USING STATISTICS: Significant Testing at Oxford Cereals 294

9.1 Fundamentals of Hypothesis-Testing Methodology 295

The Null and Alternative Hypotheses 295

The Critical Value of the Test Statistic 296

Regions of Rejection and Nonrejection 297

Risks in Decision Making Using Hypothesis Testing 297

Z Test for the Mean (σ Known) 300

Hypothesis Testing Using the Critical Value Approach 300

EXHIBIT: The Critical Value Approach to Hypothesis Testing 301

Hypothesis Testing Using the p -Value Approach 303

EXHIBIT: The p -Value Approach to Hypothesis Testing 304

A Connection Between Confidence Interval Estimation and Hypothesis Testing 305

Can You Ever Know the Population Standard Deviation? 306

9.2 t Test of Hypothesis for the Mean (σ Unknown) 308

The Critical Value Approach 308

p -Value Approach 310

Checking the Normality Assumption 310

9.3 One-Tail Tests 314

The Critical Value Approach 314

The p -Value Approach 315

EXHIBIT: The Null and Alternative Hypotheses in One-Tail Tests 317

9.4 Z Test of Hypothesis for the Proportion 318

The Critical Value Approach 319

The p -Value Approach 320

9.5 Potential Hypothesis-Testing Pitfalls and Ethical Issues 322

EXHIBIT: Questions for the Planning Stage of Hypothesis Testing 322

Statistical Significance Versus Practical Significance 323

Statistical *Insignificance* Versus Importance 323

Reporting of Findings 323

Ethical Issues 323

9.6 Power of the Test 324

USING STATISTICS: Significant Testing. . . , Revisited 324

SUMMARY 324

REFERENCES 325

KEY EQUATIONS 325

KEY TERMS 325

CHECKING YOUR UNDERSTANDING 325

CHAPTER REVIEW PROBLEMS 326

CASES FOR CHAPTER 9 328

- Managing Ashland MultiComm Services 328
- Digital Case 328
- Sure Value Convenience Stores 328

CHAPTER 9 EXCEL GUIDE 329

- EG9.1 Fundamentals of Hypothesis-Testing Methodology 329
- EG9.2 t Test of Hypothesis for the Mean (σ Unknown) 329
- EG9.3 One-Tail Tests 330
- EG9.4 Z Test of Hypothesis for the Proportion 330

10 Two-Sample Tests 331**USING STATISTICS: Differing Means for Selling Streaming Media Players at Arlington's? 331**

- 10.1** Comparing the Means of Two Independent Populations 332
 - Pooled-Variance t Test for the Difference Between Two Means 332
 - Confidence Interval Estimate for the Difference Between Two Means 337
 - t Test for the Difference Between Two Means, Assuming Unequal Variances 338

CONSIDER THIS: Do People Really Do This? 339

- 10.2** Comparing the Means of Two Related Populations 341
 - Paired t Test 342
 - Confidence Interval Estimate for the Mean Difference 347
- 10.3** Comparing the Proportions of Two Independent Populations 349
 - Z Test for the Difference Between Two Proportions 350
 - Confidence Interval Estimate for the Difference Between Two Proportions 354

10.4 F Test for the Ratio of Two Variances 356**10.5 Effect Size 360****USING STATISTICS: Differing Means for Selling. . . , Revisited 361****SUMMARY 361****REFERENCES 362****KEY EQUATIONS 362****KEY TERMS 363****CHECKING YOUR UNDERSTANDING 363****CHAPTER REVIEW PROBLEMS 363****CASES FOR CHAPTER 10 365**

- Managing Ashland MultiComm Services 365
- Digital Case 366
- Sure Value Convenience Stores 366
- CardioGood Fitness 366
- More Descriptive Choices Follow-Up 366
- Clear Mountain State Student Survey 366

CHAPTER 10 EXCEL GUIDE 367

- EG10.1 Comparing The Means of Two Independent Populations 367
- EG10.2 Comparing the Means of Two Related Populations 369
- EG10.3 Comparing the Proportions of Two Independent Populations 370
- EG10.4 F Test for the Ratio of Two Variances 371

11 Analysis of Variance 372**USING STATISTICS: The Means to Find Differences at Arlington's 372**

- 11.1** The Completely Randomized Design: One-Way ANOVA 373
 - Analyzing Variation in One-Way ANOVA 374
 - F Test for Differences Among More Than Two Means 376
 - One-Way ANOVA F Test Assumptions 380
 - Levene Test for Homogeneity of Variance 381
 - Multiple Comparisons: The Tukey-Kramer Procedure 382
 - The Analysis of Means (ANOM) 384

- 11.2** The Factorial Design: Two-Way ANOVA 387
 - Factor and Interaction Effects 388
 - Testing for Factor and Interaction Effects 390
 - Multiple Comparisons: The Tukey Procedure 393
 - Visualizing Interaction Effects: The Cell Means Plot 395
 - Interpreting Interaction Effects 395

11.3 The Randomized Block Design 399**11.4 Fixed Effects, Random Effects, and Mixed Effects Models 399****USING STATISTICS: The Means to Find Differences at Arlington's Revisited 399****SUMMARY 400****REFERENCES 400****KEY EQUATIONS 400****KEY TERMS 401****CHECKING YOUR UNDERSTANDING 402****CHAPTER REVIEW PROBLEMS 402****CASES FOR CHAPTER 11 404**

- Managing Ashland MultiComm Services 404
- PHASE 1 404
- PHASE 2 404
- Digital Case 405
- Sure Value Convenience Stores 405
- CardioGood Fitness 405
- More Descriptive Choices Follow-Up 405
- Clear Mountain State Student Survey 405

CHAPTER 11 EXCEL GUIDE 406

- EG11.1 The Completely Randomized Design: One-Way ANOVA 406
- EG11.2 The Factorial Design: Two-Way ANOVA 408

12 Chi-Square and Nonparametric Tests 410**USING STATISTICS: Avoiding Guesswork about Resort Guests 410**

- 12.1** Chi-Square Test for the Difference Between Two Proportions 411
- 12.2** Chi-Square Test for Differences Among More Than Two Proportions 418
 - The Marascuilo Procedure 421
 - The Analysis of Proportions (ANOP) 423
- 12.3** Chi-Square Test of Independence 424

- 12.4** Wilcoxon Rank Sum Test: A Nonparametric Method for Two Independent Populations 430
- 12.5** Kruskal-Wallis Rank Test: A Nonparametric Method for the One-Way ANOVA 436
 - Assumptions 439
- 12.6** McNemar Test for the Difference Between Two Proportions (Related Samples) 441
- 12.7** Chi-Square Test for the Variance or Standard Deviation 441

USING STATISTICS: Avoiding Guesswork. . . , Revisited 442

SUMMARY 442

REFERENCES 443

KEY EQUATIONS 443

KEY TERMS 444

CHECKING YOUR UNDERSTANDING 444

CHAPTER REVIEW PROBLEMS 444

CASES FOR CHAPTER 12 446

- Managing Ashland MultiComm Services 446
- PHASE 1 446
- PHASE 2 446
- Digital Case 447
- Sure Value Convenience Stores 447
- CardioGood Fitness 447
- More Descriptive Choices Follow-Up 447
- Clear Mountain State Student Survey 447

CHAPTER 12 EXCEL GUIDE 448

- EG12.1 Chi-Square Test for the Difference Between Two Proportions 448
- EG12.2 Chi-Square Test for Differences Among More Than Two Proportions 448
- EG12.3 Chi-Square Test of Independence 449
- EG12.4 Wilcoxon Rank Sum Test: a Nonparametric Method for Two Independent Populations 449
- EG12.5 Kruskal-Wallis Rank Test: a Nonparametric Method for the One-Way ANOVA 450

13 Simple Linear Regression 451

USING STATISTICS: Knowing Customers at Sunflowers Apparel 451

- 13.1** Types of Regression Models 452
 - Simple Linear Regression Models 453
- 13.2** Determining the Simple Linear Regression Equation 454
 - The Least-Squares Method 454
 - Predictions in Regression Analysis: Interpolation Versus Extrapolation 457
 - Computing the Y Intercept, b_0 and the Slope, b_1 457
 - VISUAL EXPLORATIONS: Exploring Simple Linear Regression Coefficients 460
- 13.3** Measures of Variation 462
 - Computing the Sum of Squares 462
 - The Coefficient of Determination 463
 - Standard Error of the Estimate 465
- 13.4** Assumptions of Regression 467
- 13.5** Residual Analysis 467
 - Evaluating the Assumptions 467

- 13.6** Measuring Autocorrelation: The Durbin-Watson Statistic 471

- Residual Plots to Detect Autocorrelation 471
- The Durbin-Watson Statistic 472

- 13.7** Inferences About the Slope and Correlation Coefficient 475

- t Test for the Slope 475
- F Test for the Slope 477
- Confidence Interval Estimate for the Slope 478
- t Test for the Correlation Coefficient 479

- 13.8** Estimation of Mean Values and Prediction of Individual Values 482

- The Confidence Interval Estimate for the Mean Response 482
- The Prediction Interval for an Individual Response 483

- 13.9** Potential Pitfalls in Regression 486

- EXHIBIT: Six Steps for Avoiding the Potential Pitfalls 486

USING STATISTICS: Knowing Customers. . . , Revisited 488

SUMMARY 488

REFERENCES 489

KEY EQUATIONS 490

KEY TERMS 491

CHECKING YOUR UNDERSTANDING 491

CHAPTER REVIEW PROBLEMS 491

CASES FOR CHAPTER 13 495

- Managing Ashland MultiComm Services 495
- Digital Case 495
- Brynne Packaging 495

CHAPTER 13 EXCEL GUIDE 496

- EG13.2 Determining the Simple Linear Regression Equation 496
- EG13.3 Measures of Variation 497
- EG13.4 Assumptions of Regression 497
- EG13.5 Residual Analysis 497
- EG13.6 Measuring Autocorrelation: The Durbin-Watson Statistic 498
- EG13.7 Inferences about the Slope and Correlation Coefficient 498
- EG13.8 Estimation of Mean Values and Prediction of Individual Values 498

14 Introduction to Multiple Regression 499

USING STATISTICS: The Multiple Effects of OmniPower Bars 499

- 14.1** Developing a Multiple Regression Model 500
 - Interpreting the Regression Coefficients 500
 - Predicting the Dependent Variable Y 503
- 14.2** r^2 , Adjusted r^2 , and the Overall F Test 505
 - Coefficient of Multiple Determination 505
 - Adjusted r^2 505
 - Test for the Significance of the Overall Multiple Regression Model 506
- 14.3** Residual Analysis for the Multiple Regression Model 508
- 14.4** Inferences Concerning the Population Regression Coefficients 510
 - Tests of Hypothesis 510
 - Confidence Interval Estimation 511
- 14.5** Testing Portions of the Multiple Regression Model 513
 - Coefficients of Partial Determination 517

- 14.6** Using Dummy Variables and Interaction Terms in Regression Models 519
Interactions 521
- 14.7** Logistic Regression 528
- USING STATISTICS:** The Multiple Effects . . . , Revisited 533
- SUMMARY** 533
- REFERENCES** 535
- KEY EQUATIONS** 535
- KEY TERMS** 536
- CHECKING YOUR UNDERSTANDING** 536
- CHAPTER REVIEW PROBLEMS** 536
- CASES FOR CHAPTER 14** 539

- Managing Ashland MultiComm Services 539
Digital Case 539

CHAPTER 14 EXCEL GUIDE 541

- EG14.1 Developing a Multiple Regression Model 541
- EG14.2 r^2 , Adjusted r^2 , and the Overall F Test 542
- EG14.3 Residual Analysis for the Multiple Regression Model 542
- EG14.4 Inferences Concerning the Population Regression Coefficients 543
- EG14.5 Testing Portions of the Multiple Regression Model 543
- EG14.6 Using Dummy Variables and Interaction Terms in Regression Models 543
- EG14.7 Logistic Regression 544

15 Multiple Regression Model Building 545

USING STATISTICS: Valuing Parsimony at WSTA-TV 545

- 15.1** Quadratic Regression Model 546
Finding the Regression Coefficients and Predicting Y 546
Testing for the Significance of the Quadratic Model 549
Testing the Quadratic Effect 549
The Coefficient of Multiple Determination 551
- 15.2** Using Transformations in Regression Models 553
The Square-Root Transformation 553
The Log Transformation 555
- 15.3** Collinearity 558
- 15.4** Model Building 559
The Stepwise Regression Approach to Model Building 561
The Best Subsets Approach to Model Building 562
Model Validation 565
EXHIBIT: Steps for Successful Model Building 566
- 15.5** Pitfalls in Multiple Regression and Ethical Issues 568
Pitfalls in Multiple Regression 568
Ethical Issues 568

USING STATISTICS: Valuing Parsimony..., Revisited 568

- SUMMARY** 569
- REFERENCES** 570
- KEY EQUATIONS** 570
- KEY TERMS** 570
- CHECKING YOUR UNDERSTANDING** 570
- CHAPTER REVIEW PROBLEMS** 570

CASES FOR CHAPTER 15 572

- The Mountain States Potato Company 572

- Sure Value Convenience Stores 573
Digital Case 573
The Craybill Instrumentation Company Case 573
More Descriptive Choices Follow-Up 574

CHAPTER 15 EXCEL GUIDE 575

- EG15.1 The Quadratic Regression Model 575
- EG15.2 Using Transformations In Regression Models 575
- EG15.3 Collinearity 576
- EG15.4 Model Building 576

16 Time-Series Forecasting 577

USING STATISTICS: Principled Forecasting 577

- 16.1** The Importance of Business Forecasting 578
- 16.2** Component Factors of Time-Series Models 578
- 16.3** Smoothing an Annual Time Series 579
Moving Averages 580
Exponential Smoothing 582
- 16.4** Least-Squares Trend Fitting and Forecasting 585
The Linear Trend Model 585
The Quadratic Trend Model 587
The Exponential Trend Model 588
Model Selection Using First, Second, and Percentage Differences 590
- 16.5** Autoregressive Modeling for Trend Fitting and Forecasting 595
Selecting an Appropriate Autoregressive Model 596
Determining the Appropriateness of a Selected Model 597
EXHIBIT: Autoregressive Modeling Steps 599
- 16.6** Choosing an Appropriate Forecasting Model 604
Performing a Residual Analysis 604
Measuring the Magnitude of the Residuals Through Squared or Absolute Differences 605
Using the Principle of Parsimony 605
A Comparison of Four Forecasting Methods 605
- 16.7** Time-Series Forecasting of Seasonal Data 607
Least-Squares Forecasting with Monthly or Quarterly Data 608
- 16.8** Index Numbers 613

CONSIDER THIS: Let the Model User Beware 613

USING STATISTICS: Principled Forecasting, Revisited 613

- SUMMARY** 614
- REFERENCES** 615
- KEY EQUATIONS** 615
- KEY TERMS** 616
- CHECKING YOUR UNDERSTANDING** 616
- CHAPTER REVIEW PROBLEMS** 616

CASES FOR CHAPTER 16 617

- Managing Ashland MultiComm Services 617
Digital Case 617

CHAPTER 16 EXCEL GUIDE 618

- EG16.3 Smoothing an Annual Time Series 618
- EG16.4 Least-Squares Trend Fitting and Forecasting 619
- EG16.5 Autoregressive Modeling for Trend Fitting and Forecasting 620
- EG16.6 Choosing an Appropriate Forecasting Model 620
- EG16.7 Time-Series Forecasting of Seasonal Data 621

17 Getting Ready to Analyze Data in the Future 622

USING STATISTICS: Mounting Future Analyses 622

17.1 Analyzing Numerical Variables 623

EXHIBIT: Questions to Ask When Analyzing Numerical Variables 623

Describe the Characteristics of a Numerical Variable? 623

Reach Conclusions about the Population Mean or the Standard Deviation? 623

Determine Whether the Mean and/or Standard Deviation Differs Depending on the Group? 624

Determine Which Factors Affect the Value of a Variable? 624
Predict the Value of a Variable Based on the Values of Other Variables? 625

Determine Whether the Values of a Variable Are Stable Over Time? 625

17.2 Analyzing Categorical Variables 625

EXHIBIT: Questions to Ask When Analyzing Categorical Variables 625

Describe the Proportion of Items of Interest in Each Category? 625

Reach Conclusions about the Proportion of Items of Interest? 625

Determine Whether the Proportion of Items of Interest Differs Depending on the Group? 626

Predict the Proportion of Items of Interest Based on the Values of Other Variables? 626

Determine Whether the Proportion of Items of Interest Is Stable Over Time? 626

USING STATISTICS: Back to Arlingtons for the Future 626

17.3 Introduction to Business Analytics 627

Data Mining 627

Power Pivot 627

17.4 Descriptive Analytics 628

Dashboards 629

Dashboard Elements 629

17.5 Predictive Analytics 630

Classification and Regression Trees 631

USING STATISTICS: The Future to be Visited 632

REFERENCES 632

CHAPTER REVIEW PROBLEMS 632

CHAPTER 17 EXCEL GUIDE 635

EG17.3 Introduction to Business Analytics 635

EG17.4 Descriptive Analytics 635

18 Statistical Applications in Quality Management (online) 18-1

USING STATISTICS: Finding Quality at the Beachcomber 18-1

18.1 The Theory of Control Charts 18-2

18.2 Control Chart for the Proportion: The p Chart 18-4

18.3 The Red Bead Experiment: Understanding Process Variability 18-10

18.4 Control Chart for an Area of Opportunity: The c Chart 18-12

18.5 Control Charts for the Range and the Mean 18-15

The R Chart 18-16

The \bar{X} Chart 18-18

18.6 Process Capability 18-21

Customer Satisfaction and Specification Limits 18-21

Capability Indices 18-23

CPL , CPU , and C_{pk} 18-24

18.7 Total Quality Management 18-26

18.8 Six Sigma 18-28

The DMAIC Model 18-29

Roles in a Six Sigma Organization 18-30

Lean Six Sigma 18-30

USING STATISTICS: Finding Quality at the Beachcomber, Revisited 18-31

SUMMARY 18-31

REFERENCES 18-32

KEY EQUATIONS 18-32

KEY TERMS 18-33

CHAPTER REVIEW PROBLEMS 18-34

CASES FOR CHAPTER 18 18-36

The Harnswell Sewing Machine Company Case 18-36

Managing Ashland Multicomm Services 18-38

CHAPTER 18 EXCEL GUIDE 18-39

EG18.1 The Theory of Control Charts 18-39

EG18.2 Control Chart for the Proportion: The p Chart 18-39

EG18.3 The Red Bead Experiment: Understanding Process Variability 18-40

EG18.4 Control Chart for an Area of Opportunity: The c Chart 18-40

EG18.5 Control Charts for the Range and the Mean 18-41

EG18.6 Process Capability 18-42

19 Decision Making (online) 19-1

USING STATISTICS: Reliable Decision Making 19-1

19.1 Payoff Tables and Decision Trees 19-2

19.2 Criteria for Decision Making 19-6

Maximax Payoff 19-6

Maximin Payoff 19-7

Expected Monetary Value 19-7

Expected Opportunity Loss 19-9

Return-to-Risk Ratio 19-11

19.3 Decision Making with Sample Information 19-16

19.4 Utility 19-21

CONSIDER THIS: Risky Business 19-22

USING STATISTICS: Reliable Decision-Making, Revisited 19-22

SUMMARY 19-23

REFERENCES 19-23

KEY EQUATIONS 19-23

KEY TERMS 19-23

CHAPTER REVIEW PROBLEMS 19-23

CASES FOR CHAPTER 19 19-26

Digital Case 19-26

CHAPTER 19 EXCEL GUIDE 19-27

EG19.1 Payoff Tables and Decision Trees 19-27

EG19.2 Criteria for Decision Making 19-27

Appendices 637

- A. Basic Math Concepts and Symbols 638
 - A.1 Rules for Arithmetic Operations 638
 - A.2 Rules for Algebra: Exponents and Square Roots 638
 - A.3 Rules for Logarithms 639
 - A.4 Summation Notation 640
 - A.5 Statistical Symbols 643
 - A.6 Greek Alphabet 643
- B Important Excel Skills and Concepts 644
 - B.1 Which Excel Do You Use? 644
 - B.2 Basic Operations 645
 - B.3 Formulas and Cell References 645
 - B.4 Entering a Formula 647
 - B.5 Formatting Cell Contents 648
 - B.6 Formatting Charts 649
 - B.7 Selecting Cell Ranges for Charts 650
 - B.8 Deleting the “Extra” Histogram Bar 651
 - B.9 Creating Histograms for Discrete Probability Distributions 651
- C. Online Resources 652
 - C.1 About the Online Resources for This Book 652
 - C.2 Accessing the Online Resources 652
 - C.3 Details of Online Resources 652
 - C.4 PHStat 659
- D. Configuring Microsoft Excel 660
 - D.1 Getting Microsoft Excel Ready for Use 660
 - D.2 Checking for the Presence of the Analysis ToolPak or Solver Add-Ins 660

- D.3 Configuring Microsoft Windows Excel Security Settings 660
- D.4 Opening Pearson-Supplied Add-Ins 661
- E. Tables 662
 - E.1 Table of Random Numbers 662
 - E.2 The Cumulative Standardized Normal Distribution 664
 - E.3 Critical Values of t 666
 - E.4 Critical Values of χ^2 668
 - E.5 Critical Values of F 669
 - E.6 Lower and Upper Critical Values, T_1 , of the Wilcoxon Rank Sum Test 673
 - E.7 Critical Values of the Studentized Range, Q 674
 - E.8 Critical Values, d_L and d_U , of the Durbin–Watson Statistic, D (Critical Values Are One-Sided) 676
 - E.9 Control Chart Factors 677
 - E.10 The Standardized Normal Distribution 678
- F. Useful Excel Knowledge 679
 - F.1 Useful Keyboard Shortcuts 679
 - F.2 Verifying Formulas and Worksheets 679
 - F.3 New Function Names 679
 - F.4 Understanding the Nonstatistical Functions 681
- G. Software FAQs 683
 - G.1 PHStat FAQs 683
 - G.2 Microsoft Excel FAQs 683

Self-Test Solutions and Answers to Selected Even-Numbered Problems 685**Index 714****Credits 721**

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 coauthors, 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 Making Statistics More Effective in Schools and Business (MSMESB) conferences. We use the ASA's Guidelines for Assessment and Instruction (GAISE) reports and combine them with our experiences teaching business statistics to a diverse student body at several universities. We also benefit from the interests and efforts of our past coauthors, Mark Berenson and Timothy Krehbiel.

Our Educational Philosophy

When writing for introductory business statistics students, 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 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 and Innovative in This Edition?

This eighth edition of *Statistics for Managers Using Microsoft Excel* contains these new and innovative features.

First Things First Chapter This new chapter provides an orientation that helps students start to understand the importance of business statistics and get ready to use Microsoft Excel even before they obtain a full copy of this book. Like its predecessor "Getting Started: Important Things to Learn First," this chapter has been developed and published to allow

distribution online even before a first class meeting. Instructors teaching online or hybrid course sections may find this to be a particularly valuable tool to get students thinking about business statistics and learning the necessary foundational concepts.

Getting Ready to Analyze Data in the Future This newly expanded version of Chapter 17 adds a second Using Statistics scenario that serves as an introduction to business analytics methods. That introduction, in turn, explains several advanced Excel features while familiarizing students with the fundamental concepts and vocabulary of business analytics. As such, the chapter provides students with a path for further growth and greater awareness about applying business statistics and analytics in their other courses and their business careers.

Expanded Excel Coverage *Workbook* instructions replace the *In-Depth Excel* instructions in the Excel Guides and discuss more fully OS X Excel (“Excel for Mac”) differences when they occur. Because the many current versions of Excel have varying capabilities, Appendix B begins by sorting through the possible confusion to ensure that students understand that not all Excel versions are alike.

In the Worksheet Notes that help explain the worksheet illustrations that in-chapter examples use as model solutions.

Many More Exhibits Stand-alone summaries of important procedures that serve as a review of chapter passages. Exhibits range from identifying best practices, such as “Best Practices for Creating Visualizations” in Chapter 2, to serving as guides to data analysis such as the pair of “Questions to Ask” exhibits in Chapter 17.

New Visual Design This edition uses a new visual design that better organizes chapter content and provides a more uncluttered, streamlined presentation.

Revised and Enhanced Content

This eighth edition of *Statistics for Managers Using Microsoft Excel* contains the following revised and enhanced content.

Revised End-of-Chapter Cases The Managing Ashland MultiComm Services case that reoccurs throughout the book has several new or updated cases. The Clear Mountain State Student Survey case, also recurring, uses new data collected from a survey of undergraduate students to practice and reinforce statistical methods learned in various chapters.

Many New Applied Examples and Problems Many of the applied examples throughout this book use new problems or revised data. Approximately 43% of the problems are new to this edition. Many of the new problems in the end-of-section and end-of-chapter problem sets contain data from *The Wall Street Journal*, *USA Today*, and other news media as well as from industry and marketing surveys from leading consultancies and market intelligence firms.

New or Revised Using Statistics Scenarios This edition contains six all-new and three revised Using Statistics scenarios. Several of the scenarios form a larger narrative when considered together even as they can all be used separately and singularly.

New “Getting Started Learning Statistics” and “Preparing to Use Microsoft Excel for Statistics” sections Included as part of the First Things First chapter, these new sections replace the “Making Best Use” section of the previous editions. The sections prepare students for learning with this book by discussing foundational statistics and Excel concepts together and explain the various ways students can work with Excel while learning business statistics with this book.

Revised Excel Appendices These appendices review the foundational skills for using Microsoft Excel, review the latest technical and relevant setup information, and discuss optional but useful knowledge about Excel.

Software FAQ Appendix This appendix provides answers to commonly-asked questions about PHStat and using Microsoft Excel and related software with this book.

Distinctive Features

This eighth edition of *Statistics for Managers Using Microsoft Excel* continues the use of the following distinctive features.

Using Statistics Business Scenarios Each chapter begins with a Using Statistics scenario, an example that highlights how statistics is used in a functional area of business such as finance, information systems, management, and marketing. Every chapter uses its scenario throughout to provide an applied context for learning concepts. Most chapters conclude with a Using Statistics, Revisited section that reinforces the statistical methods and applications that a chapter discusses.

Emphasis on Data Analysis and Interpretation of Excel Results Our focus emphasizes analyzing data by interpreting results while reducing emphasis on doing calculations. For example, in the coverage of tables and charts in Chapter 2, we help students interpret various charts and explain when to use each chart discussed. Our coverage of hypothesis testing in Chapters 9 through 12 and regression and multiple regression in Chapters 13–15 include extensive software results so that the p -value approach can be emphasized.

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

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

Digital Cases These cases ask students to examine interactive PDF documents to sift through various claims and information and discover the data most relevant to a business case scenario. In doing so, students determine whether the data support the conclusions and claims made by the characters in the case as well as learn how to identify 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 A special section at the end of this book provides answers to most of the even-numbered exercises of this book.

Flexibility Using Excel For almost every statistical method discussed, students can use Excel Guide model workbook solutions with the *Workbook* instructions or the *PHStat* instructions to produce the worksheet solutions that the book discusses and presents. And, whenever possible, the book provides *Analysis ToolPak* instructions to create similar solutions.

Extensive Support for Using Excel For readers using the *Workbook* instructions, this book explains operational differences among current Excel versions and provides alternate instructions when necessary.

PHStat PHStat is the Pearson Education Statistics add-in that makes operating Excel as distraction-free as possible. PHStat executes for you the low-level menu selection and worksheet entry tasks that are associated with Excel-based solutions. Students studying statistics can focus solely on mastering statistical concepts and not worry about having to become expert Excel users simultaneously.

PHStat creates the “live,” dynamic worksheets and chart sheets that match chapter illustrations and from which students can learn more about Excel. PHStat includes over 60 procedures including:

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

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

Sampling: sampling distributions simulation

Confidence interval estimation: for the mean, sigma unknown; for the mean, sigma known, for the population variance, for the proportion, and for the total difference

Sample size determination: for the mean and the proportion

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

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

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

Multiple-sample tests: chi-square test, Marascuilo procedure Kruskal-Wallis rank test, Levene test, one-way ANOVA, Tukey-Kramer procedure, randomized block design, and two-way ANOVA with replication

Regression: simple linear regression, multiple regression, best subsets, stepwise regression, and logistic regression

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

Decision-making: covariance and portfolio management, expected monetary value, expected opportunity loss, and opportunity loss

Data preparation: stack and unstack data

To learn more about PHStat, see Appendix C.

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

Chapter-by-Chapter Changes Made for This Edition

As authors, we take pride in updating the content of our chapters *and* our problem sets. Besides incorporating the new and innovative features that the previous section discusses, each chapter of the eighth edition of *Statistics for Managers Using Microsoft Excel* contains specific changes that refine and enhance our past editions as well as many new or revised problems.

The new **First Things First** chapter replaces the seventh edition's Let's Get Started chapter, keeping that chapter's strength while immediately drawing readers into the changing face of statistics and business analytics with a new opening Using Statistics scenario. And like the previous edition's opening chapter, Pearson Education openly posts this chapter so students can get started learning business statistics even before they obtain their textbooks.

Chapter 1 builds on the opening chapter with a new Using Statistics scenario that offers a cautionary tale about the importance of defining and collecting data. Rewritten Sections 1.1 ("Defining Variables") and 1.2 ("Collecting Data") use lessons from the scenario to underscore important points. Over one-third of the problems in this chapter are new or updated.

Chapter 2 features several new or updated data sets, including a new data set of 407 mutual funds that illustrate a number of descriptive methods. The chapter now discusses doughnut charts and sparklines and contains a reorganized section on organizing and visualizing a mix of variables. Section 2.7 (“The Challenge in Organizing and Visualizing Variables”) expands on previous editions’ discussions that focused solely on visualization issues. This chapter uses an updated Clear Mountain State student survey as well. Over half of the problems in this chapter are new or updated.

Chapter 3 also uses the new set of 407 mutual funds and uses new or updated data sets for almost all examples that the chapter presents. Updated data sets include the restaurant meal cost samples and the NBA values data. This chapter also uses an updated Clear Mountain State student survey. Just under one-half of the problems in this chapter are new or updated.

Chapter 4 uses an updated Using Statistics scenario while preserving the best features of this chapter. The chapter now starts a section on Bayes’ theorem which completes as an online section, and 43% of the problems in the chapter are new or updated.

Chapter 5 has been streamlined with the sections “Covariance of a Probability Distribution and Its Application in Finance” and “Hypergeometric Distribution” becoming online sections. Nearly 40% of the problems in this chapter are new or updated.

Chapter 6 features an updated Using Statistics scenario and the section “Exponential Distribution” has become an online section. This chapter also uses an updated Clear Mountain State student survey. Over one-third of the problems in this chapter are new or updated.

Chapter 7 now contains an additional example on sampling distributions from a larger population, and one-in-three problems are new or updated.

Chapter 8 has been revised to provide enhanced explanations of Excel worksheet solutions and contains a rewritten “Managing Ashland MultiComm Services” case. This chapter also uses an updated Clear Mountain State student survey, and new or updated problems comprise 39% of the problems.

Chapter 9 contains refreshed data for its examples and enhanced Excel coverage that provides greater details about the hypothesis test worksheets that the chapter uses. Over 40% of the problems in this chapter are new or updated.

Chapter 10 contains a new Using Statistics scenario that relates to sales of streaming video players and that connects to Using Statistics scenarios in Chapters 11 and 17. This chapter gains a new online section on effect size. The Clear Mountain State survey has been updated, and over 40% of the problems in this chapter are new or updated.

Chapter 11 expands on the Chapter 10 Using Statistics scenario that concerns the sales of mobile electronics. The Clear Mountain State survey has been updated. Over one-quarter of the problems in this chapter are new or updated.

Chapter 12 now incorporates material that was formerly part of the “Short Takes” for the chapter. The chapter also includes updated “Managing Ashland MultiComm Services” and Clear Mountain State student survey cases and 41% of the problems in this chapter are new or updated.

Chapter 13 features a brand new opening passage that better sets the stage for the discussion of regression that continues in subsequent chapters. Chapter 13 also features substantially revised and expanded Excel coverage that describes more fully the details of regression results worksheets. Nearly one-half of the problems in this chapter are new or updated.

Chapter 14 likewise contains expanded Excel coverage, with some Excel Guides sections completely rewritten. As with Chapter 13, nearly one-half of the problems in this chapter are new or updated.

Chapter 15 contains a revised opening passage, and the “Using Transformations with Regression Models” section has been greatly expanded with additional examples. Over 40% of the problems in this chapter are new or updated.

Chapter 16 contains updated chapter examples concerning movie attendance data and Cola-Cola Company and Wal-Mart Stores revenues. Two-thirds of the problems in this chapter are new or updated.

Chapter 17 has been retitled “Getting Ready to Analyze Data in the Future” and now includes sections on Business Analytics that return to issues that the First Things First Chapter scenario raises and that provide students with a path to future learning and application of business statistics. The chapter presents several Excel-based descriptive analytics techniques and illustrates how advanced statistical programs can work with worksheet data created in Excel. One-half of the problems in this chapter are new or updated.

A Note of Thanks

Creating a new edition of a textbook is a team effort, and we would like to thank our Pearson Education editorial, marketing, and production teammates: Suzanna Bainbridge, Chere Bemelmans, Sherry Berg, Tiffany Bitzel, Deirdre Lynch, Jean Choe, and Joe Vetere. We also thank our statistical readers and accuracy checkers James Lapp, Susan Herring, Dirk Tempelaar, Paul Lorzczak, Doug Cashing, and Stanley Seltzer for their diligence in checking our work and Nancy Kincade of Lumina Datamatics. We also thank the following people for their helpful comments that we have used to improve this new edition: Anusua Datta, Philadelphia University; Doug Dotterweich, East Tennessee State University; Gary Evans, Purdue University; Chris Maurer, University of Tampa; Bharatendra Rai, University of Massachusetts Dartmouth; Joseph Snider and Keith Stracher, Indiana Wesleyan University; Leonie Stone, SUNY Geneseo; and Patrick Thompson, University of Florida.

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.

Pearson would also like to thank Walid D. Al-Wagfi, Gulf University for Science and Technology; Håkan Carlqvist, Luleå University of Technology; Rosie Ching, Singapore Management University; Ahmed ElMelegy, American University in Dubai; Sanjay Nadkarni, The Emirates Academy of Hospitality Management; and Ralph Scheubrein, Baden-Wuerttemberg Cooperative State University, for their work on the Global Edition.

Contact Us!

Please email us at authors@davidlevinestatics.com or tweet us @BusStatBooks with your questions about the contents of this book. Please include the hashtag #SMUME8 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.

We are happy to answer all types of questions, but if you need assistance using Excel or PHStat, please contact your local support person or Pearson Technical Support at 247pearsoned.custhelp.com. They have the resources to resolve and walk you through a solution to many technical issues in a way we do not.

We invite you to visit us at smume8.davidlevinestatics.com (bit.ly/1I8Lv2K), where you will find additional information and support for this book that we furnish in addition to all the resources that Pearson Education offers you on our book’s behalf (see pages 23 and 24).

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

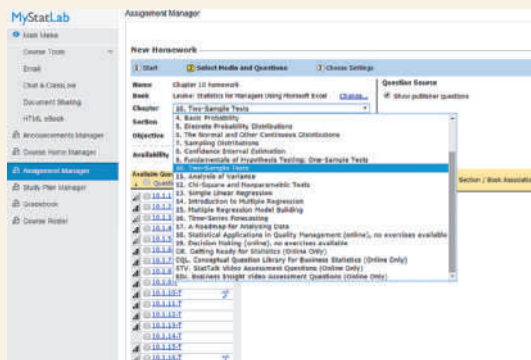
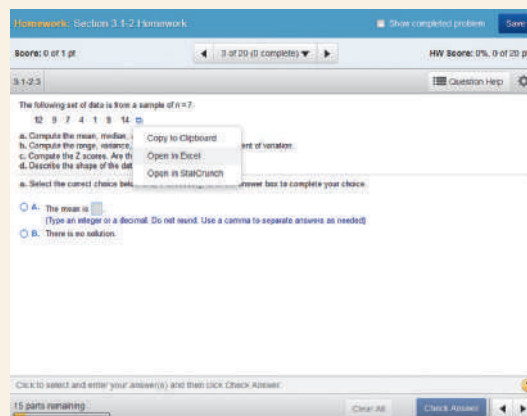
Resources for Success

MyStatLab™ Online Course for Statistics for Managers Using Microsoft® Excel by Levine/Stephan/Szabat (access code required)

MyStatLab 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 MyStatLab course, making learning the material as seamless as possible.

New! Launch Exercise Data in Excel

Students are now able to quickly and seamlessly launch data sets from exercises within MyStatLab into a Microsoft Excel spreadsheet for easy analysis. As always, students may also copy and paste exercise data sets into most other software programs.

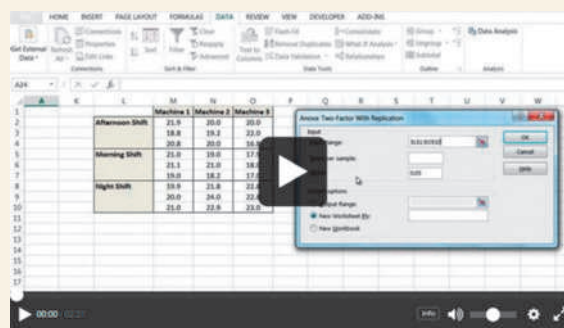


Diverse Question Libraries

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.

Technology Tutorials and Study Cards

Excel® tutorials provide brief video walkthroughs and step-by-step instructional study cards on common statistical procedures such as Confidence Intervals, ANOVA, Simple & Multiple Regression, and Hypothesis Testing. Tutorials will capture methods in Microsoft Windows Excel® 2010, 2013, and 2016 versions.



Resources for Success

Instructor Resources

Instructor's Solutions Manual, by Professor Pin Tian Ng of Northern Arizona University, includes solutions for end-of-section and end-of-chapter problems, answers to case questions, where applicable, and teaching tips for each chapter. The Instructor's Solutions Manual is available at the Instructor's Resource Center (www.pearsonglobaleditions.com/Levine) or in MyStatLab.

Lecture PowerPoint Presentations, by Professor Patrick Schur of Miami University (Ohio), are available for each chapter. The PowerPoint slides provide an instructor with individual lecture outlines to accompany the text. The slides include many of the figures and tables from the text. Instructors can use these lecture notes as is or can easily modify the notes to reflect specific presentation needs. The PowerPoint slides are available at the Instructor's Resource Center (www.pearsonglobaleditions.com/Levine) or in MyStatLab.

Test Bank, by Professor Pin Tian Ng of Northern Arizona University, contains true/false, multiple-choice, fill-in, and problem-solving questions based on the definitions, concepts, and ideas developed in each chapter of the text. New to this edition are specific test questions that use Excel datasets. The Test Bank is available at the Instructor's Resource Center (www.pearsonglobaleditions.com/Levine) or in MyStatLab.

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.

Online resources

The complete set of online resources are discussed fully in Appendix C. For adopting instructors, the following resources are among those available at the Instructor's Resource Center (www.pearsonglobaleditions.com/Levine) or in MyStatLab.

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

- FTF.1** Think Differently About Statistics
- FTF.2** Business Analytics: The Changing Face of Statistics
- FTF.3** Getting Started Learning Statistics
- FTF.4** Preparing to Use Microsoft Excel for Statistics

EXCEL GUIDE

- EG.1** Entering Data
- EG.2** Reviewing Worksheets
- EG.3** If You Plan to Use the *Workbook* Instructions

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.

studentTIP

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.)

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** provide one such framework.

DCOVA Framework

- **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

You have already read that business statistics shows a preference for using software to perform calculations. You can perform calculations *faster and more accurately* using software than you can if you performed those calculations by hand.

When you use software, you do more than just enter data. You need to review and modify, and possibly create, solutions. In Microsoft Excel, you use worksheet solutions that contain a mix of *organized* data and instructions that perform calculations on that data. Being able to review and modify worksheet solutions requires analytical skills more than arithmetic skills.

Allowing individuals to create new solutions from scratch in business can create risk. For example, in the aftermath of the 2012 “London Whale” trading debacle, JP Morgan Chase discovered a worksheet that could greatly miscalculate the volatility of a trading portfolio (see reference 4). To avoid this unnecessary risk, businesses prefer to use **templates**, *reusable* worksheet solutions that have been previously audited and verified.

When templates prove impractical, businesses seek to use *model worksheet solutions*. These solutions provide employees a basis for modification that is more extensive than changes one would make to a template. Whether you use the Excel Guide workbooks or PHStat with this book, you will reflect business practice by working with templates and model solutions as you use this book to learn statistics. You will not find many from-scratch construction tasks other than for the tasks of organizing and visualizing data in this book.

student TIP

Examining the structure of worksheet templates and models can also be helpful if learning more about Excel is one of your secondary learning goals.

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.

student TIP

Because you cannot “download” a big data collection, this book uses conventional structured (worksheet) files, both small and large, to demonstrate some of the principles and methods of business analytics in selected chapters, including Chapter 17, which introduce you to business analytics.

“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 other than the traditional structured forms such as data processing records, files, and tables and worksheets. These attributes of volume, velocity, and variety (see reference 5) distinguish big data from a set of data that contains a large number of similarly structured records or rows that you can place into a file or worksheet for browsing. In contrast, you cannot directly view big data; information system and statistical methods typically combine and summarize big data for you and then present the results of that processing.

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).

Structured Versus Unstructured Data

Statistics has traditionally used **structured data**, data that exist in repeating records or rows of similar format, such as the data found in the worksheet data files that this book describes in Appendix C. In contrast, **unstructured data** has very little or no repeating internal structure.

For example, to deeply analyze a group of companies, you might collect structured data in the form of published tables of financial data and the contents of fill-in-the-blank documents that record information from surveys you distributed. However, you might also collect unstructured data such as social media posts and tweets that do not have an internal repeating structure.

Typically, you preprocess or filter unstructured data before performing deep analysis. For example, to analyze social media posts you could use business analytics methods that determine whether the content of each post is a positive, neutral, or negative comment. The “type of comment” can become a new variable that can be inserted into a *structured* record, along with other attributes of the post, such as the number of words, and demographic data about the writer of the post.

Unstructured data can form part of a big data collection. When analyzed as part of a big data collection, you typically see the results of the preprocessing and not the unstructured data itself. Because unstructured data usually has some (external) structure, some authorities prefer to use the term *semistructured data*. If you are familiar with that term, understand that this book’s use of the phrase *unstructured data* incorporates that category.

FTF.3 Getting Started Learning Statistics

Learning the **operational definitions**, precise definitions and explanations that all can understand clearly, of several basic terms is a good way to get started learning statistics. Previously, you learned that *data* are the facts about the world that one seeks to study and explore. A related term, *variable of interest*, commonly shortened to *variable*, can be used to precisely define data in its statistical sense.

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 title and number of chapters, as these facts can vary from book to book. For a given item, variables have a specific value. For this book, the value of the variable title would be “Statistics for Managers Using Microsoft Excel,” and “17” would be the value for the variable number of chapters.

Using the definition of variable, you can state the definition of data, in its statistical sense, 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 when referring to this book? Without this rule, you might say that the single list “Levine, Stephan, Szabat” is the value. However, applying this rule, we say that the variable author has the three separate values: “Levine”, “Stephan”, and “Szabat”. This distinction of using only *single-value data* has the practical benefit of simplifying the task of entering your data into a computer system for analysis.

Using the definitions of data and variable, you can restate the definition of statistics 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.

Do not confuse this use of the word statistics with the noun *statistic*, the plural of which is, confusingly, *statistics*.

Statistic

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.

student TIP

Business analytics, discussed in Chapter 17, combine mostly inferential methods with methods from other disciplines.

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, invalid statistics can be produced if any tasks in the DCOVA framework are applied incorrectly. As discussed in later chapters, 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 discussed later in this book, 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 Preparing to Use Microsoft Excel for Statistics

As Section FTF.1 explains, the proper use of business statistics requires a framework to apply statistics correctly, analytic skills, and software to automate calculation. This book uses Microsoft Excel to demonstrate the integral role of software in applying statistics to decision making, and preparing to use Microsoft Excel is one of the first things you can do to prepare yourself to learn business statistics from this book.

Microsoft Excel is the data analysis component of Microsoft Office that evolved from earlier electronic spreadsheets used in accounting and financial applications. In Excel, you use **worksheets** (also known as spreadsheets) that organize data in tabular blocks as well as store **formulas**, instructions to process that data. You make entries in worksheet **cells** that are formed by the intersections of worksheet rows and columns. You refer to individual cells by their column letter and row number address, such as A1 for the uppermost left top cell (in column A and row 1). Into each cell, you place a single data value or a formula. With the proper design, worksheets can also present summaries of data and results of applying a particular statistical method.

“Excel files” are not single worksheets but **workbooks**, collections of one or more worksheets and **chart sheets**, sheets that display visualizations of data. Because workbooks contain collections, you can clearly present information in more than one way on different