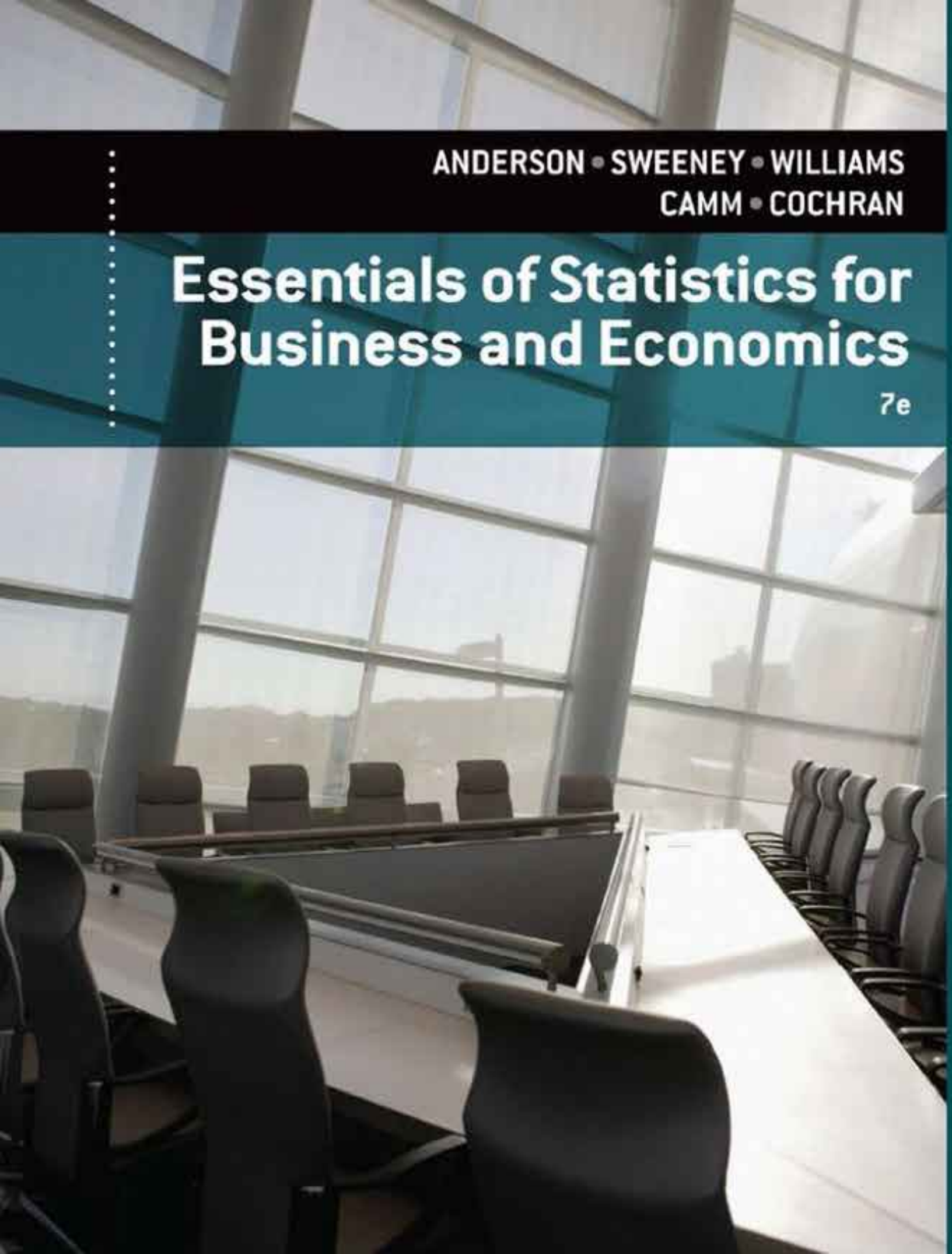


ANDERSON • SWEENEY • WILLIAMS
CAMM • COCHRAN

Essentials of Statistics for Business and Economics

7e

ESSENTIALS OF STATISTICS FOR BUSINESS AND ECONOMICS 7e

ESSENTIALS OF STATISTICS FOR BUSINESS AND ECONOMICS 7e

David R. Anderson
University of Cincinnati

Dennis J. Sweeney
University of Cincinnati

Thomas A. Williams
Rochester Institute of Technology

Jeffrey D. Camm
University of Cincinnati

James J. Cochran
Louisiana Tech University

CENGAGE Learning®

Australia • Brazil • Japan • Korea • Mexico • Singapore • Spain • United Kingdom • United States

This is an electronic version of the print textbook. Due to electronic rights restrictions, some third party content may be suppressed. Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. The publisher reserves the right to remove content from this title at any time if subsequent rights restrictions require it. For valuable information on pricing, previous editions, changes to current editions, and alternate formats, please visit www.cengage.com/highered to search by ISBN#, author, title, or keyword for materials in your areas of interest.

CENGAGE Learning

Essentials of Statistics for Business and Economics, Seventh Edition
David R. Anderson, Dennis J. Sweeney, Thomas A. Williams, Jeffrey D. Camm, James J. Cochran

Senior Vice President, Global Product Manager, Higher Education: Jack W. Calhoun

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

Product Director: Joe Sabatino

Product Manager: Aaron Arnsparger

Content Developer: Maggie Kubale

Product Assistant: Anne Merrill

Sr. Content Project Manager: Tamborah Moore

Media Developer: Chris Valentine

Manufacturing Planner: Ron Montgomery

Marketing Director: Natalie Livingston

Associate Marketing Manager: Roy Rosa

Marketing Manager: Heather Mooney

Marketing Communications Manager: Sarah Greber

Production Service: MPS

Sr. Art Director: Stacy Jenkins Shirley

Internal Designer: Michael Stratton/ cmiller design

Cover Designer: Craig Ramsdell

Cover Image: © Mark Edward Atkinson

Rights Acquisitions Specialist Text and Images: John Hill

Text Permissions Researcher: Pinky Subi/PMG

Image Permissions: Dharanivel Baskar/PMG

©2015, 2011 Cengage Learning

WCN: 02-200-203

ALL RIGHTS RESERVED. No part of this work covered by the copyright herein may be reproduced, transmitted, stored, or used in any form or by any means graphic, electronic, or mechanical, including but not limited to photocopying, recording, scanning, digitizing, taping, web distribution, information networks, or information storage and retrieval systems, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without the prior written permission of the publisher.

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**

Microsoft Excel® is a registered trademark of Microsoft Corporation. © 2013 Microsoft.

Library of Congress Control Number: 2013932884

Pkg ISBN-13: 978-1-133-62965-8

Pkg ISBN-10: 1-133-62965-2

Student Edition ISBN 13: 978-1-133-58779-8

Student Edition ISBN 10: 1-133-58779-8

Cengage Learning
200 First Stamford Place, 4th Floor
Stamford, CT 06902
USA

Cengage Learning is a leading provider of customized learning solutions with office locations around the globe, including Singapore, the United Kingdom, Australia, Mexico, Brazil, and Japan. Locate your local office at: **www.cengage.com/global**

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**

Printed in Canada
1 2 3 4 5 6 7 17 16 15 14 13

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

Brief Contents

Preface xxi

About the Authors xxv

Chapter 1 Data and Statistics 1

Chapter 2 Descriptive Statistics: Tabular and Graphical Displays 33

Chapter 3 Descriptive Statistics: Numerical Measures 105

Chapter 4 Introduction to Probability 176

Chapter 5 Discrete Probability Distributions 222

Chapter 6 Continuous Probability Distributions 263

Chapter 7 Sampling and Sampling Distributions 296

Chapter 8 Interval Estimation 335

Chapter 9 Hypothesis Tests 375

Chapter 10 Comparisons Involving Means, Experimental Design, and Analysis of Variance 424

Chapter 11 Comparisons Involving Proportions and a Test of Independence 481

Chapter 12 Simple Linear Regression 520

Chapter 13 Multiple Regression 588

Appendix A References and Bibliography 639

Appendix B Tables 640

Appendix C Summation Notation online only

Appendix D Self-Test Solutions and Answers to Even-Numbered Exercises 667

Appendix E Microsoft Excel 2013 and Tools for Statistical Analysis 706

Appendix F Computing *p*-Values Using Minitab and Excel 718

Index 722

Contents

Preface xxi
About the Authors xxv

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
Statistical Studies 12
Data Acquisition Errors 14
1.4 Descriptive Statistics 14
1.5 Statistical Inference 16
1.6 Computers and Statistical Analysis 18
1.7 Data Mining 18
1.8 Ethical Guidelines for Statistical Practice 19
Summary 21
Glossary 21
Supplementary Exercises 22
Appendix An Introduction to StatTools 29

Chapter 2 Descriptive Statistics: Tabular and Graphical Displays 33

Statistics in Practice: Colgate-Palmolive Company 34
2.1 Summarizing Data for a Categorical Variable 35
Frequency Distribution 35

Copyright 2013 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Relative Frequency and Percent Frequency Distributions 36
Bar Charts and Pie Charts 36

2.2 Summarizing Data for a Quantitative Variable 42
Frequency Distribution 42
Relative Frequency and Percent Frequency Distributions 43
Dot Plot 44
Histogram 44
Cumulative Distributions 46
Stem-and-Leaf Display 47

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 70
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
Appendix 2.1 Using Minitab for Tabular and Graphical Presentations 86
Appendix 2.2 Using Excel for Tabular and Graphical Presentations 89
Appendix 2.3 Using StatTools for Tabular and Graphical Presentations 103

Chapter 3 Descriptive Statistics: Numerical Measures 105

Statistics in Practice: Small Fry Design 106

3.1 Measures of Location 107
Mean 107
Weighted Mean 109
Median 110
Geometric Mean 112
Mode 113
Percentiles 114
Quartiles 115

Copyright 2013 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

3.2 Measures of Variability 122
Range 122
Interquartile Range 123
Variance 123
Standard Deviation 124
Coefficient of Variation 125
3.3 Measures of Distribution Shape, Relative Location, and Detecting Outliers 129
Distribution Shape 129
z-Scores 129
Chebyshev's Theorem 131
Empirical Rule 132
Detecting Outliers 133
3.4 Five-Number Summaries and Box Plots 136
Five-Number Summary 137
Box Plot 137
3.5 Measures of Association Between Two Variables 141
Covariance 142
Interpretation of the Covariance 144
Correlation Coefficient 146
Interpretation of the Correlation Coefficient 147
3.6 Data Dashboards: Adding Numerical Measures to Improve Effectiveness 151
Summary 155
Glossary 155
Key Formulas 156
Supplementary Exercises 158
Case Problem 1 Pelican Stores 163
Case Problem 2 Motion Picture Industry 164
Case Problem 3 Business Schools of Asia-Pacific 165
Case Problem 4 Heavenly Chocolates Website Transactions 167
Case Problem 5 African Elephant Populations 168
Appendix 3.1 Descriptive Statistics Using Minitab 169
Appendix 3.2 Descriptive Statistics Using Excel 171
Appendix 3.3 Descriptive Statistics Using StatTools 174

Chapter 4 Introduction to Probability 176

Statistics in Practice: National Aeronautics and Space Administration 177
4.1 Experiments, Counting Rules, and Assigning Probabilities 178
Counting Rules, Combinations, and Permutations 179
Assigning Probabilities 183
Probabilities for the KP&L Project 185

Copyright 2013 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

4.2 Events and Their Probabilities 188
4.3 Some Basic Relationships of Probability 192
Complement of an Event 192
Addition Law 193
4.4 Conditional Probability 199
Independent Events 202
Multiplication Law 202
4.5 Bayes' Theorem 207
Tabular Approach 210
Summary 213
Glossary 213
Key Formulas 214
Supplementary Exercises 215
Case Problem Hamilton County Judges 219

Chapter 5 Discrete Probability Distributions 222

Statistics in Practice: Citibank 223
5.1 Random Variables 224
Discrete Random Variables 224
Continuous Random Variables 224
5.2 Developing Discrete Probability Distributions 227
5.3 Expected Value and Variance 232
Expected Value 232
Variance 232
5.4 Binomial Probability Distribution 237
A Binomial Experiment 237
Martin Clothing Store Problem 239
Using Tables of Binomial Probabilities 243
Expected Value and Variance for the Binomial Distribution 244
5.5 Poisson Probability Distribution 248
An Example Involving Time Intervals 249
An Example Involving Length or Distance Intervals 249
5.6 Hypergeometric Probability Distribution 252
Summary 255
Glossary 256
Key Formulas 257
Supplementary Exercises 258
Appendix 5.1 Discrete Probability Distributions with Minitab 261
Appendix 5.2 Discrete Probability Distributions with Excel 261

Copyright 2013 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Chapter 6 Continuous Probability Distributions 263

Statistics in Practice: Procter & Gamble 264

6.1 Uniform Probability Distribution 265
Area as a Measure of Probability 266

6.2 Normal Probability Distribution 269
Normal Curve 269
Standard Normal Probability Distribution 271
Computing Probabilities for Any Normal Probability Distribution 276
Grear Tire Company Problem 277

6.3 Normal Approximation of Binomial Probabilities 281

6.4 Exponential Probability Distribution 285
Computing Probabilities for the Exponential Distribution 285
Relationship Between the Poisson and Exponential Distributions 286

Summary 288
Glossary 289
Key Formulas 289
Supplementary Exercises 289
Case Problem Specialty Toys 293
Appendix 6.1 Continuous Probability Distributions with Minitab 294
Appendix 6.2 Continuous Probability Distributions with Excel 295

Chapter 7 Sampling and Sampling Distributions 296

Statistics in Practice: Meadwestvaco Corporation 297

7.1 The Electronics Associates Sampling Problem 298

7.2 Selecting a Sample 299
Sampling from a Finite Population 299
Sampling from an Infinite Population 301

7.3 Point Estimation 304
Practical Advice 306

7.4 Introduction to Sampling Distributions 308

7.5 Sampling Distribution of $\bar{x}$ 310
Expected Value of $\bar{x}$ 310
Standard Deviation of $\bar{x}$ 311
Form of the Sampling Distribution of $\bar{x}$ 312
Sampling Distribution of $\bar{x}$ for the EAI Problem 314
Practical Value of the Sampling Distribution of $\bar{x}$ 315
Relationship Between the Sample Size and the Sampling Distribution of $\bar{x}$ 316

7.6 Sampling Distribution of $\bar{p}$ 320
Expected Value of $\bar{p}$ 321
Standard Deviation of $\bar{p}$ 321

Copyright 2013 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Form of the Sampling Distribution of $\bar{p}$ 322
Practical Value of the Sampling Distribution of $\bar{p}$ 322

7.7 Other Sampling Methods 326
Stratified Random Sampling 326
Cluster Sampling 327
Systematic Sampling 327
Convenience Sampling 327
Judgment Sampling 328

Summary 328

Glossary 329

Key Formulas 330

Supplementary Exercises 330

Appendix 7.1 Random Sampling with Minitab 333

Appendix 7.2 Random Sampling with Excel 333

Appendix 7.3 Random Sampling with StatTools 334

Chapter 8 Interval Estimation 335

Statistics in Practice: Food Lion 336

8.1 Population Mean: σ Known 337
Margin of Error and the Interval Estimate 337
Practical Advice 341

8.2 Population Mean: σ Unknown 343
Margin of Error and the Interval Estimate 344
Practical Advice 347
Using a Small Sample 347
Summary of Interval Estimation Procedures 349

8.3 Determining the Sample Size 352

8.4 Population Proportion 355
Determining the Sample Size 357

Summary 360

Glossary 361

Key Formulas 362

Supplementary Exercises 362

Case Problem 1 *Young Professional* Magazine 365

Case Problem 2 Gulf Real Estate Properties 366

Case Problem 3 Metropolitan Research, Inc. 368

Appendix 8.1 Interval Estimation with Minitab 368

Appendix 8.2 Interval Estimation Using Excel 370

Appendix 8.3 Interval Estimation with StatTools 373

Copyright 2013 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Chapter 9 Hypothesis Tests 375

Statistics in Practice: John Morrell & Company 376

9.1 Developing Null and Alternative Hypotheses 377
The Alternative Hypothesis as a Research Hypothesis 377
The Null Hypothesis as an Assumption to Be Challenged 378
Summary of Forms for Null and Alternative Hypotheses 379

9.2 Type I and Type II Errors 380

9.3 Population Mean: σ Known 383
One-Tailed Test 383
Two-Tailed Test 389
Summary and Practical Advice 391
Relationship Between Interval Estimation and Hypothesis Testing 393

9.4 Population Mean: σ Unknown 398
One-Tailed Test 398
Two-Tailed Test 399
Summary and Practical Advice 401

9.5 Population Proportion 404
Summary 406

Summary 409

Glossary 410

Key Formulas 410

Supplementary Exercises 410

Case Problem 1 Quality Associates, Inc. 413

Case Problem 2 Ethical Behavior of Business Students at Bayview University 415

Appendix 9.1 Hypothesis Testing with Minitab 416

Appendix 9.2 Hypothesis Testing with Excel 418

Appendix 9.3 Hypothesis Testing with StatTools 422

Chapter 10 Comparisons Involving Means, Experimental Design, and Analysis of Variance 424

Statistics in Practice: U.S. Food and Drug Administration 425

10.1 Inferences About the Difference Between Two Population Means: σ_1 and σ_2 Known 426
Interval Estimation of $\mu_1 - \mu_2$ 426
Hypothesis Tests About $\mu_1 - \mu_2$ 429
Practical Advice 430

10.2 Inferences About the Difference Between Two Population Means: σ_1 and σ_2 Unknown 433
Interval Estimation of $\mu_1 - \mu_2$ 433

Copyright 2013 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Hypothesis Tests About $\mu_1 - \mu_2$ 435
Practical Advice 437

10.3 Inferences About the Difference Between Two Population Means: Matched Samples 441

10.4 An Introduction to Experimental Design and Analysis of Variance 447
Data Collection 448
Assumptions for Analysis of Variance 449
Analysis of Variance: A Conceptual Overview 449

10.5 Analysis of Variance and the Completely Randomized Design 452
Between-Treatments Estimate of Population Variance 453
Within-Treatments Estimate of Population Variance 454
Comparing the Variance Estimates: The *F* Test 455
ANOVA Table 456
Computer Results for Analysis of Variance 457
Testing for the Equality of *k* Population Means: An Observational Study 459

Summary 463
Glossary 464
Key Formulas 464
Supplementary Exercises 466
Case Problem 1 Par, Inc. 471
Case Problem 2 Wentworth Medical Center 472
Case Problem 3 Compensation for Sales Professionals 473
Appendix 10.1 Inferences About Two Populations Using Minitab 474
Appendix 10.2 Analysis of Variance with Minitab 475
Appendix 10.3 Inferences About Two Populations Using Excel 475
Appendix 10.4 Analysis of Variance with Excel 477
Appendix 10.5 Inferences About Two Populations Using StatTools 478
Appendix 10.6 Analysis of a Completely Randomized Design Using StatTools 480

Chapter 11 Comparisons Involving Proportions and a Test of Independence 481

Statistics in Practice: United Way 482

11.1 Inferences About the Difference Between Two Population Proportions 483
Interval Estimation of $p_1 - p_2$ 483
Hypothesis Tests About $p_1 - p_2$ 485

11.2 Testing the Equality of Population Proportions for Three or More Populations 489
A Multiple Comparison Procedure 495

11.3 Test of Independence 500

Copyright 2013 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Summary 508
Glossary 508
Key Formulas 508
Supplementary Exercises 509
Case Problem 1 A Bipartisan Agenda for Change 514
Appendix 11.1 Inferences About Two Population Proportions Using Minitab 515
Appendix 11.2 Chi-Square Tests Using Minitab 516
Appendix 11.3 Chi-Square Tests Using Excel 516
Appendix 11.4 Inferences About Two Population Proportions Using StatTools 518
Appendix 11.5 Chi-Square Tests Using StatTools 519

Chapter 12 Simple Linear Regression 520

Statistics in Practice: Alliance Data Systems 521
12.1 Simple Linear Regression Model 522
Regression Model and Regression Equation 522
Estimated Regression Equation 523
12.2 Least Squares Method 525
12.3 Coefficient of Determination 536
Correlation Coefficient 539
12.4 Model Assumptions 543
12.5 Testing for Significance 544
Estimate of σ^2 544
t Test 546
Confidence Interval for β_1 548
F Test 548
Some Cautions About the Interpretation of Significance Tests 550
12.6 Using the Estimated Regression Equation for Estimation and Prediction 554
Interval Estimation 555
Confidence Interval for the Mean Value of y 555
Prediction Interval for an Individual Value of y 556
12.7 Computer Solution 561
12.8 Residual Analysis: Validating Model Assumptions 565
Residual Plot Against x 566
Residual Plot Against $\hat{y}$ 569
Summary 571
Glossary 572
Key Formulas 572
Supplementary Exercises 574
Case Problem 1 Measuring Stock Market Risk 580

Copyright 2013 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Case Problem 2 U.S. Department of Transportation 581
Case Problem 3 Selecting a Point-and-Shoot Digital Camera 581
Case Problem 4 Finding the Best Car Value 583
Appendix 12.1 Regression Analysis with Minitab 584
Appendix 12.2 Regression Analysis with Excel 584
Appendix 12.3 Regression Analysis Using StatTools 587

Chapter 13 Multiple Regression 588

Statistics in Practice: dunnhumby 589
13.1 Multiple Regression Model 590
Regression Model and Regression Equation 590
Estimated Multiple Regression Equation 590
13.2 Least Squares Method 591
An Example: Butler Trucking Company 592
Note on Interpretation of Coefficients 594
13.3 Multiple Coefficient of Determination 600
13.4 Model Assumptions 604
13.5 Testing for Significance 605
F Test 605
t Test 608
Multicollinearity 609
13.6 Using the Estimated Regression Equation for Estimation and Prediction 612
13.7 Categorical Independent Variables 615
An Example: Johnson Filtration, Inc. 615
Interpreting the Parameters 617
More Complex Categorical Variables 619
Summary 623
Glossary 623
Key Formulas 624
Supplementary Exercises 625
Case Problem 1 Consumer Research, Inc. 631
Case Problem 2 Predicting Winnings for NASCAR Drivers 632
Case Problem 3 Finding the Best Car Value 634
Appendix 13.1 Multiple Regression with Minitab 635
Appendix 13.2 Multiple Regression with Excel 635
Appendix 13.3 Multiple Regression Analysis Using StatTools 636

Appendix A: References and Bibliography 639

Appendix B: Tables 640

Copyright 2013 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Appendix C: Summation Notation online only

Appendix D: Self-Test Solutions and Answers to Even-Numbered Exercises 667

Appendix E: Microsoft Excel 2013 and Tools for Statistical Analysis 706

Appendix F: Computing *p*-Values Using Minitab and Excel 718

Index 722

Copyright 2013 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Copyright 2013 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Preface

This text is the 7th edition of *ESSENTIALS OF STATISTICS FOR BUSINESS AND ECONOMICS.* With this edition we welcome two eminent scholars to our author team: Jeffrey D. Camm of the University of Cincinnati and James J. Cochran of Louisiana Tech University. Both Jeff and Jim are accomplished teachers, researchers, and practitioners in the fields of statistics and business analytics. Jim is a fellow of the American Statistical Association. You can read more about their accomplishments in the About the Authors section that follows this preface. We believe that the addition of Jeff and Jim as our coauthors will both maintain and improve the effectiveness of *Essentials of Statistics for Business and Economics.*

The purpose of *Essentials of Statistics for Business and Economics* 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 knowledge 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 16 and Microsoft® Office Excel 2013 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 appendixes so that instructors have the flexibility of using as much computer emphasis as desired for the course. StatTools, a commercial Excel add-in developed by Palisade Corporation, extends the range of statistical options for Excel users. We show how to download and install StatTools in an appendix to Chapter 1, and most chapters include a chapter appendix that shows the steps required to accomplish a statistical procedure using StatTools. We have made the use of StatTools optional so that instructors who want to teach using only the standard tools available in Excel can do so.

Changes in the Seventh Edition

We appreciate the acceptance and positive response to the previous editions of *Essentials 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 significant changes in the new edition are summarized here.

Copyright 2013 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Content Revisions

- **Descriptive Statistics—Chapters 2 and 3.** We have substantially revised these chapters to incorporate new material on data visualization, best practices, and much more. Chapter 2 has been reorganized to include new material on side-by-side and stacked bar charts and a new section has been added on data visualization and best practices in creating effective displays. Chapter 3 now includes coverage of the geometric mean in the section on measures of location. The geometric mean has many applications in the computation of growth rates for financial assets, annual percentage rates, and so on. Chapter 3 also includes a new section on data dashboards and how summary statistics can be incorporated to enhance their effectiveness.
- **Comparisons Involving Proportions and a Test of Independence—Chapter 11.** This chapter has undergone a major revision. We have replaced the section on goodness of fit tests with a new section on testing the equality of three or more population proportions. This section includes a procedure for making multiple comparison tests between all pairs of population proportions. The section on the test of independence has been rewritten to clarify that the test concerns the independence of two categorical variables. Revised appendices with step-by-step instructions for Minitab, Excel, and StatTools are included.
- **New Case Problems.** We have added 7 new case problems to this edition; the total number of cases is 25. Three new descriptive statistics cases have been added to Chapters 2 and 3. Four new case problems involving regression appear in Chapters 12 and 13. These case problems provide students with the opportunity to analyze larger data sets and prepare managerial reports based on the results of their analysis.
- **New Statistics in Practice Applications.** Each chapter begins with a Statistics in Practice vignette that describes an application of the statistical methodology to be covered in the chapter. New to this edition is a Statistics in Practice for Chapter 2 describing the use of data dashboards and data visualization at the Cincinnati Zoo. We have also added a new Statistics in Practice to Chapter 4 describing how a NASA team used probability to assist the rescue of 33 Chilean miners trapped by a cave-in.
- **New Examples and Exercises Based on Real Data.** We continue to make a significant 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 approximately 200 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 to develop explanations and to create exercises that demonstrate the many uses of statistics in business and economics. We believe that the use of real data helps generate more student interest in the material and enables the student to learn about both the statistical methodology and its application. The seventh edition contains over 300 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.

Copyright 2013 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

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. The data sets are available in both Minitab and Excel formats. Webfile 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.

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

David H. Carhart
Bentley University

Joan M. Donohue
University of South Carolina

Patrick Jaska
University of Mary Hardin-Baylor

Andres Jauregui
Columbus State University

C. P. Kartha
University of Michigan—Flint

Joseph A. Scazzero
Eastern Michigan University

Timothy Scheppa
Concordia University Wisconsin

Matthew J. Stollak
St. Norbert College

Daniel R. Strang
SUNY Geneseo

Daniel A. Talley
Dakota State University

David M. Taurisano
Utica College

Rahmat O. Tavallali
Walsh University

Jennifer VanGilder
Ursinus College

Ahmad Vessal
California State University Northridge

Tatsuma Wada
Wayne State University

Bruce Watson
Boston University

Carol A. Keeth Williams
Central Virginia Community College

Mark Wilson
St. Bonaventure University

Zachary Wong
Sonoma State University

Steven T. Yen
University of Tennessee

Jiang Zhang
Robert B. Willumstad School of Business
Adelphi University

Copyright 2013 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

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:

Alan Smith
Robert Morris College

Ali Arshad
College of Santa Fe

Bennie Waller
Francis Marion University

Carlton Scott
University of California—Irvine

Charles Reichert
University of Wisconsin—Superior

Charles Zimmerman
Robert Morris College

Dale DeBoer
University of Colorado—Colorado Springs

Elaine Parks
Laramie County Community College

Gary Nelson
Central Community College—Columbus Campus

Gipsie Ranney
Belmont University

Habtu Braha
Coppin State College

Karen Gutermuth
Virginia Military Institute

Larry Scheuermann
University of Louisiana, Lafayette

Md. Mahbubul Kabir
Lyon College

Nader Ebrahimi
University of New Mexico

Raj Devasagayam
St. Norbert College

Robert Cochran
University of Wyoming

H. Robert Gadd
Southern Adventist University

Stephen Smith
Gordon College

Timothy Bergquist
Northwest Christian College

Wibawa Sutanto
Prairie View A&M University

Yan Yu
University of Cincinnati

Zhiwei Zhu
University of Louisiana at Lafayette

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 Director, Joe Sabatino; our Product Manager Aaron Arnsparger; our Content Developer Maggie Kubale; our Content Project Manager, Tamborah Moore; our Project Manager at MPS, Lynn Lustberg; our Media Developer Chris Valentine; and others at Cengage Learning 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

Copyright 2013 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

About the Authors

David R. Anderson. David R. Anderson is Professor Emeritus of Quantitative Analysis in the Lindner College of Business 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 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 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 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.

Copyright 2013 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Jeffrey D. Camm. Jeffrey D. Camm is Professor of Quantitative Analysis, Head of the Department of Operations, Business Analytics, and Information Systems and College of Business Research Fellow in the Lindner College of Business at the University of Cincinnati. Born in Cincinnati, Ohio, he holds a B.S. from Xavier University and a Ph.D. from Clemson University. He has been at the University of Cincinnati since 1984 and has 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. He has published his research in *Science, Management Science, Operations Research, Interfaces,* and other professional journals. At the University of Cincinnati, he was named the Dornoff Fellow of Teaching Excellence 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 the Bank of Ruston Barnes, Thompson, & Thurman Endowed Research Professor of Quantitative Analysis at Louisiana Tech University. 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 Louisiana Tech University since 2000 and has been a visiting scholar at Stanford University, Universidad de Talca, the University of South Africa, and Pôle Universitaire Léonard De Vinci.

Professor Cochran has published over two 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, European Journal of Operational Research, Journal of Combinatorial Optimization,* 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. A strong advocate for effective operations research and statistics 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; and Buea, Cameroon. He has served as an operations research consultant to numerous companies and not-for-profit organizations. From 2007 through 2012 he served as editor-in-chief of *INFORMS Transactions on Education* and is on the editorial board of *Interfaces,* the *Journal of the Chilean Institute of Operations Research,* the *Journal of Quantitative Analysis in Sports,* and *ORiON.*

Copyright 2013 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

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
Statistical Studies
Data Acquisition Errors

1.4 DESCRIPTIVE STATISTICS

1.5 STATISTICAL INFERENCE

1.6 COMPUTERS AND STATISTICAL ANALYSIS

1.7 DATA MINING

1.8 ETHICAL GUIDELINES FOR STATISTICAL PRACTICE

Copyright 2013 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

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* 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.*

Bloomberg Businessweek uses statistical facts and summaries in many of its articles. © Kyodo/Photoshot.

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.

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

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

- The median price for an existing single-family home is $186,000, up 7.6% from a year earlier (*The Wall Street Journal,* November 8, 2012).
- Women account for 14.1% of the executive officers in Fortune 500 companies (*The Wall Street Journal,* April 30, 2012).

Copyright 2013 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

- The average annual cost for a college education is $17,100 for public, in-state universities and $38,600 for private universities (*Money Magazine,* March 2012).
- A Yahoo Finance survey reported 51% of workers say the key to getting ahead is internal politics, while 27% say the key to getting ahead is hard work (*USA Today,* September 29, 2012).
- The median age of a first marriage is 29 for men and 26 for women (Associated Press, December 25, 2011).
- The percentage of U.S. workers getting less than six hours of sleep per night is 30% (*The Wall Street Journal,* August 4, 2012).
- The average credit card debt is $5,204 per person (PRWeb website, April 5, 2012).

The numerical facts in the preceding statements ($186,000, 7.6%, 14.1%, $17,100, $38,600, 51%, 27%, 29, 26, 30%, and $5,204) 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 three sections of Chapter 1 provide the role of the computer in statistical analysis, an introduction to data mining, and a discussion of ethical guidelines for statistical practice. A chapter-ending appendix includes an introduction to the add-in StatTools which can be used to extend the statistical options for users of Microsoft Excel.

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.

Copyright 2013 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

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* (March 19, 2012) reported that the average dividend yield for the S&P 500 companies was 2.2%. Microsoft showed a dividend yield of 2.42%. In this case, the statistical information on dividend yield indicates a higher 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 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 x-bar value, is plotted on an 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 x-bar values fall between the chart's upper and lower control limits. Properly interpreted, an 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.

Copyright 2013 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

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 dispute.

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 output of 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 total dollar value of the nation's imports and total dollar value of the nation's exports.
- Fitch Rating: The nation's sovereign credit rating as appraised by the Fitch Group[1]; 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 contains the following measurements:

[1]The 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.

Copyright 2013 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

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

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

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

© Cengage Learning

Copyright 2013 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

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 contains the following measurements: Member, 40,800, −33,304,157, AAA, and 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 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 numeric. 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.

Copyright 2013 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

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.

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

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 for 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.

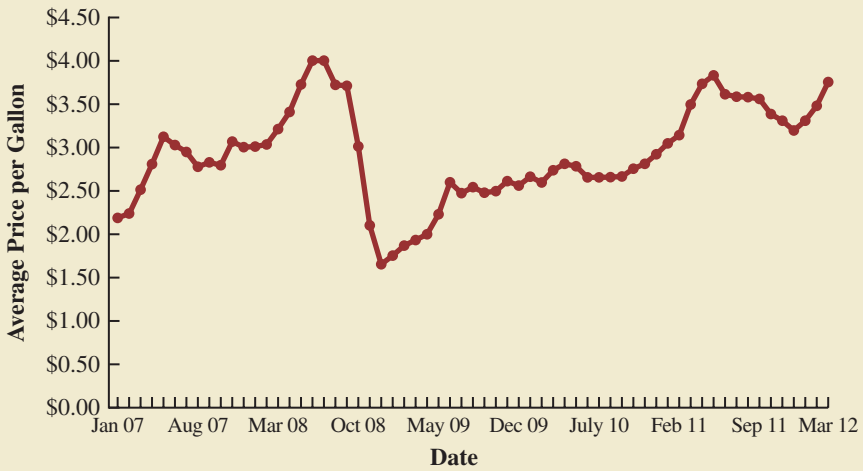
Cross-Sectional and Time Series Data

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 2007 and 2012. Note that gasoline prices peaked in the summer of 2008 and then dropped sharply in the fall of 2008. Since 2008, the average price per gallon has continued to climb steadily, approaching an all-time high again in 2012.

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 2002 to 2012. In April 2002, the popular stock market index was near 10,000. Over the next five years the index rose to its all-time high of slightly over 14,000 in October 2007. However, notice the sharp decline in the time series after the high in 2007. By March 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 and was above 13,000 in early 2012.

Copyright 2013 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

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



Source: Energy Information Administration, U.S. Department of Energy, March 2012.

© Cengage Learning

The graph in Panel (B) shows the net income of McDonald's Inc. from 2005 to 2011. 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.

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.

Copyright 2013 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.