



Statistics for Managers

Using Microsoft® Excel

8TH 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>

This page intentionally left blank

Statistics for Managers Using Microsoft® Excel

8th 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

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

Editorial Director: Chris Hoag
Editor in Chief: Deirdre Lynch
Acquisitions Editor: Suzanna Bainbridge
Senior Editorial Assistant: Justin Billing
Program Manager: Chere Bemelmans
Project Manager: Sherry Berg
Program Management Team Lead: Karen Wernholm
Project Management Team Lead: Peter Silvia
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: Cenveo Publisher Services
Production Coordination, Composition, and Illustrations: Lumina Datamatics
Cover Image: Serp/Shutterstock

Copyright © 2017, 2014, 2011 by Pearson Education, Inc. All Rights Reserved. Printed in the United States of America. This publication is protected by copyright, and permission should be obtained from the publisher prior to any prohibited reproduction, storage in a retrieval system, or transmission in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise. For information regarding permissions, request forms and the appropriate contacts within the Pearson Education Global Rights & Permissions department, please visit www.pearsoned.com/permissions/.

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

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

MICROSOFT® 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.

Library of Congress Cataloging-in-Publication Data

Levine, David M., 1946– author.

Statistics for managers using Microsoft Excel / David M. Levine, David F. Stephan, Kathryn A. Szabat.—8th edition.

pages cm

Includes bibliographical references and index.

ISBN 978-0-13-417305-4 (hardcover)

1. Microsoft Excel (Computer file) 2. Management—Statistical methods. 3. Commercial statistics. 4. Electronic spreadsheets.

5. Management—Statistical methods—Computer programs. 6. Commercial statistics—Computer programs. I. Stephan,

David F., author. II. Szabat, Kathryn A., author. III. Title.

HD30.215.S73 2017

519.50285/554—dc23

2015027139

1 2 3 4 5 6 7 8 9 10—CRK—19 18 17 16 15

PEARSON

www.pearsonhighered.com

ISBN-10: 0-13-417305-8
ISBN-13: 978-0-13-417305-4

*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 xvii

First Things First 1

- 1 Defining and Collecting Data 12
- 2 Organizing and Visualizing Variables 32
- 3 Numerical Descriptive Measures 95
- 4 Basic Probability 141
- 5 Discrete Probability Distributions 166
- 6 The Normal Distribution and Other Continuous Distributions 189
- 7 Sampling Distributions 216
- 8 Confidence Interval Estimation 237
- 9 Fundamentals of Hypothesis Testing: One-Sample Tests 270
- 10 Two-Sample Tests 307
- 11 Analysis of Variance 348
- 12 Chi-Square Tests and Nonparametric Tests 386
- 13 Simple Linear Regression 427
- 14 Introduction to Multiple Regression 475
- 15 Multiple Regression Model Building 521
- 16 Time-Series Forecasting 553
- 17 Getting Ready To Analyze Data In The Future 598
- 18 Statistical Applications in Quality Management (*online*) 18-1
- 19 Decision Making (*online*) 19-1

Appendices A–G 613

Self-Test Solutions and Answers to Selected Even-Numbered Problems 661

Index 692

Credits 699

Contents

Preface xvii

First Things First 1

USING STATISTICS: “The Price of Admission” 1

Now Appearing on Broadway . . . and Everywhere Else 2

FTF.1 Think Differently About Statistics 2

Statistics: A Way of Thinking 2

Analytical Skills More Important than Arithmetic Skills 3

Statistics: An Important Part of Your Business Education 3

FTF.2 Business Analytics: The Changing Face of Statistics 4 “Big Data” 4

Structured Versus Unstructured Data 4

FTF.3 Getting Started Learning Statistics 5

Statistic 5

Can Statistics (*p*., Statistic) Lie? 6

FTF.4 Preparing to Use Microsoft Excel for Statistics 6

Reusability Through Recalculation 7

Practical Matters: Skills You Need 7

Ways of Working with Excel 7

Excel Guides 8

Which Excel Version to Use? 8

Conventions Used 8

REFERENCES 9

KEY TERMS 9

EXCEL GUIDE 10

EG.1 Entering Data 10

EG.2 Reviewing Worksheets 10

EG.3 If You Plan to Use the *Workbook* Instructions 11

1 Defining and Collecting Data 12

USING STATISTICS: Defining Moments 12

1.1 Defining Variables 13

Classifying Variables by Type 14

Measurement Scales 14

1.2 Collecting Data 15

Populations and Samples 16

Data Sources 16

1.3 Types of Sampling Methods 17

Simple Random Sample 18

Systematic Sample 18

Stratified Sample 19

Cluster Sample 19

1.4 Data Preparation 20

Data Cleaning 20

Data Formatting 21

Stacked and Unstacked Variables 21

Recoding Variables 22

1.5 Types of Survey Errors 23

Coverage Error 23

Nonresponse Error 23

Sampling Error 23

Measurement Error 24

Ethical Issues About Surveys 24

CONSIDER THIS: New Media Surveys/Old Survey Errors 24

USING STATISTICS: Defining Moments, Revisited 26

SUMMARY 26

REFERENCES 26

KEY TERMS 26

CHECKING YOUR UNDERSTANDING 27

CHAPTER REVIEW PROBLEMS 27

CASES FOR CHAPTER 1 28

Managing Ashland MultiComm Services 28

CardioGood Fitness 28

Clear Mountain State Student Survey 29

Learning with the Digital Cases 29

CHAPTER 1 EXCEL GUIDE 30

EG1.1 Defining Variables 30

EG1.2 Collecting Data 30

EG1.3 Types of Sampling Methods 31

2 Organizing and Visualizing Variables 32

USING STATISTICS: “The Choice Is Yours” 32

2.1 Organizing Categorical Variables 33

The Summary Table 33

The Contingency Table 34

2.2 Organizing Numerical Variables 37

The Frequency Distribution 38

Classes and Excel Bins 40

The Relative Frequency Distribution and the Percentage Distribution 41

The Cumulative Distribution 43

2.3 Visualizing Categorical Variables 46

The Bar Chart 46

The Pie Chart and the Doughnut Chart 47

	The Pareto Chart	48
	Visualizing Two Categorical Variables	50
2.4	Visualizing Numerical Variables	52
	The Stem-and-Leaf Display	53
	The Histogram	54
	The Percentage Polygon	55
	The Cumulative Percentage Polygon (Ogive)	56
2.5	Visualizing Two Numerical Variables	59
	The Scatter Plot	59
	The Time-Series Plot	61
2.6	Organizing and Visualizing a Mix of Variables	63
	Multidimensional Contingency Table	63
	Adding a Numerical Variable to a Multidimensional Contingency Table	64
	Drill Down	64
	Excel Slicers	65
	PivotChart	66
	Sparklines	66
2.7	The Challenge in Organizing and Visualizing Variables	68
	Obscuring Data	68
	Creating False Impressions	69
	Chartjunk	70
	EXHIBIT: Best Practices for Creating Visualizations	72
	USING STATISTICS: The Choice Is Yours, Revisited	73
	SUMMARY	73
	REFERENCES	74
	KEY EQUATIONS	74
	KEY TERMS	75
	CHECKING YOUR UNDERSTANDING	75
	CHAPTER REVIEW PROBLEMS	75
	CASES FOR CHAPTER 2 80	
	Managing Ashland MultiComm Services	80
	Digital Case	80
	CardioGood Fitness	81
	The Choice Is Yours Follow-Up	81
	Clear Mountain State Student Survey	81
	CHAPTER 2 EXCEL GUIDE 82	
	EG2.1 Organizing Categorical Variables	82
	EG2.2 Organizing Numerical Variables	84
	EG2.3 Visualizing Categorical Variables	86
	EG2.4 Visualizing Numerical Variables	88
	EG2.5 Visualizing Two Numerical Variables	92
	EG2.6 Organizing and Visualizing a Set of Variables	92

3 Numerical Descriptive Measures 95

USING STATISTICS: More Descriptive Choices 95

3.1	Central Tendency	96
	The Mean	96
	The Median	98
	The Mode	99
	The Geometric Mean	100
3.2	Variation and Shape	101
	The Range	101

	The Variance and the Standard Deviation	102
	EXHIBIT: Manually Calculating the Sample Variance, S^2 , and Sample Standard Deviation, S	103
	The Coefficient of Variation	105
	Z Scores	106
	Shape: Skewness	108
	Shape: Kurtosis	108

3.3	Exploring Numerical Data	113
	Quartiles	113
	EXHIBIT: Rules for Calculating the Quartiles from a Set of Ranked Values	113
	The Interquartile Range	115
	The Five-Number Summary	115
	The Boxplot	117
3.4	Numerical Descriptive Measures for a Population	119
	The Population Mean	120
	The Population Variance and Standard Deviation	120
	The Empirical Rule	121
	Chebyshev's Theorem	122
3.5	The Covariance and the Coefficient of Correlation	124
	The Covariance	124
	The Coefficient of Correlation	125
3.6	Statistics: Pitfalls and Ethical Issues	130
	USING STATISTICS: More Descriptive Choices, Revisited	130

	SUMMARY	130
	REFERENCES	131
	KEY EQUATIONS	131
	KEY TERMS	132
	CHECKING YOUR UNDERSTANDING	132
	CHAPTER REVIEW PROBLEMS	133

	CASES FOR CHAPTER 3 136	
	Managing Ashland MultiComm Services	136
	Digital Case	136
	CardioGood Fitness	136
	More Descriptive Choices Follow-up	136
	Clear Mountain State Student Survey	136

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

4 Basic Probability 141

USING STATISTICS: Possibilities at M&R Electronics World 141

4.1	Basic Probability Concepts	142
	Events and Sample Spaces	143
	Contingency Tables	145
	Simple Probability	145
	Joint Probability	146
	Marginal Probability	147
	General Addition Rule	147

- 4.2 Conditional Probability 151
 - Computing Conditional Probabilities 151
 - Decision Trees 152
 - Independence 154
 - Multiplication Rules 155
 - Marginal Probability Using the General Multiplication Rule 156

4.3 Ethical Issues and Probability 158

4.4 Bayes' Theorem 159

CONSIDER THIS: Divine Providence and Spam 159

4.5 Counting Rules 160

USING STATISTICS: Possibilities at M&R Electronics World, Revisited 161

SUMMARY 161

REFERENCES 161

KEY EQUATIONS 161

KEY TERMS 162

CHECKING YOUR UNDERSTANDING 162

CHAPTER REVIEW PROBLEMS 162

CASES FOR CHAPTER 4 164

- Digital Case 164
- CardioGood Fitness 164
- The Choice Is Yours Follow-Up 164
- Clear Mountain State Student Survey 164

CHAPTER 4 EXCEL GUIDE 165

- EG4.1 Basic Probability Concepts 165
- EG4.4 Bayes' Theorem 165

5 Discrete Probability Distributions 166

USING STATISTICS: Events of Interest at Ricknel Home Centers 166

- 5.1 The Probability Distribution for a Discrete Variable 167
 - Expected Value of a Discrete Variable 167
 - Variance and Standard Deviation of a Discrete Variable 168

5.2 Binomial Distribution 171

5.3 Poisson Distribution 178

5.4 Covariance of a Probability Distribution and its Application in Finance 181

5.5 Hypergeometric Distribution 182

USING STATISTICS: Events of Interest..., Revisited 182

SUMMARY 182

REFERENCES 182

KEY EQUATIONS 182

KEY TERMS 183

CHECKING YOUR UNDERSTANDING 183

CHAPTER REVIEW PROBLEMS 183

CASES FOR CHAPTER 5 185

- Managing Ashland MultiComm Services 185
- Digital Case 186

CHAPTER 5 EXCEL GUIDE 187

- EG5.1 The Probability Distribution for a Discrete Variable 187

EG5.2 Binomial Distribution 187

EG5.3 Poisson Distribution 188

6 The Normal Distribution and Other Continuous Distributions 189

USING STATISTICS: Normal Load Times at MyTVLab 189

6.1 Continuous Probability Distributions 190

6.2 The Normal Distribution 191

EXHIBIT: Normal Distribution Important Theoretical Properties 191

Computing Normal Probabilities 192

VISUAL EXPLORATIONS: Exploring the Normal Distribution 198

Finding X Values 198

CONSIDER THIS: What Is Normal? 202

6.3 Evaluating Normality 203

Comparing Data Characteristics to Theoretical Properties 204

Constructing the Normal Probability Plot 205

6.4 The Uniform Distribution 207

6.5 The Exponential Distribution 209

6.6 The Normal Approximation to the Binomial Distribution 209

USING STATISTICS: Normal Load Times..., Revisited 210

SUMMARY 210

REFERENCES 210

KEY EQUATIONS 211

KEY TERMS 211

CHECKING YOUR UNDERSTANDING 211

CHAPTER REVIEW PROBLEMS 211

CASES FOR CHAPTER 6 213

- Managing Ashland MultiComm Services 213
- CardioGood Fitness 213
- More Descriptive Choices Follow-up 213
- Clear Mountain State Student Survey 213
- Digital Case 213

CHAPTER 6 EXCEL GUIDE 214

- EG6.1 Continuous Probability Distributions 214
- EG6.2 The Normal Distribution 214
- EG6.3 Evaluating Normality 124

7 Sampling Distributions 216

USING STATISTICS: Sampling Oxford Cereals 216

7.1 Sampling Distributions 217

7.2 Sampling Distribution of the Mean 217

The Unbiased Property of the Sample Mean 217

Standard Error of the Mean 219

Sampling from Normally Distributed Populations 220

Sampling from Non-normally Distributed Populations—The Central Limit Theorem 223

EXHIBIT: Normality and the Sampling Distribution of the Mean 224

VISUAL EXPLORATIONS: Exploring Sampling Distributions 227

7.3 Sampling Distribution of the Proportion 228

USING STATISTICS: Sampling Oxford Cereals, Revisited 231

SUMMARY 232

REFERENCES 232

KEY EQUATIONS 232

KEY TERMS 232

CHECKING YOUR UNDERSTANDING 233

CHAPTER REVIEW PROBLEMS 233

CASES FOR CHAPTER 7 235

Managing Ashland Multicomm Services 235

Digital Case 235

CHAPTER 7 EXCEL GUIDE 236

EG7.2 Sampling Distribution of the Mean 236

8 Confidence Interval Estimation 237

USING STATISTICS: Getting Estimates at Ricknel Home Centers 237

8.1 Confidence Interval Estimate for the Mean (σ Known) 238

Can You Ever Know the Population Standard Deviation? 243

8.2 Confidence Interval Estimate for the Mean (σ Unknown) 244

Student's t Distribution 244

Properties of the t Distribution 245

The Concept of Degrees of Freedom 246

The Confidence Interval Statement 247

8.3 Confidence Interval Estimate for the Proportion 252

8.4 Determining Sample Size 255

Sample Size Determination for the Mean 255

Sample Size Determination for the Proportion 257

8.5 Confidence Interval Estimation and Ethical Issues 260

8.6 Application of Confidence Interval Estimation in Auditing 261

8.7 Estimation and Sample Size Estimation for Finite Populations 261

8.8 Bootstrapping 261

USING STATISTICS: Getting Estimates. . . , Revisited 261

SUMMARY 262

REFERENCES 262

KEY EQUATIONS 262

KEY TERMS 263

CHECKING YOUR UNDERSTANDING 263

CHAPTER REVIEW PROBLEMS 263

CASES FOR CHAPTER 8 266

Managing Ashland MultiComm Services 266

Digital Case 267

Sure Value Convenience Stores 267

CardioGood Fitness 267

More Descriptive Choices Follow-Up 267

Clear Mountain State Student Survey 267

CHAPTER 8 EXCEL GUIDE 268

EG8.1 Confidence Interval Estimate for the Mean (σ Known) 268

EG8.2 Confidence Interval Estimate for the Mean (σ Unknown) 268

EG8.3 Confidence Interval Estimate for the Proportion 269

EG8.4 Determining Sample Size 269

9 Fundamentals of Hypothesis Testing: One-Sample Tests 270

USING STATISTICS: Significant Testing at Oxford Cereals 270

9.1 Fundamentals of Hypothesis-Testing Methodology 271

The Null and Alternative Hypotheses 271

The Critical Value of the Test Statistic 272

Regions of Rejection and Nonrejection 273

Risks in Decision Making Using Hypothesis Testing 273

Z Test for the Mean (σ Known) 276

Hypothesis Testing Using the Critical Value Approach 276

EXHIBIT: The Critical Value Approach to Hypothesis Testing 277

Hypothesis Testing Using the p -Value Approach 279

EXHIBIT: The p -Value Approach to Hypothesis Testing 280

A Connection Between Confidence Interval Estimation and Hypothesis Testing 281

Can You Ever Know the Population Standard Deviation? 282

9.2 t Test of Hypothesis for the Mean (σ Unknown) 284

The Critical Value Approach 284

p -Value Approach 286

Checking the Normality Assumption 286

9.3 One-Tail Tests 290

The Critical Value Approach 290

The p -Value Approach 291

EXHIBIT: The Null and Alternative Hypotheses in One-Tail Tests 293

9.4 Z Test of Hypothesis for the Proportion 294

The Critical Value Approach 295

The p -Value Approach 296

9.5 Potential Hypothesis-Testing Pitfalls and Ethical Issues 298

EXHIBIT: Questions for the Planning Stage of Hypothesis Testing 298

Statistical Significance Versus Practical Significance 299

Statistical *Insignificance* Versus Importance 299

Reporting of Findings 299

Ethical Issues 299

9.6 Power of the Test 300

USING STATISTICS: Significant Testing. . . , Revisited 300

SUMMARY 300

REFERENCES 301

KEY EQUATIONS 301

KEY TERMS 301

CHECKING YOUR UNDERSTANDING 301

CHAPTER REVIEW PROBLEMS 302

CASES FOR CHAPTER 9 304

- Managing Ashland MultiComm Services 304
- Digital Case 304
- Sure Value Convenience Stores 304

CHAPTER 9 EXCEL GUIDE 305

- EG9.1 Fundamentals of Hypothesis-Testing Methodology 305
- EG9.2 t Test of Hypothesis for the Mean (σ Unknown) 305
- EG9.3 One-Tail Tests 306
- EG9.4 Z Test of Hypothesis for the Proportion 306

10 Two-Sample Tests 307**USING STATISTICS: Differing Means for Selling Streaming Media Players at Arlington's? 307****10.1 Comparing the Means of Two Independent Populations 308**

- Pooled-Variance t Test for the Difference Between Two Means 308
- Confidence Interval Estimate for the Difference Between Two Means 313
- t Test for the Difference Between Two Means, Assuming Unequal Variances 314

CONSIDER THIS: Do People Really Do This? 315**10.2 Comparing the Means of Two Related Populations 317**

- Paired t Test 318
- Confidence Interval Estimate for the Mean Difference 323

10.3 Comparing the Proportions of Two Independent Populations 325

- Z Test for the Difference Between Two Proportions 326
- Confidence Interval Estimate for the Difference Between Two Proportions 330

10.4 F Test for the Ratio of Two Variances 332**10.5 Effect Size 336****USING STATISTICS: Differing Means for Selling. . . , Revisited 337****SUMMARY 337****REFERENCES 338****KEY EQUATIONS 338****KEY TERMS 339****CHECKING YOUR UNDERSTANDING 339****CHAPTER REVIEW PROBLEMS 339****CASES FOR CHAPTER 10 341**

- Managing Ashland MultiComm Services 341
- Digital Case 342
- Sure Value Convenience Stores 342
- CardioGood Fitness 342
- More Descriptive Choices Follow-Up 342
- Clear Mountain State Student Survey 342

CHAPTER 10 EXCEL GUIDE 343

- EG10.1 Comparing The Means of Two Independent Populations 343
- EG10.2 Comparing the Means of Two Related Populations 345
- EG10.3 Comparing the Proportions of Two Independent Populations 346
- EG10.4 F Test for the Ratio of Two Variances 347

11 Analysis of Variance 348**USING STATISTICS: The Means to Find Differences at Arlington's 348****11.1 The Completely Randomized Design: One-Way ANOVA 349**

- Analyzing Variation in One-Way ANOVA 350
- F Test for Differences Among More Than Two Means 352
- One-Way ANOVA F Test Assumptions 356
- Levene Test for Homogeneity of Variance 357
- Multiple Comparisons: The Tukey-Kramer Procedure 358
- The Analysis of Means (ANOM) 360

11.2 The Factorial Design: Two-Way ANOVA 363

- Factor and Interaction Effects 364
- Testing for Factor and Interaction Effects 366
- Multiple Comparisons: The Tukey Procedure 369
- Visualizing Interaction Effects: The Cell Means Plot 371
- Interpreting Interaction Effects 371

11.3 The Randomized Block Design 375**11.4 Fixed Effects, Random Effects, and Mixed Effects Models 375****USING STATISTICS: The Means to Find Differences at Arlington's Revisited 375****SUMMARY 376****REFERENCES 376****KEY EQUATIONS 376****KEY TERMS 377****CHECKING YOUR UNDERSTANDING 378****CHAPTER REVIEW PROBLEMS 378****CASES FOR CHAPTER 11 380**

- Managing Ashland MultiComm Services 380
- PHASE 1 380
- PHASE 2 380
- Digital Case 381
- Sure Value Convenience Stores 381
- CardioGood Fitness 381
- More Descriptive Choices Follow-Up 381
- Clear Mountain State Student Survey 381

CHAPTER 11 EXCEL GUIDE 382

- EG11.1 The Completely Randomized Design: One-Way ANOVA 382
- EG11.2 The Factorial Design: Two-Way ANOVA 384

12 Chi-Square and Nonparametric Tests 386**USING STATISTICS: Avoiding Guesswork about Resort Guests 386****12.1 Chi-Square Test for the Difference Between Two Proportions 387****12.2 Chi-Square Test for Differences Among More Than Two Proportions 394**

- The Marascuilo Procedure 397
- The Analysis of Proportions (ANOP) 399

12.3 Chi-Square Test of Independence 400

- 12.4** Wilcoxon Rank Sum Test: A Nonparametric Method for Two Independent Populations 406
- 12.5** Kruskal-Wallis Rank Test: A Nonparametric Method for the One-Way ANOVA 412
 - Assumptions 415
- 12.6** McNemar Test for the Difference Between Two Proportions (Related Samples) 417
- 12.7** Chi-Square Test for the Variance or Standard Deviation 417

USING STATISTICS: Avoiding Guesswork. . . , Revisited 418

SUMMARY 418

REFERENCES 419

KEY EQUATIONS 419

KEY TERMS 420

CHECKING YOUR UNDERSTANDING 420

CHAPTER REVIEW PROBLEMS 420

CASES FOR CHAPTER 12 422

- Managing Ashland MultiComm Services 422
- PHASE 1 422
- PHASE 2 422
- Digital Case 423
- Sure Value Convenience Stores 423
- CardioGood Fitness 423
- More Descriptive Choices Follow-Up 423
- Clear Mountain State Student Survey 423

CHAPTER 12 EXCEL GUIDE 424

- EG12.1 Chi-Square Test for the Difference Between Two Proportions 424
- EG12.2 Chi-Square Test for Differences Among More Than Two Proportions 424
- EG12.3 Chi-Square Test of Independence 425
- EG12.4 Wilcoxon Rank Sum Test: a Nonparametric Method for Two Independent Populations 425
- EG12.5 Kruskal-Wallis Rank Test: a Nonparametric Method for the One-Way ANOVA 426

13 Simple Linear Regression 427

USING STATISTICS: Knowing Customers at Sunflowers Apparel 427

- 13.1** Types of Regression Models 428
 - Simple Linear Regression Models 429
- 13.2** Determining the Simple Linear Regression Equation 430
 - The Least-Squares Method 430
 - Predictions in Regression Analysis: Interpolation Versus Extrapolation 433
 - Computing the Y Intercept, b_0 and the Slope, b_1 433
 - VISUAL EXPLORATIONS: Exploring Simple Linear Regression Coefficients 436
- 13.3** Measures of Variation 438
 - Computing the Sum of Squares 438
 - The Coefficient of Determination 439
 - Standard Error of the Estimate 441
- 13.4** Assumptions of Regression 443
- 13.5** Residual Analysis 443
 - Evaluating the Assumptions 443

- 13.6** Measuring Autocorrelation: The Durbin-Watson Statistic 447

- Residual Plots to Detect Autocorrelation 447
- The Durbin-Watson Statistic 448

- 13.7** Inferences About the Slope and Correlation Coefficient 451

- t Test for the Slope 451
- F Test for the Slope 453
- Confidence Interval Estimate for the Slope 454
- t Test for the Correlation Coefficient 455

- 13.8** Estimation of Mean Values and Prediction of Individual Values 458

- The Confidence Interval Estimate for the Mean Response 458
- The Prediction Interval for an Individual Response 459

- 13.9** Potential Pitfalls in Regression 462

- EXHIBIT: Six Steps for Avoiding the Potential Pitfalls 462

USING STATISTICS: Knowing Customers. . . , Revisited 464

SUMMARY 464

REFERENCES 465

KEY EQUATIONS 466

KEY TERMS 467

CHECKING YOUR UNDERSTANDING 467

CHAPTER REVIEW PROBLEMS 467

CASES FOR CHAPTER 13 471

- Managing Ashland MultiComm Services 471
- Digital Case 471
- Brynne Packaging 471

CHAPTER 13 EXCEL GUIDE 472

- EG13.2 Determining the Simple Linear Regression Equation 472
- EG13.3 Measures of Variation 473
- EG13.4 Assumptions of Regression 473
- EG13.5 Residual Analysis 473
- EG13.6 Measuring Autocorrelation: The Durbin-Watson Statistic 474
- EG13.7 Inferences about the Slope and Correlation Coefficient 474
- EG13.8 Estimation of Mean Values and Prediction of Individual Values 474

14 Introduction to Multiple Regression 475

USING STATISTICS: The Multiple Effects of OmniPower Bars 475

- 14.1** Developing a Multiple Regression Model 476
 - Interpreting the Regression Coefficients 476
 - Predicting the Dependent Variable Y 479
- 14.2** r^2 , Adjusted r^2 , and the Overall F Test 481
 - Coefficient of Multiple Determination 481
 - Adjusted r^2 481
 - Test for the Significance of the Overall Multiple Regression Model 482
- 14.3** Residual Analysis for the Multiple Regression Model 484
- 14.4** Inferences Concerning the Population Regression Coefficients 486
 - Tests of Hypothesis 486
 - Confidence Interval Estimation 487
- 14.5** Testing Portions of the Multiple Regression Model 489
 - Coefficients of Partial Determination 493

14.6 Using Dummy Variables and Interaction Terms in Regression Models 495
 Interactions 497

14.7 Logistic Regression 504

USING STATISTICS: The Multiple Effects . . . , Revisited 509

SUMMARY 509

REFERENCES 511

KEY EQUATIONS 511

KEY TERMS 512

CHECKING YOUR UNDERSTANDING 512

CHAPTER REVIEW PROBLEMS 512

CASES FOR CHAPTER 14 515
 Managing Ashland MultiComm Services 515
 Digital Case 515

CHAPTER 14 EXCEL GUIDE 517
 EG14.1 Developing a Multiple Regression Model 517
 EG14.2 r^2 , Adjusted r^2 , and the Overall F Test 518
 EG14.3 Residual Analysis for the Multiple Regression Model 518
 EG14.4 Inferences Concerning the Population Regression Coefficients 519
 EG14.5 Testing Portions of the Multiple Regression Model 519
 EG14.6 Using Dummy Variables and Interaction Terms in Regression Models 519
 EG14.7 Logistic Regression 520

15 Multiple Regression Model Building 521

USING STATISTICS: Valuing Parsimony at WSTA-TV 521

15.1 Quadratic Regression Model 522
 Finding the Regression Coefficients and Predicting Y 522
 Testing for the Significance of the Quadratic Model 525
 Testing the Quadratic Effect 525
 The Coefficient of Multiple Determination 527

15.2 Using Transformations in Regression Models 529
 The Square-Root Transformation 529
 The Log Transformation 531

15.3 Collinearity 534

15.4 Model Building 535
 The Stepwise Regression Approach to Model Building 537
 The Best Subsets Approach to Model Building 538
 Model Validation 541
 EXHIBIT: Steps for Successful Model Building 542

15.5 Pitfalls in Multiple Regression and Ethical Issues 544
 Pitfalls in Multiple Regression 544
 Ethical Issues 544

USING STATISTICS: Valuing Parsimony..., Revisited 544

SUMMARY 545

REFERENCES 546

KEY EQUATIONS 546

KEY TERMS 546

CHECKING YOUR UNDERSTANDING 546

CHAPTER REVIEW PROBLEMS 546

CASES FOR CHAPTER 15 548
 The Mountain States Potato Company 548

Sure Value Convenience Stores 549
 Digital Case 549
 The Craybill Instrumentation Company Case 549
 More Descriptive Choices Follow-Up 550

CHAPTER 15 EXCEL GUIDE 551

EG15.1 The Quadratic Regression Model 551
 EG15.2 Using Transformations In Regression Models 551
 EG15.3 Collinearity 552
 EG15.4 Model Building 552

16 Time-Series Forecasting 553

USING STATISTICS: Principled Forecasting 553

16.1 The Importance of Business Forecasting 554

16.2 Component Factors of Time-Series Models 554

16.3 Smoothing an Annual Time Series 555
 Moving Averages 556
 Exponential Smoothing 558

16.4 Least-Squares Trend Fitting and Forecasting 561
 The Linear Trend Model 561
 The Quadratic Trend Model 563
 The Exponential Trend Model 564
 Model Selection Using First, Second, and Percentage Differences 566

16.5 Autoregressive Modeling for Trend Fitting and Forecasting 571
 Selecting an Appropriate Autoregressive Model 572
 Determining the Appropriateness of a Selected Model 573
 EXHIBIT: Autoregressive Modeling Steps 575

16.6 Choosing an Appropriate Forecasting Model 580
 Performing a Residual Analysis 580
 Measuring the Magnitude of the Residuals Through Squared or Absolute Differences 581
 Using the Principle of Parsimony 581
 A Comparison of Four Forecasting Methods 581

16.7 Time-Series Forecasting of Seasonal Data 583
 Least-Squares Forecasting with Monthly or Quarterly Data 584

16.8 Index Numbers 589

CONSIDER THIS: Let the Model User Beware 589

USING STATISTICS: Principled Forecasting, Revisited 589

SUMMARY 590

REFERENCES 591

KEY EQUATIONS 591

KEY TERMS 592

CHECKING YOUR UNDERSTANDING 592

CHAPTER REVIEW PROBLEMS 592

CASES FOR CHAPTER 16 593
 Managing Ashland MultiComm Services 593
 Digital Case 593

CHAPTER 16 EXCEL GUIDE 594
 EG16.3 Smoothing an Annual Time Series 594
 EG16.4 Least-Squares Trend Fitting and Forecasting 595
 EG16.5 Autoregressive Modeling for Trend Fitting and Forecasting 596
 EG16.6 Choosing an Appropriate Forecasting Model 596
 EG16.7 Time-Series Forecasting of Seasonal Data 597

17 Getting Ready to Analyze Data in the Future 598

USING STATISTICS: Mounting Future Analyses 598

17.1 Analyzing Numerical Variables 599

EXHIBIT: Questions to Ask When Analyzing Numerical Variables 599

Describe the Characteristics of a Numerical Variable? 599

Reach Conclusions about the Population Mean or the Standard Deviation? 599

Determine Whether the Mean and/or Standard Deviation Differs Depending on the Group? 600

Determine Which Factors Affect the Value of a Variable? 600

Predict the Value of a Variable Based on the Values of Other Variables? 601

Determine Whether the Values of a Variable Are Stable Over Time? 601

17.2 Analyzing Categorical Variables 601

EXHIBIT: Questions to Ask When Analyzing Categorical Variables 601

Describe the Proportion of Items of Interest in Each Category? 601

Reach Conclusions about the Proportion of Items of Interest? 601

Determine Whether the Proportion of Items of Interest Differs Depending on the Group? 602

Predict the Proportion of Items of Interest Based on the Values of Other Variables? 602

Determine Whether the Proportion of Items of Interest Is Stable Over Time? 602

USING STATISTICS: Back to Arlingtons for the Future 602

17.3 Introduction to Business Analytics 603

Data Mining 603

Power Pivot 603

17.4 Descriptive Analytics 604

Dashboards 605

Dashboard Elements 605

17.5 Predictive Analytics 606

Classification and Regression Trees 607

USING STATISTICS: The Future to be Visited 608

REFERENCES 608

CHAPTER REVIEW PROBLEMS 608

CHAPTER 17 EXCEL GUIDE 611

EG17.3 Introduction to Business Analytics 611

EG17.4 Descriptive Analytics 611

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 613

- A. Basic Math Concepts and Symbols 614
 - A.1 Rules for Arithmetic Operations 614
 - A.2 Rules for Algebra: Exponents and Square Roots 614
 - A.3 Rules for Logarithms 615
 - A.4 Summation Notation 616
 - A.5 Statistical Symbols 619
 - A.6 Greek Alphabet 619
- B Important Excel Skills and Concepts 620
 - B.1 Which Excel Do You Use? 620
 - B.2 Basic Operations 621
 - B.3 Formulas and Cell References 621
 - B.4 Entering a Formula 623
 - B.5 Formatting Cell Contents 624
 - B.6 Formatting Charts 625
 - B.7 Selecting Cell Ranges for Charts 626
 - B.8 Deleting the “Extra” Histogram Bar 627
 - B.9 Creating Histograms for Discrete Probability Distributions 627
- C. Online Resources 628
 - C.1 About the Online Resources for This Book 628
 - C.2 Accessing the Online Resources 628
 - C.3 Details of Online Resources 628
 - C.4 PHStat 635
- D. Configuring Microsoft Excel 636
 - D.1 Getting Microsoft Excel Ready for Use 636
 - D.2 Checking for the Presence of the Analysis ToolPak or Solver Add-Ins 636

- D.3 Configuring Microsoft Windows Excel Security Settings 636
- D.4 Opening Pearson-Supplied Add-Ins 637
- E. Tables 638
 - E.1 Table of Random Numbers 638
 - E.2 The Cumulative Standardized Normal Distribution 640
 - E.3 Critical Values of t 642
 - E.4 Critical Values of χ^2 644
 - E.5 Critical Values of F 645
 - E.6 Lower and Upper Critical Values, T_1 , of the Wilcoxon Rank Sum Test 649
 - E.7 Critical Values of the Studentized Range, Q 650
 - E.8 Critical Values, d_L and d_U , of the Durbin–Watson Statistic, D (Critical Values Are One-Sided) 652
 - E.9 Control Chart Factors 653
 - E.10 The Standardized Normal Distribution 654
- F. Useful Excel Knowledge 655
 - F.1 Useful Keyboard Shortcuts 655
 - F.2 Verifying Formulas and Worksheets 655
 - F.3 New Function Names 655
 - F.4 Understanding the Nonstatistical Functions 657
- G. Software FAQs 659
 - G.1 PHStat FAQs 659
 - G.2 Microsoft Excel FAQs 659

Self-Test Solutions and Answers to Selected Even-Numbered Problems 661**Index 692****Credits 699**

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 Lorczak, 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.

Contact Us!

Please email us at authors@davidlevinestatistics.com or tweet us @[BusStatBooks](https://twitter.com/BusStatBooks) with your questions about the contents of this book. Please include the hashtag #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](https://www.pearsoned.com/custhelp). 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.davidlevinestatistics.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 xxiii and xxiv).

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

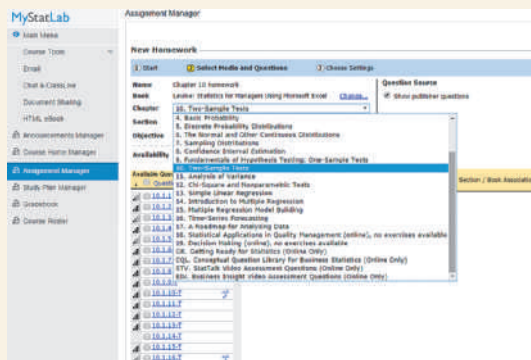
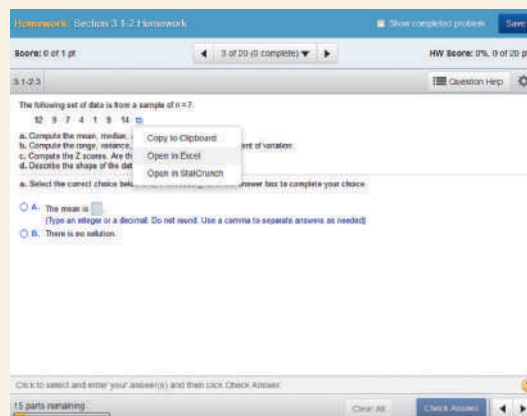
Resources for Success

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.

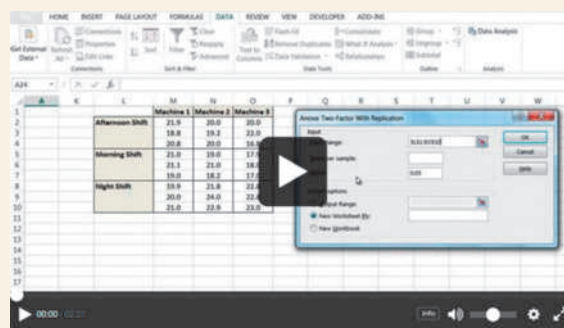


Diverse Question Libraries

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.pearsonhighered.com/irc) 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.pearsonhighered.com/irc) 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.pearsonhighered.com/irc) 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.

Student Resources

Student's Solutions Manual, by Professor Pin Tian Ng of Northern Arizona University, provides detailed solutions to virtually all the even-numbered exercises and worked-out solutions to the self-test problems (ISBN-13: 978-0-13-417382-5).

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.pearsonhighered.com/irc) 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.