CENGAGE

Anderson Sweeney Williams Camm Cochran Fry Ohlmann

Essentials of Modern Business Statistics

with Microsoft® Excel®

iStockPhoto.com/alongkot-s

Essentials of Modern Business Statistics 8e

with Microsoft® Excel®

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
The University of Alabama

Michael J. Fry
University of Cincinnati

Jeffrey W. Ohlmann
University of Iowa

CENGAGE

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

Copyright 2020 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.

***Essentials of Modern Business Statistics with Microsoft® Excel®*, 8e**
David R. Anderson, Dennis J. Sweeney, Thomas A. Williams, Jeffrey D. Camm, James J. Cochran, Michael J. Fry, Jeffrey W. Ohlmann

Senior Vice President, Higher Education & Skills Product: Erin Joyner

Product Director: Jason Fremder

Senior Product Manager: Aaron Arnsparger

Content Manager: Conor Allen

Product Assistant: Maggie Russo

Marketing Manager: Chris Walz

Senior Learning Designer: Brandon Foltz

Associate Subject Matter Expert: Nancy Marchant

Digital Delivery Lead: Mark Hopkinson

Intellectual Property Analyst: Ashley Maynard

Intellectual Property Project Manager: Kelli Besse

Production Service: MPS Limited

Senior Project Manager, MPS Limited: Manoj Kumar

Art Director: Chris Doughman

Text Designer: Beckmeyer Design

Cover Designer: Beckmeyer Design

Cover Image: iStockPhoto.com/alongkot-s

© 2020, 2018 Cengage Learning, Inc.

Unless otherwise noted, all content is © Cengage.

WCN: 02-300

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 Customer & Sales Support, 1-800-354-9706** or **support.cengage.com.**

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

Library of Congress Control Number: 2019915732

Student Edition:
ISBN: 978-0-357-13162-6

Loose-leaf Edition:
ISBN: 978-0-357-13153-4

Cengage
200 Pier 4 Boulevard
Boston, MA 02210
USA

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

To learn more about Cengage platforms and services, register or access your online learning solution, or purchase materials for your course, visit **www.cengage.com**.

Printed in the United States of America
Print Number: 01 Print Year: 2019

Copyright 2020 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.

Brief Contents

PREFACE xix
ABOUT THE AUTHORS xxv

CHAPTER 1 Data and Statistics 1
CHAPTER 2 Descriptive Statistics: Tabular and Graphical Displays 35
CHAPTER 3 Descriptive Statistics: Numerical Measures 103
CHAPTER 4 Introduction to Probability 171
CHAPTER 5 Discrete Probability Distributions 217
CHAPTER 6 Continuous Probability Distributions 273
CHAPTER 7 Sampling and Sampling Distributions 305
CHAPTER 8 Interval Estimation 355
CHAPTER 9 Hypothesis Tests 397
CHAPTER 10 Inference About Means and Proportions with Two Populations 445
CHAPTER 11 Inferences About Population Variances 489
CHAPTER 12 Tests of Goodness of Fit, Independence, and Multiple Proportions 517
CHAPTER 13 Experimental Design and Analysis of Variance 551
CHAPTER 14 Simple Linear Regression 605
CHAPTER 15 Multiple Regression 685
APPENDIX A References and Bibliography 734
APPENDIX B Tables 736
APPENDIX C Summation Notation 747
APPENDIX D Answers to Even-Numbered Exercises (MindTap Reader)
APPENDIX E Microsoft Excel and Tools for Statistical Analysis 749
APPENDIX F Microsoft Excel Online and Tools for Statistical Analysis 757

INDEX 765

Copyright 2020 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 2020 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.

Contents

PREFACE xix
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 3
Marketing 4
Production 4
Economics 4
Information Systems 4
1.2 Data 5
Elements, Variables, and Observations 5
Scales of Measurement 5
Categorical and Quantitative Data 7
Cross-Sectional and Time Series Data 8
1.3 Data Sources 10
Existing Sources 10
Observational Study 11
Experiment 12
Time and Cost Issues 13
Data Acquisition Errors 13
1.4 Descriptive Statistics 13
1.5 Statistical Inference 15
1.6 Statistical Analysis Using Microsoft Excel 16
Data Sets and Excel Worksheets 17
Using Excel for Statistical Analysis 18
1.7 Analytics 20
1.8 Big Data and Data Mining 21
1.9 Ethical Guidelines for Statistical Practice 22
Summary 24
Glossary 24
Supplementary Exercises 25
Appendix 1.1 Getting Started with R and RStudio (MindTap Reader)
Appendix 1.2 Basic Data Manipulation in R (MindTap Reader)

CHAPTER 2 **Descriptive Statistics: Tabular and Graphical Displays 35**
Statistics in Practice: Colgate-Palmolive Company 36
2.1 Summarizing Data for a Categorical Variable 37
Frequency Distribution 37
Relative Frequency and Percent Frequency Distributions 38

Copyright 2020 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.

Using Excel to Construct a Frequency Distribution, a Relative Frequency Distribution, and a Percent Frequency Distribution 39
Bar Charts and Pie Charts 40
Using Excel to Construct a Bar Chart 42
2.2 Summarizing Data for a Quantitative Variable 47
Frequency Distribution 47
Relative Frequency and Percent Frequency Distributions 49
Using Excel to Construct a Frequency Distribution 50
Dot Plot 51
Histogram 52
Using Excel's Recommended Charts Tool to Construct a Histogram 54
Cumulative Distributions 55
Stem-and-Leaf Display 56
2.3 Summarizing Data for Two Variables Using Tables 65
Crosstabulation 65
Using Excel's PivotTable Tool to Construct a Crosstabulation 68
Simpson's Paradox 69
2.4 Summarizing Data for Two Variables Using Graphical Displays 75
Scatter Diagram and Trendline 76
Using Excel to Construct a Scatter Diagram and a Trendline 77
Side-by-Side and Stacked Bar Charts 79
Using Excel's Recommended Charts Tool to Construct Side-by-Side and Stacked Bar Charts 81
2.5 Data Visualization: Best Practices in Creating Effective Graphical Displays 85
Creating Effective Graphical Displays 85
Choosing the Type of Graphical Display 86
Data Dashboards 86
Data Visualization in Practice: Cincinnati Zoo and Botanical Garden 88
Summary 90
Glossary 91
Key Formulas 92
Supplementary Exercises 93
Case Problem 1: Pelican Stores 98
Case Problem 2: Movie Theater Releases 99
Case Problem 3: Queen City 100
Case Problem 4: Cut-Rate Machining, Inc. 100
Appendix 2.1 Creating Tabular and Graphical Presentations with R (MindTap Reader)

CHAPTER 3 **Descriptive Statistics: Numerical Measures 103**
Statistics in Practice: Small Fry Design 104
3.1 Measures of Location 105
Mean 105

Copyright 2020 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.

Median 107
Mode 108
Using Excel to Compute the Mean, Median, and Mode 109
Weighted Mean 109
Geometric Mean 111
Using Excel to Compute the Geometric Mean 112
Percentiles 113
Quartiles 114
Using Excel to Compute Percentiles and Quartiles 115
3.2 Measures of Variability 121
Range 122
Interquartile Range 122
Variance 122
Standard Deviation 124
Using Excel to Compute the Sample Variance and Sample Standard Deviation 125
Coefficient of Variation 126
Using Excel's Descriptive Statistics Tool 126
3.3 Measures of Distribution Shape, Relative Location, and Detecting Outliers 130
Distribution Shape 130
z-Scores 131
Chebyshev's Theorem 132
Empirical Rule 133
Detecting Outliers 134
3.4 Five-Number Summaries and Boxplots 138
Five-Number Summary 138
Boxplot 138
Using Excel to Construct a Boxplot 139
Comparative Analysis Using Boxplots 139
Using Excel to Construct a Comparative Analysis Using Boxplots 140
3.5 Measures of Association Between Two Variables 144
Covariance 144
Interpretation of the Covariance 146
Correlation Coefficient 148
Interpretation of the Correlation Coefficient 149
Using Excel to Compute the Sample Covariance and Sample Correlation Coefficient 151
3.6 Data Dashboards: Adding Numerical Measures to Improve Effectiveness 153
Summary 156
Glossary 157
Key Formulas 158
Supplementary Exercises 159
Case Problem 1: Pelican Stores 165

Copyright 2020 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: Movie Theater Releases 166
Case Problem 3: Business Schools of Asia-Pacific 167
Case Problem 4: Heavenly Chocolates Website Transactions 167
Case Problem 5: African Elephant Populations 169
Appendix 3.1 Descriptive Statistics with R (MindTap Reader)

CHAPTER 4 **Introduction to Probability 171**
Statistics in Practice: National Aeronautics and Space Administration 172
4.1 Experiments, Counting Rules, and Assigning Probabilities 173
Counting Rules, Combinations, and Permutations 174
Assigning Probabilities 178
Probabilities for the KP&L Project 179
4.2 Events and Their Probabilities 183
4.3 Some Basic Relationships of Probability 187
Complement of an Event 187
Addition Law 188
4.4 Conditional Probability 193
Independent Events 196
Multiplication Law 196
4.5 Bayes' Theorem 201
Tabular Approach 204
Summary 206
Glossary 207
Key Formulas 208
Supplementary Exercises 208
Case Problem 1: Hamilton County Judges 213
Case Problem 2: Rob's Market 215

CHAPTER 5 **Discrete Probability Distributions 217**
Statistics in Practice: Voter Waiting Times in Elections 218
5.1 Random Variables 218
Discrete Random Variables 219
Continuous Random Variables 220
5.2 Developing Discrete Probability Distributions 221
5.3 Expected Value and Variance 226
Expected Value 226
Variance 227
Using Excel to Compute the Expected Value, Variance, and Standard Deviation 228
5.4 Bivariate Distributions, Covariance, and Financial Portfolios 233
A Bivariate Empirical Discrete Probability Distribution 233
Financial Applications 236
Summary 239
5.5 Binomial Probability Distribution 242
A Binomial Experiment 242

Copyright 2020 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.

Martin Clothing Store Problem 244
Using Excel to Compute Binomial Probabilities 248
Expected Value and Variance for the Binomial Distribution 249
5.6 Poisson Probability Distribution 252
An Example Involving Time Intervals 253
An Example Involving Length or Distance Intervals 254
Using Excel to Compute Poisson Probabilities 254
5.7 Hypergeometric Probability Distribution 257
Using Excel to Compute Hypergeometric Probabilities 259
Summary 261
Glossary 262
Key Formulas 263
Supplementary Exercises 264
Case Problem 1: *Go Bananas*! Breakfast Cereal 268
Case Problem 2: McNeil's Auto Mall 269
Case Problem 3: Grievance Committee at Tuglar Corporation 270
Case Problem 4: Sagittarius Casino 270
Appendix 5.1 Discrete Probability Distributions with R (MindTap Reader)

CHAPTER 6 Continuous Probability Distributions 273
Statistics in Practice: Procter & Gamble 274
6.1 Uniform Probability Distribution 275
Area as a Measure of Probability 276
6.2 Normal Probability Distribution 279
Normal Curve 279
Standard Normal Probability Distribution 281
Computing Probabilities for Any Normal Probability Distribution 285
Grear Tire Company Problem 286
Using Excel to Compute Normal Probabilities 288
6.3 Exponential Probability Distribution 293
Computing Probabilities for the Exponential Distribution 294
Relationship Between the Poisson and Exponential Distributions 295
Using Excel to Compute Exponential Probabilities 295
Summary 298
Glossary 298
Key Formulas 298
Supplementary Exercises 299
Case Problem 1: Specialty Toys 301
Case Problem 2: Gebhardt Electronics 302
Appendix 6.1 Continuous Probability Distributions with R (MindTap Reader)

Copyright 2020 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 7 Sampling and Sampling Distributions 305
Statistics in Practice: The Food and Agriculture Organization 306
7.1 The Electronics Associates Sampling Problem 307
7.2 Selecting a Sample 308
Sampling from a Finite Population 308
Sampling from an Infinite Population 312
7.3 Point Estimation 316
Practical Advice 317
7.4 Introduction to Sampling Distributions 319
7.5 Sampling Distribution of $\bar{x}$ 322
Expected Value of $\bar{x}$ 322
Standard Deviation of $\bar{x}$ 322
Form of the Sampling Distribution of $\bar{x}$ 324
Sampling Distribution of $\bar{x}$ for the EAI Problem 324
Practical Value of the Sampling Distribution of $\bar{x}$ 325
Relationship Between the Sample Size and the Sampling Distribution of $\bar{x}$ 327
7.6 Sampling Distribution of $\bar{p}$ 331
Expected Value of $\bar{p}$ 332
Standard Deviation of $\bar{p}$ 332
Form of the Sampling Distribution of $\bar{p}$ 333
Practical Value of the Sampling Distribution of $\bar{p}$ 333
7.7 Other Sampling Methods 337
Stratified Random Sampling 337
Cluster Sampling 337
Systematic Sampling 338
Convenience Sampling 338
Judgment Sampling 339
7.8 Practical Advice: Big Data and Errors in Sampling 339
Sampling Error 339
Nonsampling Error 340
Big Data 341
Understanding What Big Data Is 342
Implications of Big Data for Sampling Error 343
Summary 348
Glossary 348
Key Formulas 349
Supplementary Exercises 350
Case Problem: Marion Dairies 353
Appendix 7.1 Random Sampling with R (MindTap Reader)

Copyright 2020 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 8 Interval Estimation 355
Statistics in Practice: Food Lion 356
8.1 Population Mean: σ Known 357
Margin of Error and the Interval Estimate 357
Using Excel 361
Practical Advice 362
8.2 Population Mean: σ Unknown 364
Margin of Error and the Interval Estimate 365
Using Excel 368
Practical Advice 369
Using a Small Sample 369
Summary of Interval Estimation Procedures 371
8.3 Determining the Sample Size 374
8.4 Population Proportion 377
Using Excel 378
Determining the Sample Size 380
8.5 Practical Advice: Big Data and Interval Estimation 384
Big Data and the Precision of Confidence Intervals 384
Implications of Big Data for Confidence Intervals 385
Summary 387
Glossary 388
Key Formulas 388
Supplementary Exercises 389
Case Problem 1: *Young Professional* Magazine 392
Case Problem 2: GULF Real Estate Properties 393
Case Problem 3: Metropolitan Research, Inc. 395
Appendix 8.1 Interval Estimation with R (MindTap Reader)

CHAPTER 9 Hypothesis Tests 397
Statistics in Practice: John Morrell & Company 398
9.1 Developing Null and Alternative Hypotheses 399
The Alternative Hypothesis as a Research Hypothesis 399
The Null Hypothesis as an Assumption to Be Challenged 400
Summary of Forms for Null and Alternative Hypotheses 401
9.2 Type I and Type II Errors 402
9.3 Population Mean: σ Known 405
One-Tailed Test 405
Two-Tailed Test 410
Using Excel 413
Summary and Practical Advice 414
Relationship Between Interval Estimation and Hypothesis Testing 415
9.4 Population Mean: σ Unknown 420
One-Tailed Test 421
Two-Tailed Test 422

Copyright 2020 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.

Using Excel 423
Summary and Practical Advice 425
9.5 Population Proportion 428
Using Excel 430
Summary 431
9.6 Practical Advice: Big Data and Hypothesis Testing 434
Big Data, Hypothesis Testing, and *p*-Values 434
Implications of Big Data in Hypothesis Testing 436
Summary 437
Glossary 438
Key Formulas 438
Supplementary Exercises 439
Case Problem 1: Quality Associates, Inc. 442
Case Problem 2: Ethical Behavior of Business Students at Bayview University 443
Appendix 9.1 Hypothesis Testing with R (MindTap Reader)

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
Using Excel to Construct a Confidence Interval 449
Hypothesis Tests About $\mu_1 - \mu_2$ 451
Using Excel to Conduct a Hypothesis Test 452
Practical Advice 454
10.2 Inferences About the Difference Between Two Population Means: σ_1 and σ_2 Unknown 456
Interval Estimation of $\mu_1 - \mu_2$ 457
Using Excel to Construct a Confidence Interval 458
Hypothesis Tests About $\mu_1 - \mu_2$ 460
Using Excel to Conduct a Hypothesis Test 462
Practical Advice 463
10.3 Inferences About the Difference Between Two Population Means: Matched Samples 467
Using Excel to Conduct a Hypothesis Test 469
10.4 Inferences About the Difference Between Two Population Proportions 474
Interval Estimation of $p_1 - p_2$ 474
Using Excel to Construct a Confidence Interval 476
Hypothesis Tests About $p_1 - p_2$ 477
Using Excel to Conduct a Hypothesis Test 479
Summary 483
Glossary 483

Copyright 2020 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.

Key Formulas 483
Supplementary Exercises 485
Case Problem: Par, Inc. 488
Appendix 10.1 Inferences About Two Populations with R (MindTap Reader)

CHAPTER 11 **Inferences About Population Variances 489**
Statistics in Practice: U.S. Government Accountability Office 490
11.1 Inferences About a Population Variance 491
Interval Estimation 491
Using Excel to Construct a Confidence Interval 495
Hypothesis Testing 496
Using Excel to Conduct a Hypothesis Test 498
11.2 Inferences About Two Population Variances 503
Using Excel to Conduct a Hypothesis Test 507
Summary 511
Key Formulas 511
Supplementary Exercises 511
Case Problem 1: Air Force Training Program 513
Case Problem 2: Meticulous Drill & Reamer 514
Appendix 11.1 Population Variances with R (MindTap Reader)

CHAPTER 12 **Tests of Goodness of Fit, Independence, and Multiple Proportions 517**
Statistics in Practice: United Way 518
12.1 Goodness of Fit Test 519
Multinomial Probability Distribution 519
Using Excel to Conduct a Goodness of Fit Test 523
12.2 Test of Independence 525
Using Excel to Conduct a Test of Independence 529
12.3 Testing for Equality of Three or More Population Proportions 534
A Multiple Comparison Procedure 537
Using Excel to Conduct a Test of Multiple Proportions 539
Summary 543
Glossary 544
Key Formulas 544
Supplementary Exercises 544
Case Problem 1: A Bipartisan Agenda for Change 547
Case Problem 2: Fuentes Salty Snacks, Inc. 548
Case Problem 3: Fresno Board Games 549
Appendix 12.1 Chi-Square Tests with R (MindTap Reader)

Copyright 2020 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 13 Experimental Design and Analysis of Variance 551
Statistics in Practice: Burke, Inc. 552
13.1 An Introduction to Experimental Design and Analysis of Variance 553
Data Collection 554
Assumptions for Analysis of Variance 556
Analysis of Variance: A Conceptual Overview 556
13.2 Analysis of Variance and the Completely Randomized Design 558
Between-Treatments Estimate of Population Variance 559
Within-Treatments Estimate of Population Variance 560
Comparing the Variance Estimates: The *F* Test 561
ANOVA Table 562
Using Excel 563
Testing for the Equality of *k* Population Means: An Observational Study 564
13.3 Multiple Comparison Procedures 570
Fisher's LSD 570
Type I Error Rates 572
13.4 Randomized Block Design 575
Air Traffic Controller Stress Test 576
ANOVA Procedure 577
Computations and Conclusions 578
Using Excel 579
13.5 Factorial Experiment 584
ANOVA Procedure 585
Computations and Conclusions 586
Using Excel 589
Summary 593
Glossary 594
Key Formulas 595
Completely Randomized Design 595
Multiple Comparison Procedures 596
Randomized Block Design 596
Factorial Experiment 596
Supplementary Exercises 596
Case Problem 1: Wentworth Medical Center 601
Case Problem 2: Compensation for Sales Professionals 602
Case Problem 3: TourisTopia Travel 603
Appendix 13.1 Analysis of Variance with R (MindTap Reader)

CHAPTER 14 Simple Linear Regression 605
Statistics in Practice: walmart.com 606
14.1 Simple Linear Regression Model 607
Regression Model and Regression Equation 607
Estimated Regression Equation 609

Copyright 2020 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.

14.2 Least Squares Method 610
Using Excel to Construct a Scatter Diagram, Display the Estimated Regression Line, and Display the Estimated Regression Equation 614
14.3 Coefficient of Determination 621
Using Excel to Compute the Coefficient of Determination 625
Correlation Coefficient 626
14.4 Model Assumptions 629
14.5 Testing for Significance 631
Estimate of σ^2 631
t Test 632
Confidence Interval for β_1 633
F Test 634
Some Cautions About the Interpretation of Significance Tests 636
14.6 Using the Estimated Regression Equation for Estimation and Prediction 639
Interval Estimation 640
Confidence Interval for the Mean Value of *y* 640
Prediction Interval for an Individual Value of *y* 641
14.7 Excel's Regression Tool 646
Using Excel's Regression Tool for the Armand's Pizza Parlors Example 646
Interpretation of Estimated Regression Equation Output 647
Interpretation of ANOVA Output 648
Interpretation of Regression Statistics Output 649
14.8 Residual Analysis: Validating Model Assumptions 651
Residual Plot Against *x* 652
Residual Plot Against $\hat{y}$ 653
Standardized Residuals 655
Using Excel to Construct a Residual Plot 657
Normal Probability Plot 660
14.9 Outliers and Influential Observations 663
Detecting Outliers 663
Detecting Influential Observations 665
14.10 Practical Advice: Big Data and Hypothesis Testing in Simple Linear Regression 670
Summary 671
Glossary 671
Key Formulas 672
Supplementary Exercises 674
Case Problem 1: Measuring Stock Market Risk 678
Case Problem 2: U.S. Department of Transportation 679
Case Problem 3: Selecting a Point-and-Shoot Digital Camera 680
Case Problem 4: Finding the Best Car Value 681
Case Problem 5: Buckeye Creek Amusement Park 682

Copyright 2020 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 14.1 Calculus-Based Derivation of Least Squares Formulas 683
Appendix 14.2 A Test for Significance Using Correlation 684
Appendix 14.3 Simple Linear Regression with R (MindTap Reader)

CHAPTER 15 **Multiple Regression 685**
Statistics in Practice: International Paper 686
15.1 Multiple Regression Model 687
Regression Model and Regression Equation 687
Estimated Multiple Regression Equation 687
15.2 Least Squares Method 688
An Example: Butler Trucking Company 689
Using Excel's Regression Tool to Develop the Estimated Multiple Regression Equation 691
Note on Interpretation of Coefficients 693
15.3 Multiple Coefficient of Determination 698
15.4 Model Assumptions 700
15.5 Testing for Significance 702
F Test 702
t Test 704
Multicollinearity 705
15.6 Using the Estimated Regression Equation for Estimation and Prediction 708
15.7 Categorical Independent Variables 710
An Example: Johnson Filtration, Inc. 710
Interpreting the Parameters 712
More Complex Categorical Variables 713
15.8 Residual Analysis 718
Residual Plot Against $\hat{y}$ 718
Standardized Residual Plot Against $\hat{y}$ 719
15.9 Practical Advice: Big Data and Hypothesis Testing in Multiple Regression 722
Summary 723
Glossary 723
Key Formulas 724
Supplementary Exercises 725
Case Problem 1: Consumer Research, Inc. 729
Case Problem 2: Predicting Winnings for NASCAR Drivers 730
Case Problem 3: Finding the Best Car Value 732
Appendix 15.1 Multiple Linear Regression with R (MindTap Reader)

Copyright 2020 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 A References and Bibliography 734

APPENDIX B Tables 736

APPENDIX C Summation Notation 747

APPENDIX D Answers to Even-Numbered Exercises (MindTap Reader)

APPENDIX E Microsoft Excel and Tools for Statistical Analysis 749

APPENDIX F Microsoft Excel Online and Tools for Statistical Analysis 757

INDEX 765

Copyright 2020 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 2020 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 eight edition of *Essentials of Modern Business Statistics with Microsoft® Excel®*. With this edition we welcome two eminent scholars to our author team: Michael J. Fry of the University of Cincinnati and Jeffrey W. Ohlmann of the University of Iowa. Both Mike and Jeff are accomplished teachers, researchers, and practitioners in the fields of statistics and business analytics. You can read more about their accomplishments in the About the Authors section that follows this preface. We believe that the addition of Mike and Jeff as our coauthors will both maintain and improve the effectiveness of *Essentials of Modern Business Statistics with Microsoft Excel*.

The purpose of *Essentials of Modern Business Statistics with Microsoft Excel* 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 applications setting, with the statistical results providing insights to decisions and solutions to applied 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.

Use of Microsoft Excel for Statistical Analysis

Essentials of Modern Business Statistics with Microsoft Excel is first and foremost a statistics textbook that emphasizes statistical concepts and applications. But since most practical problems are too large to be solved using hand calculations, some type of statistical software package is required to solve these problems. There are several excellent statistical packages available today. However, because most students and potential employers value spreadsheet experience, many schools now use a spreadsheet package in their statistics courses. Microsoft Excel is the most widely used spreadsheet package in business as well as in colleges and universities. We have written *Essentials of Modern Business Statistics with Microsoft Excel* especially for statistics courses in which Microsoft Excel is used as the software package.

Excel has been integrated within each of the chapters and plays an integral part in providing an application orientation. Although we assume that readers using this text are familiar with Excel basics such as selecting cells, entering formulas, and copying we do not assume that readers are familiar with Excel or Excel's tools for statistical analysis. As a result, we have included Appendix E, which provides an introduction to Excel and tools for statistical analysis.

Throughout the text the discussion of using Excel to perform a statistical procedure appears in a subsection immediately following the discussion of the statistical procedure. We believe that this style enables us to fully integrate the use of Excel throughout the text, but still maintain the primary emphasis on the statistical methodology being discussed. In each of these subsections, we use a standard format for using Excel for statistical analysis. There are four primary tasks: Enter/Access Data, Enter Functions and Formulas, Apply Tools, and Editing Options. We believe a consistent framework for applying Excel helps users to focus on the statistical methodology without getting bogged down in the details of using Excel.

In presenting worksheet figures we often use a nested approach in which the worksheet shown in the background of the figure displays the formulas and the worksheet shown in the foreground shows the values computed using the formulas. Different colors and shades of colors are used to differentiate worksheet cells containing data, highlight cells containing

Copyright 2020 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.

Excel functions and formulas, and highlight material printed by Excel as a result of using one or more data analysis tools.

Changes in the Eighth Edition

We appreciate the acceptance and positive response to the previous editions of *Essentials of Modern Business Statistics with Microsoft Excel*. Accordingly, in making modifications for this new edition, we have maintained the presentation style and readability of those editions. The significant changes in the new edition are summarized here.

- **Software.** In addition to step-by-step instructions and screen captures that show how to use the latest version of Excel to implement statistical procedures, we also provide instructions for Excel Online and R through the MindTap Reader.
- **New Examples and Exercises Based on Real Data.** In this edition, we have added headers to all Applications exercises to make the application of each exercise more clear. We have also added over 160 new examples and exercises based on real data and referenced sources. By using data from sources also used by *The Wall Street Journal*, *USA Today*, *The Financial Times*, *Forbes*, 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 generates greater student interest in the material and enables the student to more effectively learn about both statistical methodology and its application.
- **Case Problems.** We have added four new case problems to this edition. The 47 case problems in the text provide students with the opportunity to analyze somewhat larger data sets and prepare managerial reports based on the results of their analysis.
- **Appendixes for Use of R.** We now provide appendixes in the MindTap Reader for many chapters that demonstrate the use of the popular open-source software R and RStudio for statistical applications. The use of R is not required to solve any problems or cases in the textbook, but the appendixes provide an introduction to R and RStudio for interested instructors and students.

Features and Pedagogy

Authors Anderson, Sweeney, Williams, Camm, Cochran, Fry, and Ohlmann 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.

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.

Copyright 2020 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.

Data Files Accompany the Text

Over 250 data files are available on the website that accompanies the text. 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.

MindTap

MindTap, featuring all new Excel Online integration powered by Microsoft, is a complete digital solution for the business statistics course. It has enhancements that take students from learning basic statistical concepts to actively engaging in critical thinking applications, while learning valuable software skills for their future careers. The R appendixes for many of the chapters in the text are also accessible through MindTap.

MindTap is a customizable digital course solution that includes an interactive eBook and autograded, algorithmic exercises from the textbook. All of these materials offer students better access to understand the materials within the course. For more information on MindTap, please contact your Cengage representative.

For Students

Online resources are available to help the student work more efficiently. The resources can be accessed at www.cengage.com/decisionsciences/anderson/embs/8e.

For Instructors

Instructor resources are available to adopters on the Instructor Companion Site, which can be found and accessed at www.cengage.com/decisionsciences/anderson/embs/8e, including:

- **Solutions Manual:** The Solutions Manual, prepared by the authors, includes solutions for all problems in the text. It is available online as well as print.
- **Solutions to Case Problems:** These are also prepared by the authors and contain solutions to all case problems presented in the text.
- **PowerPoint Presentation Slides:** The presentation slides contain a teaching outline that incorporates figures to complement instructor lectures.
- **Test Bank:** Cengage Learning Testing Powered by Cognero is a flexible, online system that allows you to:
  - author, edit, and manage test bank content from multiple Cengage Learning solutions,
  - create multiple test versions in an instant, and
  - deliver tests from your LMS, your classroom, or wherever you want.

Acknowledgments

A special thanks goes to 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 senior product manager, Aaron Arnsparger; our content manager, Conor Allen; senior learning designer, Brandon Foltz; digital delivery lead, Mark Hopkinson; and our senior project managers at MPS Limited, Santosh Pandey & Manoj Kumar, for their editorial counsel and support during the preparation of this text.

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:

Copyright 2020 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.

Jamal Abdul-Hafidh
University of Missouri–St. Louis

Chris Adalikwu
Concordia College

Eugene Allevato
Woodbury University

Solomon Antony
Murray State University

Ardavan Asef-Vaziri
California State University, Northridge

S. Scott Bailey
Troy University

Robert J. Banis
University of Missouri–St. Louis

Wayne Bedford
University of West Alabama

Enoch K. Beraho
South Carolina State University

Timothy M. Bergquist
Northwest Christian College

Darl Bien
University of Denver

William H. Bleuel
Pepperdine University

Gary Bliss
Florida State University–Panama City

Leslie M. Bobb
New York Institute of Technology

Michelle Boddy
Baker College

Thomas W. Bolland
Ohio University

Derrick S. Boone, Sr.
Wake Forest University

Lawrence Bos
Cornerstone University

Alan Brokaw
Michigan Tech University

Nancy Brooks
University of Vermont

Yvonne Brown
Pima Community College

Dawn Bulriss
Maricopa Community Colleges

Robert Burgess
Georgia Tech

Von L. Burton
Athens State University

John R. Carpenter
Cornerstone University

Jasmine Chang
Georgia State University

Si Chen
Murray State University

Alan S. Chesen
Wright State University

Michael Cicero
Highline Community College

Robert Collins
Marquette University

Ping Deng
Maryville University

Sarvanan Devaraj
Notre Dame University

Terry Dielman
Texas Christian University

Cassandra DiRienzo
Elon University

Anne Drougas
Dominican University

Jianjun Du
University of Houston, Victoria

John N. Dyer
Georgia Southern University

Hossein Eftekari
University of Wisconsin–River Falls

Mohammed A. El-Saidi
Ferris State University

Robert M. Escudero
Pepperdine University

Allessandra Faggian
Ohio State University

Nicolas Farnum
California State University, Fullerton

Abe Feinberg
California State University, Northridge

Maggie Williams Flint
Northeast State Tech Community College

Alfonso Flores-Lagunes
University of Arizona

James Flynn
Cleveland State University

Alan F. Foltz
Drury University

Ronald L. Friesen
Bluffton College

Richard Gebhart
University of Tulsa

Paul Gentine
Bethany College

Deborah J. Gougeon
University of Scranton

Jeffrey Gropp
DePauw University

V. Daniel Guide
Duquesne University

Aravind Narasipur
Chennai Business School

Rhonda Hensley
North Carolina A&T University

Erick Hofacker
University of Wisconsin–River Falls

Amy C. Hooper
Gettysburg College

Paul Hudec
Milwaukee School of Engineering

Alan Humphrey
University of Rhode Island

Wade Jackson
University of Memphis

Timmy James
Northwest Shoals Community College

Copyright 2020 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.

Eugene Jones
The Ohio State University

Naser Kamleh
Wallace Community College

Mark P. Karscig
Central Missouri State University

Howard Kittleson
Riverland Community College

Kenneth Klassen
California State University, Northridge

Eileen Quinn Knight
St. Xavier University, Chicago

Bharat Kolluri
University of Hartford

Joseph Kosler
Indiana University of Pennsylvania

David A. Kravitz
George Mason University

Laura Kuhl
University of Phoenix, Cleveland Campus

June Lapidus
Roosevelt University

John Lawrence
California State University, Fullerton

Tenpao Lee
Niagara University

Daniel Light
Northwest State College

Robert Lindsey
College of Charleston

B. Lucas
A&M College

Michael Machiorlatti
City College of San Francisco

Malik B. Malik
University of Maryland Eastern Shore

Lee McClain
Western Washington University

Timothy E. McDaniel
Buena Vista University

Kim I. Melton
North Georgia College & State University

Brian Metz
Cabrini College

John M. Miller
Sam Houston State University

Patricia A. Mullins
University of Wisconsin–Madison

Jack Muryn
University of Wisconsin, Washington County

Muhammad Mustafa
South Carolina State University

Anthony Narsing
Macon State College

Kenneth F. O'Brien
Farmingdale State College

Ceyhun Ozgur
Valparaiso University

Michael Parzen
Emory University

Barry Pasternack
California State University, Fullerton

Lynne Pastor
Carnegie Mellon University

Ranjna Patel
Bethune-Cookman College

Tremaine Pimperl
Faulkner State Community College

Jennifer M. Platania
Elon University

Von Roderick Plessner
Northwest State Community College

Glenn Potts
University of Wisconsin–River Falls

Irene Powell
Grinnell College

Leonard Presby
William Paterson University

W. N. Pruitt
South Carolina State University

Narseeyappa Rajanikanth
Mississippi Valley State University

Elizabeth L. Rankin
Centenary College of Louisiana

Tim Raynor
Albertus Magnus College

Carolyn Renier
Pellissippi State Tech Community College

Ronny Richardson
Southern Polytechnic State University

Leonard E. Ross
California State University, Pomona

Probir Roy
University of Missouri, Kansas City

Randall K. Russell
Yavapai College

Alan Safer
California State University, Long Beach

David Satava
University of Houston, Victoria

Richard W. Schrader
Bellarmine University

Larry Seifert
Webster University

John Seydel
Arkansas State University

Jim Shi
Robinson College of Business
Georgia State University

Philip Shaw
Fairfield University

Robert Simoneau
Keene State College

Copyright 2020 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.

Harvey A. Singer
George Mason University

Donald R. Smith
Monmouth University

Toni M. Somers
Wayne State University

Clifford Sowell
Berea College

Keith Starcher
Indiana Wesleyan University

William Stein
Texas A&M University

Jason Stine
Troy University

William Struning
Seton Hall University

Timothy S. Sullivan
Southern Illinois University, Edwardsville

Peter Wibawa Sutanto
Prairie View A&M University

Lee Tangedahk
University of Montana

Sudhir Thakur
California State University, Sacramento

Alexander Thomson
Schoolcraft College

Suzanne Tilleman
Montana State University Northern

Daniel Tschopp
Daemen College (NY)

Sushila Umashankar
University of Arizona

Jack Vaughn
University of Texas, El Paso

Dave Vinson
Pellissippi State Community College

John Vogt
Newman University

Geoffrey L. Wallace
University of Wisconsin, Madison

Michael Wiemann
Metro Community Colleges

Charles Wilf
Duquesne University

John Wiorkowski
University of Texas, Dallas

Joyce A. Zadzilka
Morrisville State College

Guoqiang Peter Zhang
Georgia Southern University

Zhe George Zhang
Western Washington University

Deborah G. Ziegler
Hannibal-LaGrange College

We would like to recognize the following individuals who have helped us in the past and continue to influence our writing.

Glen Archibald
University of Mississippi

Mike Bourke
Houston Baptist University

Peter Bryant
University of Colorado

Terri L. Byczkowski
University of Cincinnati

Ying Chien
University of Scranton

Robert Cochran
University of Wyoming

Murray Côté
University of Florida

David W. Cravens
Texas Christian University

Robert Carver
Stonehill College

Tom Dahlstrom
Eastern College

Ronald Ehresman
Baldwin-Wallace College

Michael Ford
Rochester Institute of Technology

Phil Fry
Boise State University

Paul Guy
California State University, Chico

Alan Humphrey
University of Rhode Island

Ann Hussein
Philadelphia College of Textiles and Science

Ben Isselhardt
Rochester Institute of Technology

Jeffery Jarrett
University of Rhode Island

Copyright 2020 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 in the Carl H. 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 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 in the Carl H. Lindner College of Business 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 Duke Energy 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 has taught 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 Business 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.

Copyright 2020 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.

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 the *INFORMS Journal on Applied Analytics* (formerly *Interfaces*).

James J. Cochran. James J. Cochran is Associate Dean for Research, 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 Pôle Universitaire Léonard de Vinci.

Professor Cochran has published over 40 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, was named a Fellow of the American Statistical Association in 2011, and was named a Fellow of INFORMS in 2017. He also received the Founders Award in 2014 and the Karl E. Peace Award in 2015 from the American Statistical Association as well as the President's Award in 2018 from INFORMS. 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 Uruguay, South Africa, Colombia, India, Argentina, Kenya, Cuba, Croatia, Cameroon, Nepal, Moldova, and Bulgaria. He has served as a statistical 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 *INFORMS Journal of Applied Analytics* (formerly *Interfaces*), *International Transactions in Operational Research*, and *Significance*.

Michael J. Fry. Michael J. Fry is Professor of Operations, Business Analytics, and Information Systems (OBAIS) and Academic Director of the Center for Business Analytics in the Carl H. Lindner College of Business at the University of Cincinnati. Born in Killeen, Texas, he earned a B.S. from Texas A&M University, and M.S.E. and Ph.D. degrees from the University of Michigan. He has been at the University of Cincinnati since 2002, where he was previously Department Head and has been named a Lindner Research Fellow. He has also been a visiting professor at the Samuel Curtis Johnson Graduate School of Management at Cornell University and the Sauder School of Business at the University of British Columbia.

Professor Fry has published more than 25 research papers in journals such as *Operations Research*, *M&SOM*, *Transportation Science*, *Naval Research Logistics*, *IIE Transactions*, *Critical Care Medicine*, and *Interfaces.* He serves on editorial boards for journals such as *Production and Operations Management*, *INFORMS Journal of Applied Analytics* (formerly *Interfaces*), *Omega*, and *Journal of Quantitative Analysis in Sports.* His research interests

Copyright 2020 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.

are in applying analytics to the areas of supply chain management, sports, and public-policy operations. He has worked with many different organizations for his research, including Dell, Inc., Starbucks Coffee Company, Great American Insurance Group, the Cincinnati Fire Department, the State of Ohio Election Commission, the Cincinnati Bengals, and the Cincinnati Zoo & Botanical Garden. He was named a finalist for the Daniel H. Wagner Prize for Excellence in Operations Research Practice, and he has been recognized for both his research and teaching excellence at the University of Cincinnati. In 2019 he led the team that was awarded the INFORMS UPS George D. Smith Prize on behalf of the OBAIS Department at the University of Cincinnati.

Jeffrey W. Ohlmann. Jeffrey W. Ohlmann is Associate Professor of Business Analytics and Huneke Research Fellow in the Tippie College of Business at the University of Iowa. Born in Valentine, Nebraska, he earned a B.S. from the University of Nebraska, and M.S. and Ph.D. degrees from the University of Michigan. He has been at the University of Iowa since 2003.

Professor Ohlmann's research on the modeling and solution of decision-making problems has produced more than two dozen research papers in journals such as *Operations Research*, *Mathematics of Operations Research*, *INFORMS Journal on Computing*, *Transportation Science*, and *European Journal of Operational Research*. He has collaborated with companies such as Transfreight, LeanCor, Cargill, the Hamilton County Board of Elections, and three National Football League franchises. Because of the relevance of his work to industry, he was bestowed the George B. Dantzig Dissertation Award and was recognized as a finalist for the Daniel H. Wagner Prize for Excellence in Operations Research Practice.

Copyright 2020 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 2020 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.

iStockPhoto.com/alongkot-s

Essentials of Modern Business Statistics 8e with Microsoft® Excel®

Copyright 2020 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 2020 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
Observational Study
Experiment
Time and Cost Issues
Data Acquisition Errors

1.4 DESCRIPTIVE STATISTICS

1.5 STATISTICAL INFERENCE

1.6 STATISTICAL ANALYSIS USING MICROSOFT EXCEL
Data Sets and Excel Worksheets
Using Excel for Statistical Analysis

1.7 ANALYTICS

1.8 BIG DATA AND DATA MINING

1.9 ETHICAL GUIDELINES FOR STATISTICAL PRACTICE

SUMMARY 24
GLOSSARY 24
SUPPLEMENTARY EXERCISES 25

APPENDIXES
APPENDIX 1.1: GETTING STARTED WITH R AND RSTUDIO (MINDTAP READER)
APPENDIX 1.2: BASIC DATA MANIPULATION IN R (MINDTAP READER)

Copyright 2020 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

Bloomberg Businessweek is one of the most widely read business magazines in the world. 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* 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 North American subscriber survey indicated 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 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.

Bloomberg Businessweek *uses statistical facts and summaries in many of its articles. AP Images/Weng lei-Imaginechina*

*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:

- Unemployment dropped to an 18-year low of 3.8% in May 2018 from 3.9% in April and after holding at 4.1% for the prior six months (*Wall Street Journal,* June 1, 2018).
- Tesla ended 2017 with around $5.4 billion of liquidity. Analysts forecast it will burn through $2.8 billion of cash this year (*Bloomberg Businessweek,* April 19, 2018).
- The biggest banks in America reported a good set of earnings for the first three months of 2018. Bank of America and Morgan Stanley made quarterly net profits of $6.9 billion and $2.7 billion, respectively (*The Economist,* April 21, 2018).
- According to a study from the Pew Research Center, 15% of U.S. adults say they have used online dating sites or mobile apps (*Wall Street Journal,* May 2, 2018).

Copyright 2020 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.

- According to the U.S. Centers for Disease Control and Prevention, in the United States alone, at least 2 million illnesses and 23,000 deaths can be attributed each year to antibiotic-resistant bacteria (*Wall Street Journal,* February 13, 2018).

The numerical facts in the preceding statements—3.8%, 3.9%, 4.1%, $5.4 billion, $2.8 billion, $6.9 billion, $2.7 billion, 15%, 2 million, 23,000—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 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 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 the role of the computer in statistical analysis, an introduction to business analytics and the role statistics plays in it, an introduction to big data and data mining, 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 average dividend yield for the S&P 500 companies for 2017 was 1.88%. Over the same period the average dividend yield for Microsoft was 1.72% (*Yahoo Finance*). In this case, the statistical information on dividend yield indicates a lower dividend yield for Microsoft

Copyright 2020 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.

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 The Nielsen Company and IRI 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.

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.

Copyright 2020 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.

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 (WTO). The WTO 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 as an observer.
- Per Capita Gross Domestic Product (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.
- 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: Member, 3615, BB–, and Stable. The second observation contains the following measurements: Member, 49755, 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

[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 2020 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.

DATA*file*
Nations

Data sets such as Nations *are available on the companion site for this title.*

TABLE 1.1 Data Set for 60 Nations in the World Trade Organization

Nation	WTO Status	Per Capita GDP ($)	Fitch Rating	Fitch Outlook
Armenia	Member	3,615	BB−	Stable
Australia	Member	49,755	AAA	Stable
Austria	Member	44,758	AAA	Stable
Azerbaijan	Observer	3,879	BBB−	Stable
Bahrain	Member	22,579	BBB	Stable
Belgium	Member	41,271	AA	Stable
Brazil	Member	8,650	BBB	Stable
Bulgaria	Member	7,469	BBB−	Stable
Canada	Member	42,349	AAA	Stable
Cape Verde	Member	2,998	B+	Stable
Chile	Member	13,793	A+	Stable
China	Member	8,123	A+	Stable
Colombia	Member	5,806	BBB−	Stable
Costa Rica	Member	11,825	BB+	Stable
Croatia	Member	12,149	BBB−	Negative
Cyprus	Member	23,541	B	Negative
Czech Republic	Member	18,484	A+	Stable
Denmark	Member	53,579	AAA	Stable
Ecuador	Member	6,019	B−	Positive
Egypt	Member	3,478	B	Negative
El Salvador	Member	4,224	BB	Negative
Estonia	Member	17,737	A+	Stable
France	Member	36,857	AAA	Negative
Georgia	Member	3,866	BB−	Stable
Germany	Member	42,161	AAA	Stable
Hungary	Member	12,820	BB+	Stable
Iceland	Member	60,530	BBB	Stable
Ireland	Member	64,175	BBB+	Stable
Israel	Member	37,181	A	Stable
Italy	Member	30,669	A−	Negative
Japan	Member	38,972	A+	Negative
Kazakhstan	Observer	7,715	BBB+	Stable
Kenya	Member	1,455	B+	Stable
Latvia	Member	14,071	BBB	Positive
Lebanon	Observer	8,257	B	Stable
Lithuania	Member	14,913	BBB	Stable
Malaysia	Member	9,508	A−	Stable
Mexico	Member	8,209	BBB	Stable
Peru	Member	6,049	BBB	Stable
Philippines	Member	2,951	BB+	Stable
Poland	Member	12,414	A−	Positive
Portugal	Member	19,872	BB+	Negative
South Korea	Member	27,539	AA−	Stable
Romania	Member	9,523	BBB−	Stable
Russia	Member	8,748	BBB	Stable
Rwanda	Member	703	B	Stable
Serbia	Observer	5,426	BB−	Negative
Singapore	Member	52,962	AAA	Stable
Slovakia	Member	16,530	A+	Stable

Copyright 2020 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.

Slovenia	Member	21,650	A−	Negative
South Africa	Member	5,275	BBB	Stable
Spain	Member	26,617	A−	Stable
Sweden	Member	51,845	AAA	Stable
Switzerland	Member	79,888	AAA	Stable
Thailand	Member	5,911	BBB	Stable
Turkey	Member	10,863	BBB−	Stable
United Kingdom	Member	40,412	AAA	Negative
United States	Member	57,638	AAA	Stable
Uruguay	Member	15,221	BB+	Positive
Zambia	Member	1,270	B+	Negative

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. 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 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

Copyright 2020 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.

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

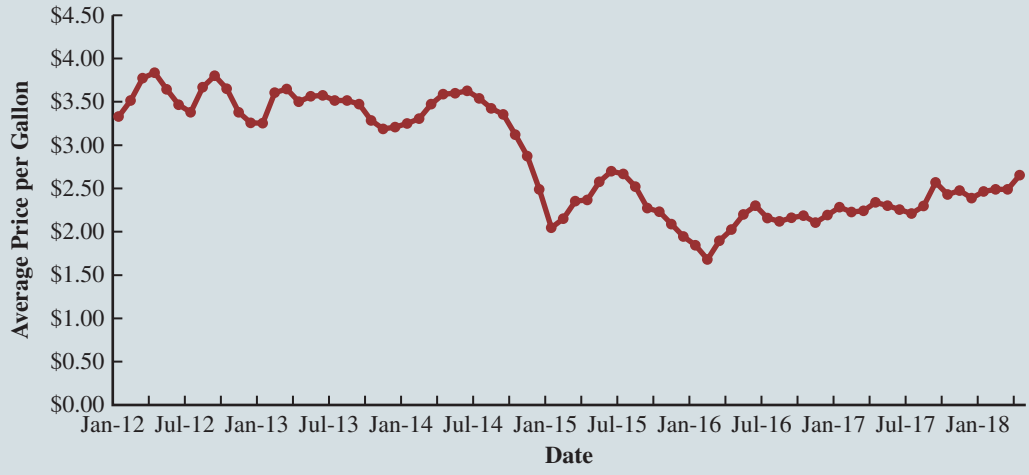
For purposes of statistical analysis, distinguishing between cross-sectional data and time series data is important. **Cross-sectional data** are data collected at the same or approximately the same point in time. The data in Table 1.1 are cross-sectional because they describe the five variables for 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 2012 and 2018. From January 2012 until June 2014, prices fluctuated between $3.19 and $3.84 per gallon before a long stretch of decreasing prices from July 2014 to January 2015. The lowest average price per gallon occurred in January 2016 ($1.68). Since then, the average price appears to be on a gradual increasing trend.

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 2008 to 2018. Poor economic conditions caused a serious drop in the index during 2008 with the low point occurring in February 2009 (7,062). After that, the index has been on a remarkable nine-year increase, reaching its peak (26,149) in January 2018.

The graph in Panel (B) shows the net income of McDonald's Inc. from 2008 to 2017. 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 peaked in 2013. After three years of relatively lower net income, their net income increased to $5.19 billion in 2017.

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

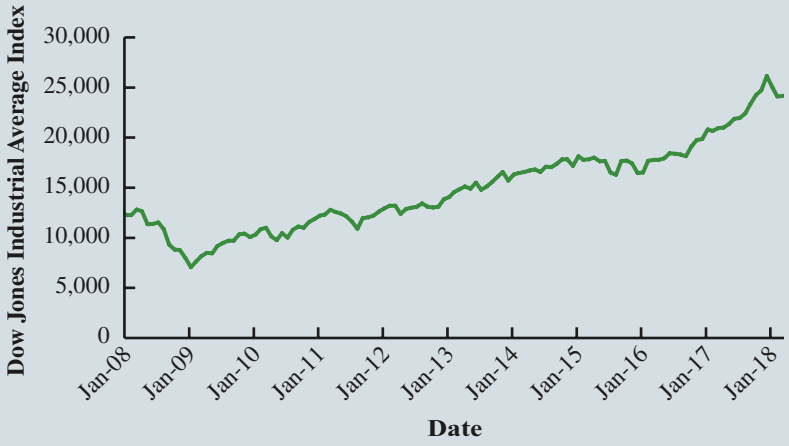
FIGURE 1.1 U.S. Average Price per Gallon for Conventional Regular Gasoline



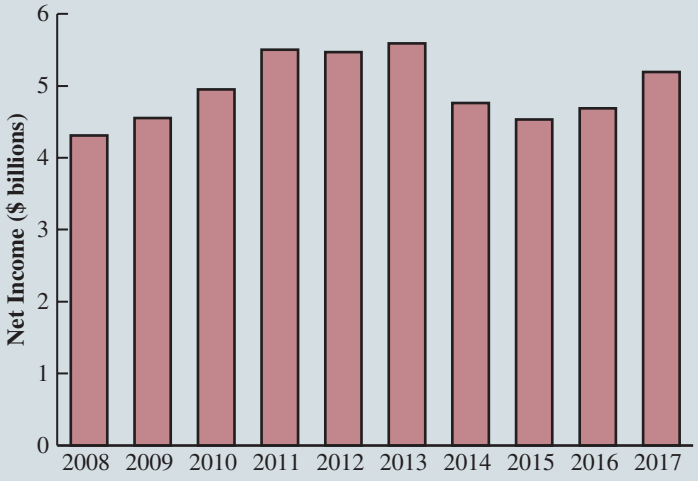
Source: *Energy Information Administration, U.S. Department of Energy.*

Copyright 2020 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.

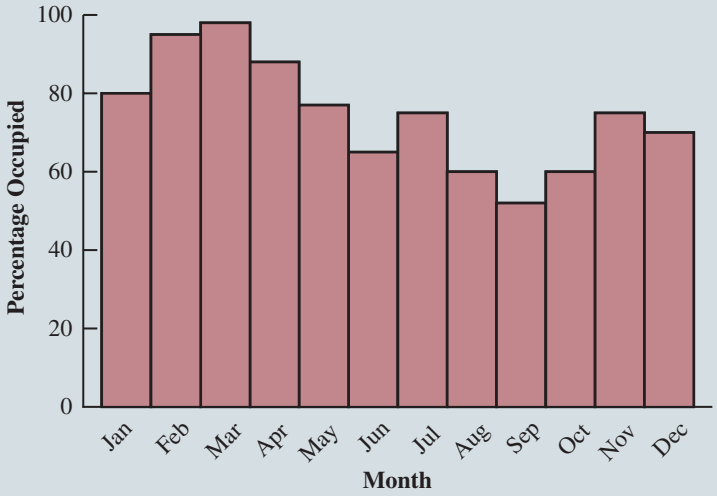
FIGURE 1.2 A Variety of Graphs of Time Series Data



(A) Dow Jones Industrial Average Index



(B) Net Income for McDonald's Inc.



(C) Occupancy Rate of South Florida Hotels

Copyright 2020 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.