

Access Code(s) for Purchased Term(s):

[Empty box for access code]

Your access code is not your password

Store your access info for safekeeping here!

Course Name: _____

Course URL (if applicable): _____

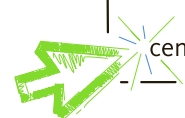
Username (Login): _____

Password: _____

GET MORE WITH THE MINDTAP MOBILE APP!

- Stay organized — never miss another assignment.
- Study whenever and wherever you want.
- Get a better grade with mobile quizzes and flashcards.

Technical Support
cengage.com/support



Additional Support
cengagebrain.com/shop/FAQ.html#Digital Products



To learn more about Cengage Learning, visit www.cengage.com

J00001

ISBN-13: 978-1-337-09419-1
ISBN-10: 1-337-09419-6



9 781337 094191

MindTap® Business Statistics for Statistics for Business & Economics, Revised 13e

Anderson/Sweeney/Williams/Camm/Cochran

ISBN: 978-1-337-09419-1



LET'S GET STARTED!

ACTIVATE INSTRUCTOR-ASSIGNED MATERIALS
FOR THIS COURSE ONLINE
USING THE ENCLOSED ACCESS CODE.

MindTap® Business Statistics, 2 Semesters
Printed Access Card for
Statistics for Business & Economics, Revised 13e

AGREEMENT URLS:

- CengageBrain Service Agreement: www.cengagebrain.com/shop/terms.html
- MindTap Service Agreement: www.cengage.com/mindtap-service-agreement
- Cengage Learning Privacy Statement: www.cengage.com/privacy

©2018 Cengage Learning

IMPORTANT: READ THIS BEFORE OPENING

The Access Code contained in this package (whether activated or not) cannot be resold or otherwise transferred, and cannot be returned once opened. MindTap and MindTap content may only be accessed within the country in which you have established your account. In addition, the Access Code and the online product or service accessible through activation of that Access Code is also governed by the additional terms contained in the agreements (which may include terms of use or terms of service) located at the Agreement URLs (Universal Resource Locator) listed above (the "Agreements"). Please read the Agreements carefully before opening this package, as opening the package will indicate your acceptance of the foregoing (including the Agreements). If you do not agree to the foregoing (including the Agreements), then Cengage is unwilling to provide the Access Code and online product or service to you, and you should promptly return the unopened package together with proof of purchase to the place of purchase and your purchase will be refunded.

CONGRATULATIONS!

You're about to unlock the content needed to ACE this course.



You only need two things to get started:

- The registration link from your instructor
- The unique access code to the right →

Have both? Good! Time to get started.

- Visit your instructor-provided registration link.
- Create an account or sign in to CengageBrain.
- When prompted, enter the access code from this package.

You now have access to your course!



Download the MindTap mobile app for convenient access anytime, anywhere! Available on both The App Store and GooglePlay.

Did something go wrong? We can help!

My instructor didn't give me a registration link.

You should ask your instructor for the registration link or visit <http://login.cengagebrain.com> and enter your access code.

Access code already in use?

If your access code came with a used book, the access code may have been used before. Access codes are single use. Return to where you purchased this product for assistance.

Need Technical Support?

See back panel for complete support information.

DO NOT DISCARD. SAVE THIS CARD TO ACCESS YOUR ONLINE PRODUCT.

REVISED EDITION

Anderson * Sweeney * Williams * Camm * Cochran



Statistics for
Business & Economics ^{13e}



Statistics for Business & Economics

13e Revised

David R. Anderson
University of Cincinnati

Dennis J. Sweeney
University of Cincinnati

Thomas A. Williams
Rochester Institute
of Technology

Jeffrey D. Camm
Wake Forest University

James J. Cochran
University of Alabama



Australia • Brazil • Mexico • Singapore • United Kingdom • United States

Statistics for Business and Economics,
Thirteenth Edition, Revised

David R. Anderson, Dennis J. Sweeney,
Thomas A. Williams, Jeffrey D. Camm,
James J. Cochran

Vice President, General Manager: Social
Science and Qualitative Business:
Erin Joyner

Product Director: Mike Schenk

Product Team Manager: Joe Sabatino

Senior Product Manager: Aaron Arnsperger

Content Developer: Anne Merrill

Senior Marketing Manager: Nate Anderson

Digital Content Designer: Brandon Foltz

Product Assistant: Audrey Jacobs

Senior Content Project Manager:
Colleen Farmer

Manufacturing Planner: Ron Montgomery

Production Service: MPS Limited

Senior Art Director: Michelle Kunkler

Internal Designer: Beckmeyer Design

Cover Designer: Beckmeyer Design

Cover Image: iStockphoto.com/alienforce

Intellectual Property

Analyst: Brittani Morgan

Project Manager: Nick Barrows

© 2018, 2015 Cengage Learning®

ALL RIGHTS RESERVED. No part of this work covered by the copyright herein may be reproduced or distributed in any form or by any means, except as permitted by U.S. copyright law, without the prior written permission of the copyright owner.

For product information and technology assistance, contact us at
Cengage Learning Customer & Sales Support, 1-800-354-9706

For permission to use material from this text or product,
submit all requests online at **www.cengage.com/permissions**

Further permissions questions can be emailed to
permissionrequest@cengage.com

Unless otherwise noted, all items © Cengage Learning.

Microsoft Excel® is a registered trademark of Microsoft Corporation.

© 2014 Microsoft.

Library of Congress Control Number: 2016954099

Package ISBN: 978-1-337-09416-0

Book only ISBN: 978-1-305-88188-4

Cengage Learning

20 Channel Center Street

Boston, MA 02210

USA

Cengage Learning is a leading provider of customized learning solutions with employees residing in nearly 40 different countries and sales in more than 125 countries around the world. Find your local representative at **www.cengage.com**

Cengage Learning products are represented in Canada by
Nelson Education, Ltd.

To learn more about Cengage Learning Solutions, visit
www.cengage.com

Purchase any of our products at your local college store or at our
preferred online store **www.cengagebrain.com**

*Dedicated to
Marcia, Cherri, Robbie, Karen, and Teresa*

Brief Contents

Preface	xxi
About the Authors	xxvi
Chapter 1	Data and Statistics 1
Chapter 2	Descriptive Statistics: Tabular and Graphical Displays 32
Chapter 3	Descriptive Statistics: Numerical Measures 102
Chapter 4	Introduction to Probability 173
Chapter 5	Discrete Probability Distributions 219
Chapter 6	Continuous Probability Distributions 271
Chapter 7	Sampling and Sampling Distributions 304
Chapter 8	Interval Estimation 348
Chapter 9	Hypothesis Tests 387
Chapter 10	Inference About Means and Proportions with Two Populations 445
Chapter 11	Inferences About Population Variances 485
Chapter 12	Comparing Multiple Proportions, Test of Independence and Goodness of Fit 509
Chapter 13	Experimental Design and Analysis of Variance 546
Chapter 14	Simple Linear Regression 600
Chapter 15	Multiple Regression 683
Chapter 16	Regression Analysis: Model Building 756
Chapter 17	Time Series Analysis and Forecasting 807
Chapter 18	Nonparametric Methods 873
Chapter 19	Statistical Methods for Quality Control 918
Chapter 20	Index Numbers 952
Chapter 21	Decision Analysis (On Website)
Chapter 22	Sample Survey (On Website)
Appendix A	References and Bibliography 974
Appendix B	Tables 976
Appendix C	Summation Notation 1003
Appendix D	Self-Test Solutions and Answers to Even-Numbered Exercises 1005
Appendix E	Microsoft Excel 2016 and Tools for Statistical Analysis 1072
Appendix F	Computing p -Values Using Minitab and Excel 1080
Index	1084

Contents

Preface	xxi
About the Authors	xxvi

Chapter 1 Data and Statistics 1

Statistics in Practice: Bloomberg Businessweek 2

1.1 Applications in Business and Economics 3

Accounting	3
Finance	4
Marketing	4
Production	4
Economics	4
Information Systems	5

1.2 Data 5

Elements, Variables, and Observations	5
Scales of Measurement	7
Categorical and Quantitative Data	8
Cross-Sectional and Time Series Data	8

1.3 Data Sources 11

Existing Sources	11
Observational Study	12
Experiment	13
Time and Cost Issues	13
Data Acquisition Errors	13

1.4 Descriptive Statistics 14

1.5 Statistical Inference 16

1.6 Analytics 17

1.7 Big Data and Data Mining 18

1.8 Computers and Statistical Analysis 20

1.9 Ethical Guidelines for Statistical Practice 20

Summary 22

Glossary 23

Supplementary Exercises 24

Chapter 2 Descriptive Statistics: Tabular and Graphical Displays 32

Statistics in Practice: Colgate-Palmolive Company 33

2.1 Summarizing Data for a Categorical Variable 34

Frequency Distribution	34
------------------------	----

	Relative Frequency and Percent Frequency Distributions	35
	Bar Charts and Pie Charts	35
2.2	Summarizing Data for a Quantitative Variable	41
	Frequency Distribution	41
	Relative Frequency and Percent Frequency Distributions	43
	Dot Plot	43
	Histogram	44
	Cumulative Distributions	45
	Stem-and-Leaf Display	46
2.3	Summarizing Data for Two Variables Using Tables	55
	Crosstabulation	55
	Simpson's Paradox	58
2.4	Summarizing Data for Two Variables Using Graphical Displays	64
	Scatter Diagram and Trendline	64
	Side-by-Side and Stacked Bar Charts	65
2.5	Data Visualization: Best Practices in Creating Effective Graphical Displays	71
	Creating Effective Graphical Displays	71
	Choosing the Type of Graphical Display	72
	Data Dashboards	72
	Data Visualization in Practice: Cincinnati Zoo and Botanical Garden	74
	Summary	77
	Glossary	78
	Key Formulas	79
	Supplementary Exercises	79
	Case Problem 1 Pelican Stores	84
	Case Problem 2 Motion Picture Industry	85
	Case Problem 3 Queen City	86
	Appendix 2.1 Using Minitab for Tabular and Graphical Presentations	87
	Appendix 2.2 Using Excel for Tabular and Graphical Presentations	90
	Chapter 3 Descriptive Statistics: Numerical Measures	102
	Statistics in Practice: Small Fry Design	103
3.1	Measures of Location	104
	Mean	104
	Weighted Mean	106
	Median	107
	Geometric Mean	109
	Mode	110
	Percentiles	111
	Quartiles	112

3.2 Measures of Variability	118
Range	118
Interquartile Range	119
Variance	119
Standard Deviation	120
Coefficient of Variation	121
3.3 Measures of Distribution Shape, Relative Location, and Detecting Outliers	125
Distribution Shape	125
z-Scores	125
Chebyshev's Theorem	127
Empirical Rule	128
Detecting Outliers	130
3.4 Five-Number Summaries and Boxplots	133
Five-Number Summary	133
Boxplot	134
Comparative Analysis Using Boxplots	135
3.5 Measures of Association Between Two Variables	138
Covariance	138
Interpretation of the Covariance	140
Correlation Coefficient	141
Interpretation of the Correlation Coefficient	143
3.6 Data Dashboards: Adding Numerical Measures to Improve Effectiveness	147
Summary	151
Glossary	152
Key Formulas	153
Supplementary Exercises	154
Case Problem 1 Pelican Stores	160
Case Problem 2 Motion Picture Industry	161
Case Problem 3 Business Schools of Asia-Pacific	162
Case Problem 4 Heavenly Chocolates Website Transactions	162
Case Problem 5 African Elephant Populations	164
Appendix 3.1 Descriptive Statistics Using Minitab	166
Appendix 3.2 Descriptive Statistics Using Excel	167

Chapter 4 Introduction to Probability 173

Statistics in Practice: National Aeronautics and Space Administration 174

4.1 Random Experiments, Counting Rules, and Assigning Probabilities 175

Counting Rules, Combinations, and Permutations 176

Assigning Probabilities 180

Probabilities for the KP&L Project 182

4.2 Events and Their Probabilities 185

4.3 Some Basic Relationships of Probability 189

Complement of an Event 189

Addition Law 190

4.4 Conditional Probability 196

Independent Events 199

Multiplication Law 199

4.5 Bayes' Theorem 204

Tabular Approach 207

Summary 210**Glossary 210****Key Formulas 211****Supplementary Exercises 212****Case Problem Hamilton County Judges 216****Chapter 5 Discrete Probability Distributions 219****Statistics in Practice: Citibank 220****5.1 Random Variables 221**

Discrete Random Variables 221

Continuous Random Variables 222

5.2 Developing Discrete Probability Distributions 224**5.3 Expected Value and Variance 229**

Expected Value 229

Variance 229

5.4 Bivariate Distributions, Covariance, and Financial Portfolios 234

A Bivariate Empirical Discrete Probability Distribution 234

Financial Applications 237

Summary 240

5.5 Binomial Probability Distribution 243

A Binomial Experiment 244

Martin Clothing Store Problem 245

Using Tables of Binomial Probabilities 249

Expected Value and Variance for the Binomial Distribution 250

5.6 Poisson Probability Distribution 254

An Example Involving Time Intervals 255

An Example Involving Length or Distance Intervals 256

5.7 Hypergeometric Probability Distribution 258**Summary 261****Glossary 262****Key Formulas 263****Supplementary Exercises 264****Case Problem *Go Bananas!* 268****Appendix 5.1 Discrete Probability Distributions with Minitab 269****Appendix 5.2 Discrete Probability Distributions with Excel 269**

Chapter 6 Continuous Probability Distributions 271

Statistics in Practice: Procter & Gamble 272

6.1 Uniform Probability Distribution 273

Area as a Measure of Probability 274

6.2 Normal Probability Distribution 277

Normal Curve 277

Standard Normal Probability Distribution 279

Computing Probabilities for Any Normal Probability Distribution 284

Great Tire Company Problem 285

6.3 Normal Approximation of Binomial Probabilities 289

6.4 Exponential Probability Distribution 293

Computing Probabilities for the Exponential Distribution 293

Relationship Between the Poisson and Exponential Distributions 294

Summary 296

Glossary 297

Key Formulas 297

Supplementary Exercises 298

Case Problem Specialty Toys 301

Appendix 6.1 Continuous Probability Distributions with Minitab 302

Appendix 6.2 Continuous Probability Distributions with Excel 303

Chapter 7 Sampling and Sampling Distributions 304

Statistics in Practice: Meadwestvaco Corporation 305

7.1 The Electronics Associates Sampling Problem 306

7.2 Selecting a Sample 307

Sampling from a Finite Population 307

Sampling from an Infinite Population 309

7.3 Point Estimation 312

Practical Advice 314

7.4 Introduction to Sampling Distributions 316

7.5 Sampling Distribution of \bar{x} 318

Expected Value of \bar{x} 319

Standard Deviation of \bar{x} 319

Form of the Sampling Distribution of \bar{x} 320

Sampling Distribution of \bar{x} for the EAI Problem 321

Practical Value of the Sampling Distribution of \bar{x} 322

Relationship Between the Sample Size and the Sampling Distribution of \bar{x} 324

7.6 Sampling Distribution of \bar{p} 328

Expected Value of \bar{p} 329

Standard Deviation of \bar{p} 329

Form of the Sampling Distribution of \bar{p} 330

Practical Value of the Sampling Distribution of \bar{p} 331

7.7 Properties of Point Estimators	334
Unbiased	334
Efficiency	335
Consistency	336
7.8 Other Sampling Methods	337
Stratified Random Sampling	337
Cluster Sampling	337
Systematic Sampling	338
Convenience Sampling	338
Judgment Sampling	339
Summary	339
Glossary	340
Key Formulas	341
Supplementary Exercises	341
Case Problem Marion Dairies	344
Appendix 7.1 The Expected Value and Standard Deviation of \bar{x}	344
Appendix 7.2 Random Sampling with Minitab	346
Appendix 7.3 Random Sampling with Excel	347
Chapter 8 Interval Estimation	348
Statistics in Practice: Food Lion	349
8.1 Population Mean: σ Known	350
Margin of Error and the Interval Estimate	350
Practical Advice	354
8.2 Population Mean: σ Unknown	356
Margin of Error and the Interval Estimate	357
Practical Advice	360
Using a Small Sample	360
Summary of Interval Estimation Procedures	362
8.3 Determining the Sample Size	365
8.4 Population Proportion	368
Determining the Sample Size	370
Summary	374
Glossary	375
Key Formulas	375
Supplementary Exercises	376
Case Problem 1 <i>Young Professional</i> Magazine	379
Case Problem 2 Gulf Real Estate Properties	380
Case Problem 3 Metropolitan Research, Inc.	380
Appendix 8.1 Interval Estimation with Minitab	382
Appendix 8.2 Interval Estimation Using Excel	384

Chapter 9 Hypothesis Tests 387

Statistics in Practice: John Morrell & Company 388

9.1 Developing Null and Alternative Hypotheses 389

The Alternative Hypothesis as a Research Hypothesis 389

The Null Hypothesis as an Assumption to Be Challenged 390

Summary of Forms for Null and Alternative Hypotheses 391

9.2 Type I and Type II Errors 392

9.3 Population Mean: σ Known 395

One-Tailed Test 395

Two-Tailed Test 401

Summary and Practical Advice 403

Relationship Between Interval Estimation and Hypothesis Testing 405

9.4 Population Mean: σ Unknown 410

One-Tailed Test 410

Two-Tailed Test 411

Summary and Practical Advice 413

9.5 Population Proportion 416

Summary 418

9.6 Hypothesis Testing and Decision Making 421

9.7 Calculating the Probability of Type II Errors 422

9.8 Determining the Sample Size for a Hypothesis Test About a Population Mean 427

Summary 430

Glossary 431

Key Formulas 432

Supplementary Exercises 432

Case Problem 1 Quality Associates, Inc. 435

Case Problem 2 Ethical Behavior of Business Students at Bayview University 437

Appendix 9.1 Hypothesis Testing with Minitab 438

Appendix 9.2 Hypothesis Testing with Excel 440

Chapter 10 Inference About Means and Proportions with Two Populations 445

Statistics in Practice: U.S. Food and Drug Administration 446

10.1 Inferences About the Difference Between Two Population Means: σ_1 and σ_2 Known 447

Interval Estimation of $\mu_1 - \mu_2$ 447

Hypothesis Tests About $\mu_1 - \mu_2$ 449

Practical Advice 451

10.2 Inferences About the Difference Between Two Population Means: σ_1 and σ_2 Unknown 454

Interval Estimation of $\mu_1 - \mu_2$ 454

Hypothesis Tests About $\mu_1 - \mu_2$	456
Practical Advice	458
10.3 Inferences About the Difference Between Two Population Means: Matched Samples	462
10.4 Inferences About the Difference Between Two Population Proportions	468
Interval Estimation of $p_1 - p_2$	468
Hypothesis Tests About $p_1 - p_2$	470
Summary	474
Glossary	474
Key Formulas	475
Supplementary Exercises	476
Case Problem Par, Inc.	479
Appendix 10.1 Inferences About Two Populations Using Minitab	480
Appendix 10.2 Inferences About Two Populations Using Excel	482
Chapter 11 Inferences About Population Variances	485
Statistics in Practice: U.S. Government Accountability Office	486
11.1 Inferences About a Population Variance	487
Interval Estimation	487
Hypothesis Testing	491
11.2 Inferences About Two Population Variances	497
Summary	504
Key Formulas	504
Supplementary Exercises	504
Case Problem Air Force Training Program	506
Appendix 11.1 Population Variances with Minitab	507
Appendix 11.2 Population Variances with Excel	508
Chapter 12 Comparing Multiple Proportions, Test of Independence and Goodness of Fit	509
Statistics in Practice: United Way	510
12.1 Testing the Equality of Population Proportions for Three or More Populations	511
A Multiple Comparison Procedure	516
12.2 Test of Independence	521
12.3 Goodness of Fit Test	529
Multinomial Probability Distribution	529
Normal Probability Distribution	532
Summary	538
Glossary	538
Key Formulas	539
Supplementary Exercises	539

Case Problem A Bipartisan Agenda for Change 542

Appendix 12.1 Chi-Square Tests Using Minitab 543

Appendix 12.2 Chi-Square Tests Using Excel 544

Chapter 13 Experimental Design and Analysis of Variance 546

Statistics in Practice: Burke Marketing Services, Inc. 547

13.1 An Introduction to Experimental Design and Analysis of Variance 548

Data Collection 549

Assumptions for Analysis of Variance 550

Analysis of Variance: A Conceptual Overview 550

13.2 Analysis of Variance and the Completely Randomized Design 553

Between-Treatments Estimate of Population Variance 554

Within-Treatments Estimate of Population Variance 555

Comparing the Variance Estimates: The F Test 556

ANOVA Table 558

Computer Results for Analysis of Variance 559

Testing for the Equality of k Population Means:
An Observational Study 560

13.3 Multiple Comparison Procedures 564

Fisher's LSD 564

Type I Error Rates 567

13.4 Randomized Block Design 570

Air Traffic Controller Stress Test 571

ANOVA Procedure 572

Computations and Conclusions 573

13.5 Factorial Experiment 577

ANOVA Procedure 579

Computations and Conclusions 579

Summary 584

Glossary 585

Key Formulas 585

Supplementary Exercises 588

Case Problem 1 Wentworth Medical Center 592

Case Problem 2 Compensation for Sales Professionals 593

Appendix 13.1 Analysis of Variance with Minitab 594

Appendix 13.2 Analysis of Variance with Excel 596

Chapter 14 Simple Linear Regression 600

Statistics in Practice: Alliance Data Systems 601

14.1 Simple Linear Regression Model 602

Regression Model and Regression Equation 602

Estimated Regression Equation 603

14.2 Least Squares Method	605
14.3 Coefficient of Determination	616
Correlation Coefficient	619
14.4 Model Assumptions	623
14.5 Testing for Significance	624
Estimate of σ^2	625
<i>t</i> Test	625
Confidence Interval for β_1	627
<i>F</i> Test	628
Some Cautions About the Interpretation of Significance Tests	630
14.6 Using the Estimated Regression Equation for Estimation and Prediction	633
Interval Estimation	634
Confidence Interval for the Mean Value of <i>y</i>	635
Prediction Interval for an Individual Value of <i>y</i>	636
14.7 Computer Solution	641
14.8 Residual Analysis: Validating Model Assumptions	645
Residual Plot Against <i>x</i>	646
Residual Plot Against \hat{y}	647
Standardized Residuals	649
Normal Probability Plot	651
14.9 Residual Analysis: Outliers and Influential Observations	654
Detecting Outliers	654
Detecting Influential Observations	656
Summary	662
Glossary	663
Key Formulas	664
Supplementary Exercises	666
Case Problem 1 Measuring Stock Market Risk	672
Case Problem 2 U.S. Department of Transportation	673
Case Problem 3 Selecting a Point-and-Shoot Digital Camera	674
Case Problem 4 Finding the Best Car Value	675
Case Problem 5 Buckeye Creek Amusement Park	676
Appendix 14.1 Calculus-Based Derivation of Least Squares Formulas	677
Appendix 14.2 A Test for Significance Using Correlation	678
Appendix 14.3 Regression Analysis with Minitab	679
Appendix 14.4 Regression Analysis with Excel	680
Chapter 15 Multiple Regression	683
Statistics in Practice: dunnhumby	684
15.1 Multiple Regression Model	685
Regression Model and Regression Equation	685
Estimated Multiple Regression Equation	685

15.2 Least Squares Method	686
An Example: Butler Trucking Company	687
Note on Interpretation of Coefficients	690
15.3 Multiple Coefficient of Determination	696
15.4 Model Assumptions	699
15.5 Testing for Significance	701
<i>F</i> Test	701
<i>t</i> Test	704
Multicollinearity	705
15.6 Using the Estimated Regression Equation for Estimation and Prediction	708
15.7 Categorical Independent Variables	711
An Example: Johnson Filtration, Inc.	711
Interpreting the Parameters	713
More Complex Categorical Variables	715
15.8 Residual Analysis	720
Detecting Outliers	722
Studentized Deleted Residuals and Outliers	722
Influential Observations	723
Using Cook's Distance Measure to Identify Influential Observations	723
15.9 Logistic Regression	727
Logistic Regression Equation	728
Estimating the Logistic Regression Equation	729
Testing for Significance	732
Managerial Use	732
Interpreting the Logistic Regression Equation	733
Logit Transformation	736
Summary	740
Glossary	740
Key Formulas	741
Supplementary Exercises	743
Case Problem 1 Consumer Research, Inc.	750
Case Problem 2 Predicting Winnings for NASCAR Drivers	751
Case Problem 3 Finding the Best Car Value	752
Appendix 15.1 Multiple Regression with Minitab	753
Appendix 15.2 Multiple Regression with Excel	753
Appendix 15.3 Logistic Regression with Minitab	755
Chapter 16 Regression Analysis: Model Building	756
Statistics in Practice: Monsanto Company	757
16.1 General Linear Model	758
Modeling Curvilinear Relationships	758
Interaction	761

Transformations Involving the Dependent Variable	765
Nonlinear Models That Are Intrinsically Linear	769
16.2 Determining When to Add or Delete Variables	773
General Case	775
Use of p -Values	776
16.3 Analysis of a Larger Problem	780
16.4 Variable Selection Procedures	784
Stepwise Regression	784
Forward Selection	786
Backward Elimination	786
Best-Subsets Regression	787
Making the Final Choice	788
16.5 Multiple Regression Approach to Experimental Design	790
16.6 Autocorrelation and the Durbin-Watson Test	795
Summary	799
Glossary	800
Key Formulas	800
Supplementary Exercises	800
Case Problem 1 Analysis of PGA Tour Statistics	803
Case Problem 2 Rating Wines from the Piedmont Region of Italy	804
Appendix 16.1 Variable Selection Procedures with Minitab	805
Chapter 17 Time Series Analysis and Forecasting	807
Statistics in Practice: Nevada Occupational Health Clinic	808
17.1 Time Series Patterns	809
Horizontal Pattern	809
Trend Pattern	811
Seasonal Pattern	811
Trend and Seasonal Pattern	812
Cyclical Pattern	812
Selecting a Forecasting Method	814
17.2 Forecast Accuracy	815
17.3 Moving Averages and Exponential Smoothing	820
Moving Averages	820
Weighted Moving Averages	823
Exponential Smoothing	823
17.4 Trend Projection	830
Linear Trend Regression	830
Nonlinear Trend Regression	835
17.5 Seasonality and Trend	841
Seasonality Without Trend	841
Seasonality and Trend	843
Models Based on Monthly Data	846

17.6 Time Series Decomposition 850

Calculating the Seasonal Indexes 851

Deseasonalizing the Time Series 855

Using the Deseasonalized Time Series to Identify Trend 855

Seasonal Adjustments 857

Models Based on Monthly Data 857

Cyclical Component 857

Summary 860**Glossary 861****Key Formulas 862****Supplementary Exercises 862****Case Problem 1 Forecasting Food and Beverage Sales 866****Case Problem 2 Forecasting Lost Sales 867****Appendix 17.1 Forecasting with Minitab 868****Appendix 17.2 Forecasting with Excel 871****Chapter 18 Nonparametric Methods 873****Statistics in Practice: West Shell Realtors 874****18.1 Sign Test 875**

Hypothesis Test About a Population Median 875

Hypothesis Test with Matched Samples 880

18.2 Wilcoxon Signed-Rank Test 883**18.3 Mann-Whitney-Wilcoxon Test 888****18.4 Kruskal-Wallis Test 899****18.5 Rank Correlation 903****Summary 908****Glossary 908****Key Formulas 909****Supplementary Exercises 910****Appendix 18.1 Nonparametric Methods with Minitab 913****Appendix 18.2 Nonparametric Methods with Excel 915****Chapter 19 Statistical Methods for Quality Control 918****Statistics in Practice: Dow Chemical Company 919****19.1 Philosophies and Frameworks 920**

Malcolm Baldrige National Quality Award 921

ISO 9000 921

Six Sigma 921

Quality in the Service Sector 924

19.2 Statistical Process Control 924

Control Charts 925

 \bar{x} Chart: Process Mean and Standard Deviation Known 926

\bar{x} Chart: Process Mean and Standard Deviation Unknown 928
R Chart 931
p Chart 933
np Chart 935
 Interpretation of Control Charts 935

19.3 Acceptance Sampling 938

KALI, Inc.: An Example of Acceptance Sampling 939
 Computing the Probability of Accepting a Lot 940
 Selecting an Acceptance Sampling Plan 943
 Multiple Sampling Plans 945

Summary 946

Glossary 946

Key Formulas 947

Supplementary Exercises 948

Appendix 19.1 Control Charts with Minitab 950

Chapter 20 Index Numbers 952

Statistics in Practice: U.S. Department of Labor, Bureau of Labor Statistics 953

20.1 Price Relatives 954

20.2 Aggregate Price Indexes 954

20.3 Computing an Aggregate Price Index from Price Relatives 958

20.4 Some Important Price Indexes 960

Consumer Price Index 960

Producer Price Index 960

Dow Jones Averages 961

20.5 Deflating a Series by Price Indexes 962

20.6 Price Indexes: Other Considerations 965

Selection of Items 965

Selection of a Base Period 965

Quality Changes 966

20.7 Quantity Indexes 966

Summary 968

Glossary 968

Key Formulas 969

Supplementary Exercises 969

Chapter 21 Decision Analysis (On Website)

Statistics in Practice: Ohio Edison Company 21-2

21.1 Problem Formulation 21-3

Payoff Tables 21-4

Decision Trees 21-4

21.2 Decision Making with Probabilities 21-5

Expected Value Approach	21-5
Expected Value of Perfect Information	21-7
21.3 Decision Analysis with Sample Information	21-13
Decision Tree	21-14
Decision Strategy	21-15
Expected Value of Sample Information	21-18
21.4 Computing Branch Probabilities Using Bayes' Theorem	21-24
Summary	21-28
Glossary	21-29
Key Formulas	21-30
Supplementary Exercises	21-30
Case Problem Lawsuit Defense Strategy	21-33
Appendix: Self-Test Solutions and Answers to Even-Numbered Exercises	21-34

Chapter 22 **Sample Survey (On Website)**

Statistics in Practice: Duke Energy	22-2
22.1 Terminology Used in Sample Surveys	22-2
22.2 Types of Surveys and Sampling Methods	22-3
22.3 Survey Errors	22-5
Nonsampling Error	22-5
Sampling Error	22-5
22.4 Simple Random Sampling	22-6
Population Mean	22-6
Population Total	22-7
Population Proportion	22-8
Determining the Sample Size	22-9
22.5 Stratified Simple Random Sampling	22-12
Population Mean	22-12
Population Total	22-14
Population Proportion	22-15
Determining the Sample Size	22-16
22.6 Cluster Sampling	22-21
Population Mean	22-23
Population Total	22-25
Population Proportion	22-25
Determining the Sample Size	22-27
22.7 Systematic Sampling	22-29
Summary	22-29
Glossary	22-30
Key Formulas	22-30
Supplementary Exercises	22-34

Appendix A	References and Bibliography	974
Appendix B	Tables	976
Appendix C	Summation Notation	1003
Appendix D	Self-Test Solutions and Answers to Even-Numbered Exercises	1005
Appendix E	Microsoft Excel 2016 and Tools for Statistical Analysis	1072
Appendix F	Computing p-Values Using Minitab and Excel	1080
Index		1084

Preface

This text is the revised 13th edition of *STATISTICS FOR BUSINESS AND ECONOMICS*. The revised edition updates the material in *STATISTICS FOR BUSINESS ECONOMICS* 13e for use with Microsoft Excel 2016 and Minitab 17. Current users of the 13th edition will find changes to the chapter-ending appendices, which now describe Excel 2016 and Minitab 17 procedures. In addition to the updated the chapter-ending appendices, we have updated the appendix to the book entitled Microsoft Excel 2016 and Tools for Statistical Analysis. This appendix provides an introduction to Excel 2016 and its tools for statistical analysis. Several of Excel's statistical functions have been upgraded and improved.

The remainder of this preface describes the authors' objectives in writing *STATISTICS FOR BUSINESS AND ECONOMICS* and the major changes that were made in developing the 13th edition. The purpose of the text is to give students, primarily those in the fields of business administration and economics, a conceptual introduction to the field of statistics and its many applications. The text is applications-oriented and written with the needs of the nonmathematician in mind; the mathematical prerequisite is understanding of algebra.

Applications of data analysis and statistical methodology are an integral part of the organization and presentation of the text material. The discussion and development of each technique is presented in an application setting, with the statistical results providing insights to decisions and solutions to problems.

Although the book is applications oriented, we have taken care to provide sound methodological development and to use notation that is generally accepted for the topic being covered. Hence, students will find that this text provides good preparation for the study of more advanced statistical material. A bibliography to guide further study is included as an appendix.

The text introduces the student to the software packages of Minitab 17 and Microsoft® Office Excel 2016 and emphasizes the role of computer software in the application of statistical analysis. Minitab is illustrated as it is one of the leading statistical software packages for both education and statistical practice. Excel is not a statistical software package, but the wide availability and use of Excel make it important for students to understand the statistical capabilities of this package. Minitab and Excel procedures are provided in appendices so that instructors have the flexibility of using as much computer emphasis as desired for the course.

Changes in the Thirteenth Edition

We appreciate the acceptance and positive response to the previous editions of *Statistics for Business and Economics*. Accordingly, in making modifications for this new edition, we have maintained the presentation style and readability of those editions. There have been many changes made throughout the text to enhance its educational effectiveness. The most substantial changes in the new edition are summarized here.

Content Revisions

- **Data and Statistics—Chapter 1.** We have expanded our section on data mining to include a discussion of big data. We have added a new section on analytics. We have also placed greater emphasis on the distinction between observed and experimental data.

- **Descriptive Statistics: Tabular and Graphical Displays—Chapter 2.** We have added instructions on how to use Excel’s recommended charts option to Appendix 2.2 at the end of this chapter. This new Excel functionality produces a gallery of suggested charts based on the data selected by the user and can help students identify the most appropriate chart(s) to use to depict their data.
- **Descriptive Statistics: Numerical Measures—Chapter 3.** We now use the method for calculating percentiles that is recommended by the National Institute of Standards and Technology (NIST). In addition to being the standard recommended by NIST, this approach is also used by a wide variety of software. The NIST recommended approach for calculating percentiles is used throughout the textbook wherever percentiles are used (for example, when creating a box plot or when calculating quantiles or an interquartile range).
- **Introduction to Probability—Chapter 4.** The discussion on experiments has been updated to draw a more clear distinction between random and designed experiments. This distinction makes it easier to understand the differences in the discussion of experiments in the probability chapters (Chapters 4, 5, and 6) and the experimental design chapter (Chapter 13).
- **Software.** We have revised all step-by-step instructions in the software appendices and all figures throughout the book that feature software output to reflect Excel 2016 and Minitab 17. This provides students exposure to and experience with the current versions of two of the most commonly used software for statistical analysis in business. In this latest edition, we no longer provide discussion of the use of StatTools.
- **Case Problems.** We have added two new case problems in this addition; the total number of cases is 33. One new probability modeling case has been added to Chapter 5, and one new simple linear regression case appears in Chapter 14. The 33 case problems in this book provide students the opportunity to work on more complex problems, analyze larger data sets, and prepare managerial reports based on the results of their analyses.
- **Examples and Exercises Based on Real Data.** We continue to make a substantial effort to update our text examples and exercises with the most current real data and referenced sources of statistical information. In this edition, we have added more than 180 new examples and exercises based on real data and referenced sources. Using data from sources also used by *The Wall Street Journal*, *USA Today*, *Barron’s*, and others, we have drawn from actual studies and applications to develop explanations and create exercises that demonstrate the many uses of statistics in business and economics. We believe that the use of real data from interesting and relevant problems helps generate more student interest in the material and enables the student to learn about both statistical methodology and its application. The 13th edition contains more than 350 examples and exercises based on real data.

Features and Pedagogy

Authors Anderson, Sweeney, Williams, Camm, and Cochran have continued many of the features that appeared in previous editions. Important ones for students are noted here.

Methods Exercises and Applications Exercises

The end-of-section exercises are split into two parts, Methods and Applications. The Methods exercises require students to use the formulas and make the necessary computations. The Applications exercises require students to use the chapter material in real-world situations. Thus, students first focus on the computational “nuts and bolts” and then move on to the subtleties of statistical application and interpretation.

Self-Test Exercises

Certain exercises are identified as “Self-Test Exercises.” Completely worked-out solutions for these exercises are provided in Appendix D. Students can attempt the Self-Test Exercises and immediately check the solution to evaluate their understanding of the concepts presented in the chapter.

Margin Annotations and Notes and Comments

Margin annotations that highlight key points and provide additional insights for the student are a key feature of this text. These annotations, which appear in the margins, are designed to provide emphasis and enhance understanding of the terms and concepts being presented in the text.

At the end of many sections, we provide Notes and Comments designed to give the student additional insights about the statistical methodology and its application. Notes and Comments include warnings about or limitations of the methodology, recommendations for application, brief descriptions of additional technical considerations, and other matters.

Data Files Accompany the Text

Over 200 data files are available on the website that accompanies the text. In previous editions, we provided data files in both Excel and Minitab formats. In this edition, to be more efficient, we provide the data files in only one format, CSV (comma separated values). In the appendices to Chapter 2, we provide instructions on how to open CSV files in both Excel and Minitab. DATAfile logos are used in the text to identify the data sets that are available on the website. Data sets for all case problems as well as data sets for larger exercises are included. In this edition, instead of supplying both Minitab and Excel data files, we provide data files in a single format (CSV format). This format is accessible to both Minitab and Excel. We give step-by-step instructions on how to open these files in Minitab and Excel in Appendices 2.1 and 2.2 at the end of Chapter 2.

Acknowledgments

We would like to acknowledge the work of our reviewers, who provided comments and suggestions of ways to continue to improve our text. Thanks to

AbouEl-Makarim Aboueissa, University of Southern Maine	Margaret E. Cochran Northwestern State University of Louisiana	Jim Grant Lewis & Clark College
Kathleen Arano Fort Hays State University	Thomas A. Dahlstrom Eastern University	Reidar Hagtvedt University of Alberta School of Business
Musa Ayar Uw-baraboo/Sauk County	Anne Drougas Dominican University	Clifford B. Hawley West Virginia University
Kathleen Burke SUNY Cortland	Fesseha Gebremikael Strayer University/ Calhoun Community College	Vance A. Hughey Western Nevada College
YC Chang University of Notre Dame	Malcolm C. Gold University of Wisconsin— Marshfield/Wood County	Tony Hunnicutt Ouachita Technical College
David Chen Rosemont College and Saint Joseph’s University	Joel Goldstein Western Connecticut State University	Stacey M. Jones Albers School of Business and Economics, Seattle University

Dukpa Kim University of Virginia	Claudiney Pereira Tulane University	Jennifer Strehler Oakton Community College
Rajaram Krishnan Earlham College	J. G. Pitt University of Toronto	Ronald Stunda Valdosta State University
Robert J. Lemke Lake Forest College	Scott A. Redenius Brandeis University	Cindy van Es Cornell University
Philip J. Mizzi Arizona State University	Sandra Robertson Thomas Nelson Community College	Jennifer VanGilder Ursinus College
Mehdi Mohaghegh Norwich University	Sunil Sapra California State University, Los Angeles	Jacqueline Wroughton Northern Kentucky University
Mihail Motzev Walla Walla University	Kyle Vann Scott Snead State Community College	Dmitry Yarushkin Grand View University
Somnath Mukhopadhyay The University of Texas at El Paso	Rodney E. Stanley Tennessee State University	David Zimmer Western Kentucky University
Kenneth E. Murphy Chapman University		
Ogbonnaya John Nwoha Grambling State University		

We continue to owe a debt to our many colleagues and friends for their helpful comments and suggestions in the development of this and earlier editions of our text. Among them are:

Mohammad Ahmadi University of Tennessee at Chattanooga	Robert Carver Stonehill College	Paul Guy California State University—Chico
Lari Arjomand Clayton College and State University	Richard Claycombe McDaniel College	Clifford Hawley West Virginia University
Robert Balough Clarion University	Robert Cochran University of Wyoming	Jim Hightower California State University, Fullerton
Philip Boudreaux University of Louisiana	Robert Collins Marquette University	Alan Humphrey University of Rhode Island
Mike Bourke Houston Baptist University	David W. Cravens Texas Christian University	Ann Hussein Philadelphia College of Textiles and Science
James Brannon University of Wisconsin— Oshkosh	Tom Dahlstrom Eastern College	C. Thomas Innis University of Cincinnati
John Bryant University of Pittsburgh	Gopal Dorai William Patterson University	Ben Isselhardt Rochester Institute of Technology
Peter Bryant University of Colorado	Nicholas Farnum California State University—Fullerton	Jeffery Jarrett University of Rhode Island
Terri L. Byczkowski University of Cincinnati	Donald Gren Salt Lake Community College	Ronald Klimberg St. Joseph's University

David A. Kravitz George Mason University	Roger Myerson Northwestern University	Hroki Tsurumi Rutgers University
David Krueger St. Cloud State University	Richard O’Connell Miami University of Ohio	David Tufte University of New Orleans
John Leschke University of Virginia	Alan Olinsky Bryant College	Victor Ukpolo Austin Peay State University
Martin S. Levy University of Cincinnati	Ceyhun Ozgur Valparaiso University	Ebenge Usip Youngstown State University
John S. Loucks St. Edward’s University	Tom Pray Rochester Institute of Technology	Cindy Van Es Cornell University
David Lucking-Reiley Vanderbilt University	Harold Rahmlow St. Joseph’s University	Jack Vaughn University of Texas-El Paso
Bala Maniam Sam Houston State University	H. V. Ramakrishna Penn State University at Great Valley	Andrew Welki John Carroll University
Don Marx University of Alaska, Anchorage	Tom Ryan Case Western Reserve University	Ari Wijetunga Morehead State University
Tom McCullough University of California— Berkeley	Bill Seaver University of Tennessee	J. E. Willis Louisiana State University
Ronald W. Michener University of Virginia	Alan Smith Robert Morris College	Mustafa Yilmaz Northeastern University
Glenn Milligan Ohio State University	Willbann Terpening Gonzaga University	Gary Yoshimoto St. Cloud State University
Mitchell Muesham Sam Houston State University	Ted Tsukahara St. Mary’s College of California	Yan Yu University of Cincinnati
		Charles Zimmerman Robert Morris College

We thank our associates from business and industry who supplied the Statistics in Practice features. We recognize them individually by a credit line in each of the articles. We are also indebted to our product manager, Aaron Arnsparger; our content developer, Anne Merrill; our content project manager, Colleen Farmer; our project manager at MPS Limited, Manoj Kumar; digital content designer, Brandon Foltz; and others at Cengage for their editorial counsel and support during the preparation of this text.

David R. Anderson
Dennis J. Sweeney
Thomas A. Williams
Jeffrey D. Camm
James J. Cochran

About the Authors

David R. Anderson. David R. Anderson is Professor Emeritus of Quantitative Analysis in the College of Business Administration at the University of Cincinnati. Born in Grand Forks, North Dakota, he earned his B.S., M.S., and Ph.D. degrees from Purdue University. Professor Anderson has served as Head of the Department of Quantitative Analysis and Operations Management and as Associate Dean of the College of Business Administration at the University of Cincinnati. In addition, he was the coordinator of the College's first Executive Program.

At the University of Cincinnati, Professor Anderson has taught introductory statistics for business students as well as graduate-level courses in regression analysis, multivariate analysis, and management science. He has also taught statistical courses at the Department of Labor in Washington, D.C. He has been honored with nominations and awards for excellence in teaching and excellence in service to student organizations.

Professor Anderson has coauthored 10 textbooks in the areas of statistics, management science, linear programming, and production and operations management. He is an active consultant in the field of sampling and statistical methods.

Dennis J. Sweeney. Dennis J. Sweeney is Professor Emeritus of Quantitative Analysis and Founder of the Center for Productivity Improvement at the University of Cincinnati. Born in Des Moines, Iowa, he earned a B.S.B.A. degree from Drake University and his M.B.A. and D.B.A. degrees from Indiana University, where he was an NDEA Fellow. Professor Sweeney has worked in the management science group at Procter & Gamble and spent a year as a visiting professor at Duke University. Professor Sweeney served as Head of the Department of Quantitative Analysis and as Associate Dean of the College of Business Administration at the University of Cincinnati.

Professor Sweeney has published more than 30 articles and monographs in the area of management science and statistics. The National Science Foundation, IBM, Procter & Gamble, Federated Department Stores, Kroger, and Cincinnati Gas & Electric have funded his research, which has been published in *Management Science*, *Operations Research*, *Mathematical Programming*, *Decision Sciences*, and other journals.

Professor Sweeney has coauthored 10 textbooks in the areas of statistics, management science, linear programming, and production and operations management.

Thomas A. Williams. Thomas A. Williams is Professor Emeritus of Management Science in the College of Business at Rochester Institute of Technology. Born in Elmira, New York, he earned his B.S. degree at Clarkson University. He did his graduate work at Rensselaer Polytechnic Institute, where he received his M.S. and Ph.D. degrees.

Before joining the College of Business at RIT, Professor Williams served for seven years as a faculty member in the College of Business Administration at the University of Cincinnati, where he developed the undergraduate program in Information Systems and then served as its coordinator. At RIT he was the first chairman of the Decision Sciences Department. He teaches courses in management science and statistics, as well as graduate courses in regression and decision analysis.

Professor Williams is the coauthor of 11 textbooks in the areas of management science, statistics, production and operations management, and mathematics. He has been a consultant for numerous *Fortune* 500 companies and has worked on projects ranging from the use of data analysis to the development of large-scale regression models.

Jeffrey D. Camm. Jeffrey D. Camm is the Inmar Presidential Chair and Associate Dean of Analytics in the School of Business at Wake Forest University. Born in Cincinnati, Ohio, he holds a B.S. from Xavier University (Ohio) and a Ph.D. from Clemson University. Prior to joining the faculty at Wake Forest, he was on the faculty of the University of Cincinnati. He has also been a visiting scholar at Stanford University and a visiting professor of business administration at the Tuck School of Business at Dartmouth College.

Dr. Camm has published over 30 papers in the general area of optimization applied to problems in operations management and marketing. He has published his research in *Science*, *Management Science*, *Operations Research*, *Interfaces*, and other professional journals. Dr. Camm was named the Dornoff Fellow of Teaching Excellence at the University of Cincinnati and he was the 2006 recipient of the INFORMS Prize for the Teaching of Operations Research Practice. A firm believer in practicing what he preaches, he has served as an operations research consultant to numerous companies and government agencies. From 2005 to 2010 he served as editor-in-chief of *Interfaces* and is currently on the editorial board of *INFORMS Transactions on Education*.

James J. Cochran. James J. Cochran is Professor of Applied Statistics and the Rogers-Spivey Faculty Fellow at the University of Alabama. Born in Dayton, Ohio, he earned his B.S., M.S., and M.B.A. degrees from Wright State University and a Ph.D. from the University of Cincinnati. He has been at the University of Alabama since 2014 and has been a visiting scholar at Stanford University, Universidad de Talca, the University of South Africa and Pole Universitaire Leonard de Vinci.

Professor Cochran has published over three dozen papers in the development and application of operations research and statistical methods. He has published his research in *Management Science*, *The American Statistician*, *Communications in Statistics—Theory and Methods*, *Annals of Operations Research*, *European Journal of Operational Research*, *Journal of Combinatorial Optimization*, *Interfaces*, *Statistics and Probability Letters*, and other professional journals. He was the 2008 recipient of the INFORMS Prize for the Teaching of Operations Research Practice and the 2010 recipient of the Mu Sigma Rho Statistical Education Award. Professor Cochran was elected to the International Statistics Institute in 2005 and named a Fellow of the American Statistical Association in 2011. He also received the Founders Award in 2014 and the Karl E. Peace Award in 2015 from the American Statistical Association. A strong advocate for effective statistics and operations research education as a means of improving the quality of applications to real problems, Professor Cochran has organized and chaired teaching effectiveness workshops in Montevideo, Uruguay; Cape Town, South Africa; Cartagena, Colombia; Jaipur, India; Buenos Aires, Argentina; Nairobi, Kenya; Buea, Cameroon; Kathmandu, Nepal; Osijek, Croatia; Havana, Cuba; and Ulaanbaatar, Mongolia. He has served as an operations research consultant to numerous companies and not-for-profit organizations. He served as editor-in-chief of *INFORMS Transactions on Education* from 2006 to 2012 and is on the editorial board of *Interfaces*, *International Transactions in Operational Research*, and *Significance*.

CHAPTER 1

Data and Statistics

CONTENTS

STATISTICS IN PRACTICE:
BLOOMBERG BUSINESSWEEK

- 1.1** APPLICATIONS IN BUSINESS AND ECONOMICS
 - Accounting
 - Finance
 - Marketing
 - Production
 - Economics
 - Information Systems
- 1.2** DATA
 - Elements, Variables, and Observations
 - Scales of Measurement
 - Categorical and Quantitative Data
 - Cross-Sectional and Time Series Data
- 1.3** DATA SOURCES
 - Existing Sources
 - Observational Study
 - Experiment
 - Time and Cost Issues
 - Data Acquisition Errors
- 1.4** DESCRIPTIVE STATISTICS
- 1.5** STATISTICAL INFERENCE
- 1.6** ANALYTICS
- 1.7** BIG DATA AND DATA MINING
- 1.8** COMPUTERS AND STATISTICAL ANALYSIS
- 1.9** ETHICAL GUIDELINES FOR STATISTICAL PRACTICE

STATISTICS *in* **PRACTICE**
BLOOMBERG BUSINESSWEEK*
 NEW YORK, NEW YORK

With a global circulation of more than 1 million, *Bloomberg Businessweek* is one of the most widely read business magazines in the world. Bloomberg's 1700 reporters in 145 service bureaus around the world enable *Bloomberg Businessweek* to deliver a variety of articles of interest to the global business and economic community. Along with feature articles on current topics, the magazine contains articles on international business, economic analysis, information processing, and science and technology. Information in the feature articles and the regular sections helps readers stay abreast of current developments and assess the impact of those developments on business and economic conditions.

Most issues of *Bloomberg Businessweek*, formerly *BusinessWeek*, provide an in-depth report on a topic of current interest. Often, the in-depth reports contain statistical facts and summaries that help the reader understand the business and economic information. Examples of articles and reports include the impact of businesses moving important work to cloud computing, the crisis facing the U.S. Postal Service, and why the debt crisis is even worse than we think. In addition, *Bloomberg Businessweek* provides a variety of statistics about the state of the economy, including production indexes, stock prices, mutual funds, and interest rates.

Bloomberg Businessweek also uses statistics and statistical information in managing its own business. For example, an annual survey of subscribers helps the company learn about subscriber demographics, reading habits, likely purchases, lifestyles, and so on. *Bloomberg Businessweek* managers use statistical summaries from the survey to provide better services to subscribers and advertisers. One recent North American subscriber survey indicated that 90% of *Bloomberg Businessweek*

*The authors are indebted to Charlene Trentham, Research Manager, for providing this Statistics in Practice.



Bloomberg Businessweek uses statistical facts and summaries in many of its articles.

subscribers use a personal computer at home and that 64% of *Bloomberg Businessweek* subscribers are involved with computer purchases at work. Such statistics alert *Bloomberg Businessweek* managers to subscriber interest in articles about new developments in computers. The results of the subscriber survey are also made available to potential advertisers. The high percentage of subscribers using personal computers at home and the high percentage of subscribers involved with computer purchases at work would be an incentive for a computer manufacturer to consider advertising in *Bloomberg Businessweek*.

In this chapter, we discuss the types of data available for statistical analysis and describe how the data are obtained. We introduce descriptive statistics and statistical inference as ways of converting data into meaningful and easily interpreted statistical information.

Frequently, we see the following types of statements in newspapers and magazines:

- Against the U.S. dollar, the euro has lost nearly 30% of its value in the last year; the Australian dollar lost almost 20% (*The Economist*, April 25th–May 1st, 2015).
- A survey conducted by the Pew Research Center reported that 68% of Internet users believe current laws are not good enough in protecting people's privacy online (*The Wall Street Journal*, March 24, 2014).

- VW Group's U.S. sales continue to slide, with total sales off by 13% from last January, to 36,930 vehicles (*Panorama*, March 2014).
- A poll of 1,320 corporate recruiters indicated that 68% of the recruiters ranked communication skills as one of the top five most important skills for new hires (*Bloomberg Businessweek* April 13–April 19, 2015).
- The California State Teachers' Retirement System has \$154.3 billion under management (*Bloomberg Businessweek*, January 21–January 27, 2013).
- At a Sotheby's art auction held on February 5, 2013, Pablo Picasso's painting *Woman Sitting Near a Window* sold for \$45 million (*The Wall Street Journal*, February 15, 2013).
- Over the past three months, the industry average for sales incentives per vehicle by GM, Chrysler, Ford, Toyota, and Honda was \$2336 (*The Wall Street Journal*, February 14, 2013).

The numerical facts in the preceding statements—30%, 20%, 68%, 13%, 36,930, 1320, 68%, \$154.3 billion, \$45 million, \$2336—are called **statistics**. In this usage, the term *statistics* refers to numerical facts such as averages, medians, percentages, and maximums that help us understand a variety of business and economic situations. However, as you will see, the field, or subject, of statistics involves much more than numerical facts. In a broader sense, statistics is the art and science of collecting, analyzing, presenting, and interpreting data. Particularly in business and economics, the information provided by collecting, analyzing, presenting, and interpreting data gives managers and decision makers a better understanding of the business and economic environment and thus enables them to make more informed and better decisions. In this text, we emphasize the use of statistics for business and economic decision making.

Chapter 1 begins with some illustrations of the applications of statistics in business and economics. In Section 1.2 we define the term *data* and introduce the concept of a data set. This section also introduces key terms such as *variables* and *observations*, discusses the difference between quantitative and categorical data, and illustrates the uses of cross-sectional and time series data. Section 1.3 discusses how data can be obtained from existing sources or through survey and experimental studies designed to obtain new data. The important role that the Internet now plays in obtaining data is also highlighted. The uses of data in developing descriptive statistics and in making statistical inferences are described in Sections 1.4 and 1.5. The last four sections of Chapter 1 provide an introduction to business analytics and the role statistics plays in it, an introduction to big data and data mining, the role of the computer in statistical analysis, and a discussion of ethical guidelines for statistical practice.

1.1

Applications in Business and Economics

In today's global business and economic environment, anyone can access vast amounts of statistical information. The most successful managers and decision makers understand the information and know how to use it effectively. In this section, we provide examples that illustrate some of the uses of statistics in business and economics.

Accounting

Public accounting firms use statistical sampling procedures when conducting audits for their clients. For instance, suppose an accounting firm wants to determine whether the amount of accounts receivable shown on a client's balance sheet fairly represents the actual amount of accounts receivable. Usually the large number of individual accounts receivable makes

reviewing and validating every account too time-consuming and expensive. As common practice in such situations, the audit staff selects a subset of the accounts called a sample. After reviewing the accuracy of the sampled accounts, the auditors draw a conclusion as to whether the accounts receivable amount shown on the client's balance sheet is acceptable.

Finance

Financial analysts use a variety of statistical information to guide their investment recommendations. In the case of stocks, analysts review financial data such as price/earnings ratios and dividend yields. By comparing the information for an individual stock with information about the stock market averages, an analyst can begin to draw a conclusion as to whether the stock is a good investment. For example, *The Wall Street Journal* (June 6, 2015) reported that the average dividend yield for the S&P 500 companies was 2%. Microsoft showed a dividend yield of 1.95%. In this case, the statistical information on dividend yield indicates a lower dividend yield for Microsoft than the average dividend yield for the S&P 500 companies. This and other information about Microsoft would help the analyst make an informed buy, sell, or hold recommendation for Microsoft stock.

Marketing

Electronic scanners at retail checkout counters collect data for a variety of marketing research applications. For example, data suppliers such as ACNielsen and Information Resources, Inc., purchase point-of-sale scanner data from grocery stores, process the data, and then sell statistical summaries of the data to manufacturers. Manufacturers spend hundreds of thousands of dollars per product category to obtain this type of scanner data. Manufacturers also purchase data and statistical summaries on promotional activities such as special pricing and the use of in-store displays. Brand managers can review the scanner statistics and the promotional activity statistics to gain a better understanding of the relationship between promotional activities and sales. Such analyses often prove helpful in establishing future marketing strategies for the various products.

Production

Today's emphasis on quality makes quality control an important application of statistics in production. A variety of statistical quality control charts are used to monitor the output of a production process. In particular, an \bar{x} -bar chart can be used to monitor the average output. Suppose, for example, that a machine fills containers with 12 ounces of a soft drink. Periodically, a production worker selects a sample of containers and computes the average number of ounces in the sample. This average, or \bar{x} -bar value, is plotted on an \bar{x} -bar chart. A plotted value above the chart's upper control limit indicates overfilling, and a plotted value below the chart's lower control limit indicates underfilling. The process is termed "in control" and allowed to continue as long as the plotted \bar{x} -bar values fall between the chart's upper and lower control limits. Properly interpreted, an \bar{x} -bar chart can help determine when adjustments are necessary to correct a production process.

Economics

Economists frequently provide forecasts about the future of the economy or some aspect of it. They use a variety of statistical information in making such forecasts. For instance, in forecasting inflation rates, economists use statistical information on such indicators as the Producer Price Index, the unemployment rate, and manufacturing capacity utilization. Often these statistical indicators are entered into computerized forecasting models that predict inflation rates.

Information Systems

Information systems administrators are responsible for the day-to-day operation of an organization's computer networks. A variety of statistical information helps administrators assess the performance of computer networks, including local area networks (LANs), wide area networks (WANs), network segments, intranets, and other data communication systems. Statistics such as the mean number of users on the system, the proportion of time any component of the system is down, and the proportion of bandwidth utilized at various times of the day are examples of statistical information that help the system administrator better understand and manage the computer network.

Applications of statistics such as those described in this section are an integral part of this text. Such examples provide an overview of the breadth of statistical applications. To supplement these examples, practitioners in the fields of business and economics provided chapter-opening Statistics in Practice articles that introduce the material covered in each chapter. The Statistics in Practice applications show the importance of statistics in a wide variety of business and economic situations.

1.2

Data

Data are the facts and figures collected, analyzed, and summarized for presentation and interpretation. All the data collected in a particular study are referred to as the **data set** for the study. Table 1.1 shows a data set containing information for 60 nations that participate in the World Trade Organization. The World Trade Organization encourages the free flow of international trade and provides a forum for resolving trade disputes.

Elements, Variables, and Observations

Elements are the entities on which data are collected. Each nation listed in Table 1.1 is an element with the nation or element name shown in the first column. With 60 nations, the data set contains 60 elements.

A **variable** is a characteristic of interest for the elements. The data set in Table 1.1 includes the following five variables:

- **WTO Status:** The nation's membership status in the World Trade Organization; this can be either as a member or an observer.
- **Per Capita GDP (\$):** The total market value (\$) of all goods and services produced by the nation divided by the number of people in the nation; this is commonly used to compare economic productivity of the nations.
- **Trade Deficit (\$1000s):** The difference between the total dollar value of the nation's imports and the total dollar value of the nation's exports.
- **Fitch Rating:** The nation's sovereign credit rating as appraised by the Fitch Group¹; the credit ratings range from a high of AAA to a low of F and can be modified by + or -.
- **Fitch Outlook:** An indication of the direction the credit rating is likely to move over the upcoming two years; the outlook can be negative, stable, or positive.

Measurements collected on each variable for every element in a study provide the data. The set of measurements obtained for a particular element is called an **observation**. Referring to Table 1.1, we see that the first observation (Armenia) contains the following measurements:

¹The Fitch Group is one of three nationally recognized statistical rating organizations designated by the U.S. Securities and Exchange Commission. The other two are Standard and Poor's and Moody's investor service.

TABLE 1.1 DATA SET FOR 60 NATIONS IN THE WORLD TRADE ORGANIZATION

DATA file
Nations

Data sets such as Nations are available on the website for this text.

Nation	WTO Status	Per Capita GDP (\$)	Trade Deficit (\$1000s)	Fitch Rating	Fitch Outlook
Armenia	Member	5,400	2,673,359	BB-	Stable
Australia	Member	40,800	-33,304,157	AAA	Stable
Austria	Member	41,700	12,796,558	AAA	Stable
Azerbaijan	Observer	5,400	-16,747,320	BBB-	Positive
Bahrain	Member	27,300	3,102,665	BBB	Stable
Belgium	Member	37,600	-14,930,833	AA+	Negative
Brazil	Member	11,600	-29,796,166	BBB	Stable
Bulgaria	Member	13,500	4,049,237	BBB-	Positive
Canada	Member	40,300	-1,611,380	AAA	Stable
Cape Verde	Member	4,000	874,459	B+	Stable
Chile	Member	16,100	-14,558,218	A+	Stable
China	Member	8,400	-156,705,311	A+	Stable
Colombia	Member	10,100	-1,561,199	BBB-	Stable
Costa Rica	Member	11,500	5,807,509	BB+	Stable
Croatia	Member	18,300	8,108,103	BBB-	Negative
Cyprus	Member	29,100	6,623,337	BBB	Negative
Czech Republic	Member	25,900	-10,749,467	A+	Positive
Denmark	Member	40,200	-15,057,343	AAA	Stable
Ecuador	Member	8,300	1,993,819	B-	Stable
Egypt	Member	6,500	28,486,933	BB	Negative
El Salvador	Member	7,600	5,019,363	BB	Stable
Estonia	Member	20,200	802,234	A+	Stable
France	Member	35,000	118,841,542	AAA	Stable
Georgia	Member	5,400	4,398,153	B+	Positive
Germany	Member	37,900	-213,367,685	AAA	Stable
Hungary	Member	19,600	-9,421,301	BBB-	Negative
Iceland	Member	38,000	-504,939	BB+	Stable
Ireland	Member	39,500	-59,093,323	BBB+	Negative
Israel	Member	31,000	6,722,291	A	Stable
Italy	Member	30,100	33,568,668	A+	Negative
Japan	Member	34,300	31,675,424	AA	Negative
Kazakhstan	Observer	13,000	-33,220,437	BBB	Positive
Kenya	Member	1,700	9,174,198	B+	Stable
Latvia	Member	15,400	2,448,053	BBB-	Positive
Lebanon	Observer	15,600	13,715,550	B	Stable
Lithuania	Member	18,700	3,359,641	BBB	Positive
Malaysia	Member	15,600	-39,420,064	A-	Stable
Mexico	Member	15,100	1,288,112	BBB	Stable
Peru	Member	10,000	-7,888,993	BBB	Stable
Philippines	Member	4,100	15,667,209	BB+	Stable
Poland	Member	20,100	19,552,976	A-	Stable
Portugal	Member	23,200	21,060,508	BBB-	Negative
South Korea	Member	31,700	-37,509,141	A+	Stable
Romania	Member	12,300	13,323,709	BBB-	Stable
Russia	Observer	16,700	-151,400,000	BBB	Positive
Rwanda	Member	1,300	939,222	B	Stable
Serbia	Observer	10,700	8,275,693	BB-	Stable
Seychelles	Observer	24,700	666,026	B	Stable
Singapore	Member	59,900	-27,110,421	AAA	Stable
Slovakia	Member	23,400	-2,110,626	A+	Stable
Slovenia	Member	29,100	2,310,617	AA-	Negative

South Africa	Member	11,000	3,321,801	BBB+	Stable
Sweden	Member	40,600	-10,903,251	AAA	Stable
Switzerland	Member	43,400	-27,197,873	AAA	Stable
Thailand	Member	9,700	2,049,669	BBB	Stable
Turkey	Member	14,600	71,612,947	BB+	Positive
UK	Member	35,900	162,316,831	AAA	Negative
Uruguay	Member	15,400	2,662,628	BB	Positive
USA	Member	48,100	784,438,559	AAA	Stable
Zambia	Member	1,600	-1,805,198	B+	Stable

Member, 5,400, 2,673,359, BB-, and Stable. The second observation (Australia) contains the following measurements: Member, 40,800, -33,304,157, AAA, Stable, and so on. A data set with 60 elements contains 60 observations.

Scales of Measurement

Data collection requires one of the following scales of measurement: nominal, ordinal, interval, or ratio. The scale of measurement determines the amount of information contained in the data and indicates the most appropriate data summarization and statistical analyses.

When the data for a variable consist of labels or names used to identify an attribute of the element, the scale of measurement is considered a **nominal scale**. For example, referring to the data in Table 1.1, the scale of measurement for the WTO Status variable is nominal because the data “member” and “observer” are labels used to identify the status category for the nation. In cases where the scale of measurement is nominal, a numerical code as well as a nonnumerical label may be used. For example, to facilitate data collection and to prepare the data for entry into a computer database, we might use a numerical code for the WTO Status variable by letting 1 denote a member nation in the World Trade Organization and 2 denote an observer nation. The scale of measurement is nominal even though the data appear as numerical values.

The scale of measurement for a variable is considered an **ordinal scale** if the data exhibit the properties of nominal data and in addition, the order or rank of the data is meaningful. For example, referring to the data in Table 1.1, the scale of measurement for the Fitch Rating is ordinal because the rating labels which range from AAA to F can be rank ordered from best credit rating AAA to poorest credit rating F. The rating letters provide the labels similar to nominal data, but in addition, the data can also be ranked or ordered based on the credit rating, which makes the measurement scale ordinal. Ordinal data can also be recorded by a numerical code, for example, your class rank in school.

The scale of measurement for a variable is an **interval scale** if the data have all the properties of ordinal data and the interval between values is expressed in terms of a fixed unit of measure. Interval data are always numerical. College admission SAT scores are an example of interval-scaled data. For example, three students with SAT math scores of 620, 550, and 470 can be ranked or ordered in terms of best performance to poorest performance in math. In addition, the differences between the scores are meaningful. For instance, student 1 scored $620 - 550 = 70$ points more than student 2, while student 2 scored $550 - 470 = 80$ points more than student 3.

The scale of measurement for a variable is a **ratio scale** if the data have all the properties of interval data and the ratio of two values is meaningful. Variables such as distance, height, weight, and time use the ratio scale of measurement. This scale requires that a zero value be included to indicate that nothing exists for the variable at the zero point.

For example, consider the cost of an automobile. A zero value for the cost would indicate that the automobile has no cost and is free. In addition, if we compare the cost of \$30,000 for one automobile to the cost of \$15,000 for a second automobile, the ratio property shows that the first automobile is $\$30,000/\$15,000 = 2$ times, or twice, the cost of the second automobile.

Categorical and Quantitative Data

Data can be classified as either categorical or quantitative. Data that can be grouped by specific categories are referred to as **categorical data**. Categorical data use either the nominal or ordinal scale of measurement. Data that use numeric values to indicate how much or how many are referred to as **quantitative data**. Quantitative data are obtained using either the interval or ratio scale of measurement.

A **categorical variable** is a variable with categorical data, and a **quantitative variable** is a variable with quantitative data. The statistical analysis appropriate for a particular variable depends upon whether the variable is categorical or quantitative. If the variable is categorical, the statistical analysis is limited. We can summarize categorical data by counting the number of observations in each category or by computing the proportion of the observations in each category. However, even when the categorical data are identified by a numerical code, arithmetic operations such as addition, subtraction, multiplication, and division do not provide meaningful results. Section 2.1 discusses ways of summarizing categorical data.

Arithmetic operations provide meaningful results for quantitative variables. For example, quantitative data may be added and then divided by the number of observations to compute the average value. This average is usually meaningful and easily interpreted. In general, more alternatives for statistical analysis are possible when data are quantitative. Section 2.2 and Chapter 3 provide ways of summarizing quantitative data.

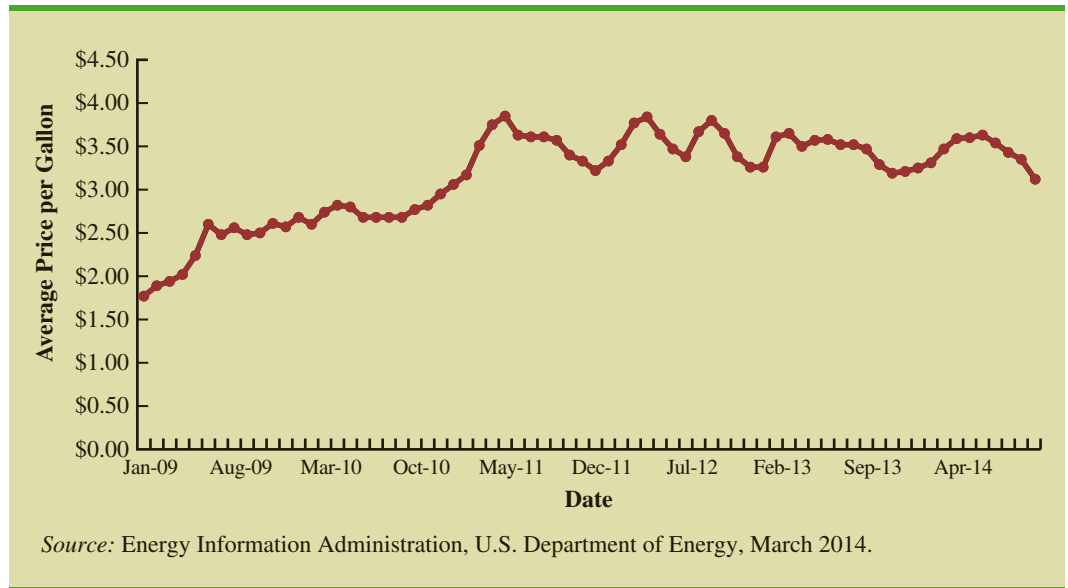
Cross-Sectional and Time Series Data

For purposes of statistical analysis, distinguishing between cross-sectional data and time series data is important. **Cross-sectional data** are data collected at the same or approximately the same point in time. The data in Table 1.1 are cross-sectional because they describe the five variables for the 60 World Trade Organization nations at the same point in time. **Time series data** are data collected over several time periods. For example, the time series in Figure 1.1 shows the U.S. average price per gallon of conventional regular gasoline between 2009 and 2014. Between January 2009 and May 2011, the average price per gallon continued to climb steadily. Since then prices have shown more fluctuation, reaching an average price per gallon of \$3.12 in October 2014.

Graphs of time series data are frequently found in business and economic publications. Such graphs help analysts understand what happened in the past, identify any trends over time, and project future values for the time series. The graphs of time series data can take on a variety of forms, as shown in Figure 1.2. With a little study, these graphs are usually easy to understand and interpret. For example, Panel (A) in Figure 1.2 is a graph that shows the Dow Jones Industrial Average Index from 2004 to 2014. In November 2004, the popular stock market index was near 10,000. The index rose to slightly over 14,000 in October 2007. However, notice the sharp decline in the time series after the high in 2007. By February 2009, poor economic conditions had caused the Dow Jones Industrial Average Index to return to the 7000 level. This was a scary and discouraging period for investors. However, by late 2009, the index was showing a recovery by reaching 10,000. The index has climbed steadily since then and was above 17,500 in November 2014.

The statistical method appropriate for summarizing data depends upon whether the data are categorical or quantitative.

FIGURE 1.1 U.S. AVERAGE PRICE PER GALLON FOR CONVENTIONAL REGULAR GASOLINE



The graph in Panel (B) shows the net income of McDonald's Inc. from 2005 to 2013. The declining economic conditions in 2008 and 2009 were actually beneficial to McDonald's as the company's net income rose to all-time highs. The growth in McDonald's net income showed that the company was thriving during the economic downturn as people were cutting back on the more expensive sit-down restaurants and seeking less-expensive alternatives offered by McDonald's. McDonald's net income continued to new all-time highs in 2010 and 2011, decreased slightly in 2012, and increased in 2013.

Panel (C) shows the time series for the occupancy rate of hotels in South Florida over a one-year period. The highest occupancy rates, 95% and 98%, occur during the months of February and March when the climate of South Florida is attractive to tourists. In fact, January to April of each year is typically the high-occupancy season for South Florida hotels. On the other hand, note the low occupancy rates during the months of August to October, with the lowest occupancy rate of 50% occurring in September. High temperatures and the hurricane season are the primary reasons for the drop in hotel occupancy during this period.

NOTES AND COMMENTS

1. An observation is the set of measurements obtained for each element in a data set. Hence, the number of observations is always the same as the number of elements. The number of measurements obtained for each element equals the number of variables. Hence, the total number of data items can be determined by multiplying the number of observations by the number of variables.
2. Quantitative data may be discrete or continuous. Quantitative data that measure how many (e.g., number of calls received in 5 minutes) are discrete. Quantitative data that measure how much (e.g., weight or time) are continuous because no separation occurs between the possible data values.