

Basic Business Statistics

Concepts and Applications

Mark L. Berenson

Department of Information Management and Business Analytics

School of Business, Montclair State University

David M. Levine

Department of Information Systems and Statistics

Zicklin School of Business, Baruch College, City University of New York

Kathryn A. Szabat

Department of Business Systems and Analytics

School of Business, La Salle University

David F. Stephan

Two Bridges Instructional Technology



Pearson

Contents

Preface 24

First Things First 33

USING STATISTICS: "The Price of Admission" 33

FTF.1 Think Differently About Statistics 34

Statistics: A Way of Thinking 34

Statistics: An Important Part of Your Business Education 35

FTF.2 Business Analytics: The Changing Face of Statistics 36

"Big Data" 36

FTF.3 Starting Point for Learning Statistics 37

Statistic 37

Can Statistics $\{pi., \text{statistic}\}$ Lie? 38

FTF.4 Starting Point for Using Software 38

Using Software Properly 40

REFERENCES 41

KEY TERMS 41

EXCEL GUIDE 42

EG.1 Getting Started with Excel 42

EG.2 Entering Data 42

EG.3 Open or Save a Workbook 42

EG.4 Working with a Workbook 43

EG.5 Print a Worksheet 43

EG.6 Reviewing Worksheets 43

EG.7 If You Use the *Workbook* Instructions 43

JMP GUIDE 44

JG.1 Getting Started with JMP 44

JG.2 Entering Data 45

JG.3 Create New Project or Data Table 45

JG.4 Open or Save Files 45

JG.5 Print Data Tables or Report Windows 45

JG.6 JMP Script Files 45

MINITAB GUIDE 46

MG.1 Getting Started with Minitab 46

MG.2 Entering Data 46

MG.3 Open or Save Files 46

MG.4 Insert or Copy Worksheets 47

MG.5 Print Worksheets 47

1 Defining and Collecting Data 48

USING STATISTICS: Defining Moments 48

1.1 Defining Variables 49

Classifying Variables by Type 49

Measurement Scales 50

1.2 Collecting Data 51

Populations and Samples 52

Data Sources 52

1.3 Types of Sampling Methods 53

Simple Random Sample 54

Systematic Sample 54

Stratified Sample 55

Cluster Sample 55

1.4 Data Cleaning 56

Invalid Variable Values 57

Coding Errors 57

Data Integration Errors 57

Missing Values 58

Algorithmic Cleaning of Extreme Numerical Values 58

1.5 Other Data Preprocessing Tasks 58

Data Formatting 58

Stacking and Unstacking Data 59

Recoding Variables 59

1.6 Types of Survey Errors 60

Coverage Error 60

Nonresponse Error 60

Sampling Error 60

Measurement Error 61

Ethical Issues About Surveys 61

CONSIDER THIS: New Media Surveys/Old Survey Errors 61

USING STATISTICS: Defining Moments, Revisited 63

SUMMARY 63

REFERENCES 63

KEY TERMS 63

CHECKING YOUR UNDERSTANDING 64

CHAPTER REVIEW PROBLEMS 64

CASES FOR CHAPTER 1 65

Managing Ashland MultiComm Services 65

CardioGood Fitness 65

Clear Mountain State Student Survey 66

Learning with the Digital Cases 66

CHAPTER 1 EXCEL GUIDE 67

EG1.1 Defining Variables 67

EG1.2 Collecting Data 67

EG1.3 Types of Sampling Methods 67

EG1.4 Data Cleaning 68

EG1.5 Other Data Preprocessing 68

CHAPTER 1 JMP GUIDE 69

JG1.1 Defining Variables 69

JG1.2 Collecting Data 69

JG1.3 Types of Sampling Methods 69

JG1.4 Data Cleaning 70

JG1.5 Other Preprocessing Tasks 71

CHAPTER 1 MINITAB GUIDE 71

- MG1.1 Defining Variables 71
- MG1.2 Collecting Data 71
- MG1.3 Types of Sampling Methods 71
- MG1.4 Data Cleaning 72
- MG1.5 Other Preprocessing Tasks 72

2 Organizing and Visualizing Variables 73

USING STATISTICS: "The Choice Is Yours" 73

- 2.1 Organizing Categorical Variables 74
 - The Summary Table 74
 - The Contingency Table 75
- 2.2 Organizing Numerical Variables 78
 - The Frequency Distribution 79
 - Classes and Excel Bins 81
 - The Relative Frequency Distribution and the Percentage Distribution 81
 - The Cumulative Distribution 83
- 2.3 Visualizing Categorical Variables 86
 - The Bar Chart 86
 - The Pie Chart and the Doughnut Chart 87
 - The Pareto Chart 88
 - Visualizing Two Categorical Variables 90
- 2.4 Visualizing Numerical Variables 93
 - The Stem-and-Leaf Display 93
 - The Histogram 93
 - The Percentage Polygon 95
 - The Cumulative Percentage Polygon (Ogive) 96
- 2.5 Visualizing Two Numerical Variables 99
 - The Scatter Plot 99
 - The Time-Series Plot 100
- 2.6 Organizing a Mix of Variables 102
 - Drill-down 103
- 2.7 Visualizing a Mix of Variables 104
 - Colored Scatter Plot 104
 - Bubble Charts 105
 - PivotChart (Excel) 105
 - Treemap (Excel, JMP) 105
 - Sparklines (Excel) 106
- 2.8 Filtering and Querying Data 107
 - Excel Slicers 107
- 2.9 Pitfalls in Organizing and Visualizing Variables 109
 - Obscuring Data 109
 - Creating False Impressions 110
 - Chartjunk 111

EXHIBIT: Best Practices for Creating Visual Summaries 112

USING STATISTICS: "The Choice Is Yours," Revisited 113**SUMMARY 113****REFERENCES 114****KEY EQUATIONS 114****KEY TERMS 115****CHECKING YOUR UNDERSTANDING 115****CHAPTER REVIEW PROBLEMS 115****CASES FOR CHAPTER 2 120**

- Managing Ashland MultiComm Services 120
- Digital Case 120
- CardioGood Fitness 121
- The Choice Is Yours Follow-Up 121
- Clear Mountain State Student Survey 121

CHAPTER 2 EXCEL GUIDE 122

- EG2.1 Organizing Categorical Variables 122
- EG2.2 Organizing Numerical Variables 124
- EG2 Charts Group Reference 126
- EG2.3 Visualizing Categorical Variables 126
- EG2.4 Visualizing Numerical Variables 128
- EG2.5 Visualizing Two Numerical Variables 131
- EG2.6 Organizing a Mix of Variables 132
- EG2.7 Visualizing a Mix of Variables 133
- EG2.8 Filtering and Querying Data 134

CHAPTER 2 JMP GUIDE 134

- JG2 JMP Choices for Creating Summaries 134
- JG2.1 Organizing Categorical Variables 135
- JG2.2 Organizing Numerical Variables 136
- JG2.3 Visualizing Categorical Variables 138
- JG2.4 Visualizing Numerical Variables 139
- JG2.5 Visualizing Two Numerical Variables 141
- JG2.6 Organizing a Mix of Variables 142
- JG2.7 Visualizing a Mix of Variables 142
- JG2.8 Filtering and Querying Data 143
- JMP Guide Gallery 144

CHAPTER 2 MINITAB GUIDE 145

- MG2.1 Organizing Categorical Variables 145
- MG2.2 Organizing Numerical Variables 145
- MG2.3 Visualizing Categorical Variables 145
- MG2.4 Visualizing Numerical Variables 147
- MG2.5 Visualizing Two Numerical Variables 149
- MG2.6 Organizing a Mix of Variables 150
- MG2.7 Visualizing a Mix of Variables 150
- MG2.8 Filtering and Querying Data 151

3 Numerical Descriptive Measures 152

USING STATISTICS: More Descriptive Choices 152

- 3.1 Measures of Central Tendency 153
 - The Mean 153
 - The Median 155
 - The Mode 156
 - The Geometric Mean 157
- 3.2 Measures of Variation and Shape 158
 - The Range 158
 - The Variance and the Standard Deviation 159
 - The Coefficient of Variation 162
 - Z Scores 162
 - Shape: Skewness 164
 - Shape: Kurtosis 164
- 3.3 Exploring Numerical Variables 169
 - Quartiles 169
 - EXHIBIT:** Rules for Calculating the Quartiles from a Set of Ranked Values 169

The Interquartile Range 171
 The Five-Number Summary 171
 The Boxplot 173

- 3.4 Numerical Descriptive Measures for a Population 175**
 The Population Mean 176
 The Population Variance and Standard Deviation 176
 The Empirical Rule 177
 Chebyshev's Theorem 178

- 3.5 The Covariance and the Coefficient of Correlation 180**
 The Covariance 180
 The Coefficient of Correlation 180

- 3.6 Descriptive Statistics: Pitfalls and Ethical Issues 184**

USING STATISTICS: More Descriptive Choices, Revisited 185

SUMMARY 185

REFERENCES 186

KEY EQUATIONS 186

KEY TERMS 186

CHECKING YOUR UNDERSTANDING 187

CHAPTER REVIEW PROBLEMS 187

CASES FOR CHAPTER 3 190

Managing Ashland MultiComm Services 190
 Digital Case 190
 CardioGood Fitness 190
 More Descriptive Choices Follow-up 191
 Clear Mountain State Student Survey 191

CHAPTER 3 EXCEL GUIDE 192

EG3.1 Measures of Central Tendency 192
 EG3.2 Measures of Variation and Shape 193
 EG3.3 Exploring Numerical Variables 193
 EG3.4 Numerical Descriptive Measures for a Population 194
 EG3.5 The Covariance and the Coefficient of Correlation 194

CHAPTER 3 JMP GUIDE 195

JG3.1 Measures of Central Tendency 195
 JG3.2 Measures of Variation and Shape 195
 JG3.3 Exploring Numerical Variables 196
 JG3.4 Numerical Descriptive Measures for a Population 196
 JG3.5 The Covariance and the Coefficient of Correlation 196

CHAPTER 3 MINITAB GUIDE 197

MG3.1 Measures of Central Tendency 197
 MG3.2 Measures of Variation and Shape 198
 MG3.3 Exploring Numerical Variables 198
 MG3.4 Numerical Descriptive Measures for a Population 199
 MG3.5 The Covariance and the Coefficient of Correlation 199

4 Basic Probability 200

USING STATISTICS: Possibilities at M&R Electronics World 200

- 4.1 Basic Probability Concepts 201**
 Events and Sample Spaces 201
 Types of Probability 202
 Summarizing Sample Spaces 203
 Simple Probability 204
 Joint Probability 205
 Marginal Probability 206
 General Addition Rule 206

- 4.2 Conditional Probability 210**
 Computing Conditional Probabilities 210
 Decision Trees 211
 Independence 213
 Multiplication Rules 214
 Marginal Probability Using the General Multiplication Rule 215

- 4.3 Ethical Issues and Probability 217**

- 4.4 Bayes' Theorem 218**

CONSIDER THIS: Divine Providence and Spam 220

- 4.5 Counting Rules 221**

USING STATISTICS: Possibilities at M&R Electronics World, Revisited 224

SUMMARY 225

REFERENCES 225

KEY EQUATIONS 225

KEY TERMS 226

CHECKING YOUR UNDERSTANDING 226

CHAPTER REVIEW PROBLEMS 226

CASES FOR CHAPTER 4 228

Digital Case 228
 CardioGood Fitness 228
 The Choice *Is* Yours Follow-Up 228
 Clear Mountain State Student Survey 228

CHAPTER 4 EXCEL GUIDE 229

EG4.1 Basic Probability Concepts 229
 EG4.4 Bayes' Theorem 229
 EG4.5 Counting Rules 229

CHAPTER 4 JMP

JG4.4 Bayes' Theorem 230

CHAPTER 4 MINITAB GUIDE 230

MG4.5 Counting Rules 230

Discrete Probability Distributions 231

USING STATISTICS: Events of Interest at Ricknel Home Centers 231

- 5.1 The Probability Distribution for a Discrete Variable 232**
 Expected Value of a Discrete Variable 232
 Variance and Standard Deviation of a Discrete Variable 233
- 5.2 Binomial Distribution 236**
EXHIBIT: Properties of the Binomial Distribution 236
 Histograms for Discrete Variables 239
 Summary Measures for the Binomial Distribution 240
- 5.3 Poisson Distribution 243**
- 5.4 Covariance of a Probability Distribution and Its Application in Finance 246**
- 5.5 Hypergeometric Distribution (*online*) 246**
- 5.6 Using the Poisson Distribution to Approximate the Binomial Distribution (*online*) 246**

USING STATISTICS: Events of Interest..., Revisited 247

SUMMARY 247

REFERENCES 247

KEY EQUATIONS 247**KEY TERMS 248****CHECKING YOUR UNDERSTANDING 248****CHAPTER REVIEW PROBLEMS 248****CASES FOR CHAPTER 5 250**

- Managing Ashland MultiComm Services 250
- Digital Case 250

CHAPTER 5 EXCEL GUIDE 251

- EG5.1 The Probability Distribution for a Discrete Variable 251
- EG5.2 Binomial Distribution 251
- EG5.3 Poisson Distribution 251

CHAPTER 5 JMP GUIDE 252

- JG5.1 The Probability Distribution for a Discrete Variable 252
- JG5.2 Binomial Distribution 252
- JG5.3 Poisson Distribution 253

CHAPTER 5 MINITAB GUIDE 253

- MG5.1 The Probability Distribution for a Discrete Variable 253
- MG5.2 Binomial Distribution 254
- MG5.3 Poisson Distribution 254

6 The Normal Distribution and Other Continuous Distributions 255

USING STATISTICS!: Normal Load Times at MyTVLab 255**6.1 Continuous Probability Distributions 256****6.2 The Normal Distribution 256**

EXHIBIT: Normal Distribution Important Theoretical Properties 257

- Role of the Mean and the Standard Deviation 258
- Calculating Normal Probabilities 259
- VISUAL EXPLORATIONS: Exploring the Normal Distribution 263
- Finding X Values 264

CONSIDER THIS: What Is Normal? 267**6.3 Evaluating Normality 269**

- Comparing Data Characteristics to Theoretical Properties 269
- Constructing the Normal Probability Plot 270

6.4 The Uniform Distribution 273**6.5 The Exponential Distribution (online) 275****6.6 The Normal Approximation to the Binomial Distribution (online) 275****USING STATISTICS: Normal Load Times..., Revisited 275****SUMMARY 275****REFERENCES 276****KEY EQUATIONS 276****KEY TERMS 276****CHECKING YOUR UNDERSTANDING 277****CHAPTER REVIEW PROBLEMS 277****CASES FOR CHAPTER 6 278**

- Managing Ashland MultiComm Services 278
- CardioGood Fitness 279
- More Descriptive Choices Follow-up 279
- Clear Mountain State Student Survey 279
- Digital Case 279

CHAPTER 6 EXCEL GUIDE 280

- EG6.2 The Normal Distribution 280
- EG6.3 Evaluating Normality 280

CHAPTER 6 JMP GUIDE 281

- JG6.2 The Normal Distribution 281
- JG6.3 Evaluating Normality 281

CHAPTER 6 MINITAB GUIDE 282

- MG6.2 The Normal Distribution 282
- MG6.3 Evaluating Normality 282

7 Sampling Distributions 284

USING STATISTICS: Sampling Oxford Cereals 284**7.1 Sampling Distributions 285****7.2 Sampling Distribution of the Mean 285**

- The Unbiased Property of the Sample Mean 285
- Standard Error of the Mean 287
- Sampling from Normally Distributed Populations 288
- Sampling from Non-normally Distributed Populations—The Central Limit Theorem 291

EXHIBIT: Normality and the Sampling Distribution of the Mean 292

VISUAL EXPLORATIONS: Exploring Sampling Distributions 295

7.3 Sampling Distribution of the Proportion 296**7.4 Sampling from Finite Populations (online) 299****USING STATISTICS: Sampling Oxford Cereals, Revisited 299****SUMMARY 300****REFERENCES 300****KEY EQUATIONS 300****KEY TERMS 300****CHECKING YOUR UNDERSTANDING 301****CHAPTER REVIEW PROBLEMS 301****CASES FOR CHAPTER 7 302**

- Managing Ashland MultiComm Services 302
- Digital Case 303

CHAPTER 7 EXCEL GUIDE 304

- EG7.2 Sampling Distribution of the Mean 304

CHAPTER 7 JMP GUIDE 305

- JG7.2 Sampling Distribution of the Mean 305

CHAPTER 7 MINITAB GUIDE 306

- MG7.2 Sampling Distribution of the Mean 306

8 Confidence Interval Estimation 307

USING STATISTICS: Getting Estimates at Ricknel Home Centers 307

- 8.1 Confidence Interval Estimate for the Mean (a Known) 308
 - Sampling Error 309
 - Can You Ever Know the Population Standard Deviation? 312
- 8.2 Confidence Interval Estimate for the Mean (a Unknown) 313
 - Student's t Distribution 314
 - The Concept of Degrees of Freedom 314
 - Properties of the f Distribution 314
 - The Confidence Interval Statement 316
- 8.3 Confidence Interval Estimate for the Proportion 321
- 8.4 Determining Sample Size 324
 - Sample Size Determination for the Mean 324
 - Sample Size Determination for the Proportion 326
- 8.5 Confidence Interval Estimation and Ethical Issues 329
- 8.6 Application of Confidence Interval Estimation in Auditing (*online*) 329
- 8.7 Estimation and Sample Size Estimation for Finite Populations (*online*) 330
- 8.8 Bootstrapping (*online*) 330

USING STATISTICS: Getting Estimates..., Revisited 330

SUMMARY¹ 330

REFERENCES 331

KEY EQUATIONS 331

KEY TERMS 331

CHECKING YOUR UNDERSTANDING 331

CHAPTER REVIEW PROBLEMS 332

CASES FOR CHAPTER 8 334

- Managing Ashland MultiComm Services 334
- Digital Case 335
- Sure Value Convenience Stores 336
- CardioGood Fitness 336
- More Descriptive Choices Follow-Up 336
- Clear Mountain State Student Survey 336

CHAPTER 8 EXCEL GUIDE 337

- EG8.1 Confidence Interval Estimate for the Mean (o Known) 337
- EG8.2 Confidence Interval Estimate for the Mean (a Unknown) 337
- EG8.3 Confidence Interval Estimate for the Proportion 338
- EG8.4 Determining Sample Size 338

CHAPTER 8 JMP GUIDE 339

- JG8.1 Confidence Interval Estimate for the Mean (a Known) 339
- JG8.2 Confidence Interval Estimate for the Mean (α Unknown) 339
- JG8.3 Confidence Interval Estimate for the Proportion 340
- JG8.4 Determining Sample Size 340

CHAPTER 8 MINITAB GUIDE 341

- MG8.1 Confidence Interval Estimate for the Mean (a Known) 341
- MG8.2 Confidence Interval Estimate for the Mean (a Unknown) 341

- MG8.3 Confidence Interval Estimate for the Proportion 342
- MG8.4 Determining Sample Size 342

Fundamentals of Hypothesis Testing: One-Sample Tests 343

USING STATISTICS: Significant Testing at Oxford Cereals 343

- 9.1 Fundamentals of Hypothesis Testing 344
 - EXHIBIT:** Fundamental Hypothesis Testing Concepts 345
 - The Critical Value of the Test Statistic 345
 - Regions of Rejection and Nonrejection 346
 - Risks in Decision Making Using Hypothesis Testing 346
 - Z Test for the Mean (a Known) 348
 - Hypothesis Testing Using the Critical Value Approach 349
 - EXHIBIT:** The Critical Value Approach to Hypothesis Testing 350
 - Hypothesis Testing Using the p-Value Approach 352
 - EXHIBIT:** The p-Value Approach to Hypothesis Testing 353
 - A Connection Between Confidence Interval Estimation and Hypothesis Testing 354
 - Can You Ever Know the Population Standard Deviation? 355
- 9.2 t Test of Hypothesis for the Mean (a Unknown) 356
 - The Critical Value Approach 357
 - p-Value Approach 358
 - Checking the Normality Assumption 359
- 9.3 One-Tail Tests 362
 - The Critical Value Approach 362
 - The p-Value Approach 363
 - EXHIBIT:** The Null and Alternative Hypotheses in One-Tail Tests 365
- 9.4 Z Test of Hypothesis for the Proportion 366
 - The Critical Value Approach 367
 - The p-Value Approach 368
- 9.5 Potential Hypothesis-Testing Pitfalls and Ethical Issues 370
 - EXHIBIT:** Questions for the Planning Stage of Hypothesis Testing 370
 - Statistical Significance Versus Practical Significance 370
 - Statistical *Insufficiency* Versus Importance 371
 - Reporting of Findings 371
 - Ethical Issues 371
- 9.6 Power of the Test (*online*) 371

USING STATISTICS: Significant Testing ... Revisited 372

SUMMARY 372

REFERENCES 372

KEY EQUATIONS 373

KEY TERMS 373

CHECKING YOUR UNDERSTANDING 373

CHAPTER REVIEW PROBLEMS 373

CASES FOR CHAPTER 9 375

- Managing Ashland MultiComm Services 375
- Digital Case 375
- Sure Value Convenience Stores 376

CHAPTER 9 EXCEL GUIDE 377

- EG9.1 Fundamentals of Hypothesis Testing 377
- EG9.2 t Test of Hypothesis for the Mean (σ Unknown) 377
- EG9.3 One-Tail Tests 378
- EG9.4 Z Test of Hypothesis for the Proportion 378

CHAPTER 9 JMP GUIDE 379

- JG9.1 Fundamentals of Hypothesis Testing 379
- JG9.2 t Test of Hypothesis for the Mean (σ Unknown) 379
- JG9.3 One-Tail Tests 380
- JG9.4 Z Test of Hypothesis for the Proportion 380

CHAPTER 9 MINITAB GUIDE 380

- MG9.1 Fundamentals of Hypothesis Testing 380
- MG9.2 t Test of Hypothesis for the Mean (σ Unknown) 381
- MG9.3 One-Tail Tests 381
- MG9.4 Z Test of Hypothesis for the Proportion 381

10 Two-Sample Tests 383**USING STATISTICS: Differing Means for Selling Streaming Media Players at Arlingtons? 383**

- 10.1** Comparing the Means of Two Independent Populations 384
 - Pooled-Variance t Test for the Difference Between Two Means Assuming Equal Variances 384
 - Evaluating the Normality Assumption 387
 - Confidence Interval Estimate for the Difference Between Two Means 389
 - Separate-Variance t Test for the Difference Between Two Means, Assuming Unequal Variances 390
- 10.2** Comparing the Means of Two Related Populations 393
 - Paired t Test 394
 - Confidence Interval Estimate for the Mean Difference 399
- 10.3** Comparing the Proportions of Two Independent Populations 401
 - Z Test for the Difference Between Two Proportions 401
 - Confidence Interval Estimate for the Difference Between Two Proportions 406
- 10.4** F Test for the Ratio of Two Variances 408
- 10.5** Effect Size (*online*) 412

USING STATISTICS: Differing Means for Selling Revisited 413**SUMMARY 413****REFERENCES 414****KEY EQUATIONS 414****KEY TERMS 415****CHECKING YOUR UNDERSTANDING 415****CHAPTER REVIEW PROBLEMS 415****CASES FOR CHAPTER 10 417**

- Managing Ashland MultiComm Services 417
- Digital Case 418
- Sure Value Convenience Stores 418
- CardioGood Fitness 418
- More Descriptive Choices Follow-Up 418
- Clear Mountain State Student Survey 419

CHAPTER 10 EXCEL GUIDE 420

- EG10.1 Comparing the Means of Two Independent Populations 420
- EG10.2 Comparing the Means of Two Related Populations 422
- EG10.3 Comparing the Proportions of Two Independent Populations 423
- EG10.4 F Test for the Ratio of Two Variances 424

CHAPTER 10 JMP GUIDE 425

- JG10.1 Comparing the Means of Two Independent Populations 425
- JG10.2 Comparing the Means of Two Related Populations 426
- JG10.3 Comparing the Proportions of Two Independent Populations 426
- JG10.4 F Test for the Ratio of Two Variances 426

CHAPTER 10 MINITAB GUIDE 427

- MG10.1 Comparing the Means of Two Independent Populations 427
- MG10.2 Comparing the Means of Two Related Populations 428
- MG10.3 Comparing the Proportions of Two Independent Populations 428
- MG10.4 F Test for the Ratio of Two Variances 429

11 Analysis of Variance 430**USING STATISTICS: The Means to Find Differences at Arlingtons 430**

- 11.1** The Completely Randomized Design: One-Way ANOVA 431
 - Analyzing Variation in One-Way ANOVA 432
 - F Test for Differences Among More Than Two Means 434
 - One-Way ANOVA F Test Assumptions 439
 - Levene Test for Homogeneity of Variance 439
 - Multiple Comparisons: The Tukey-Kramer Procedure 441
 - The Analysis of Means (ANOM) 443
- 11.2** The Factorial Design: Two-Way ANOVA 446
 - Factor and Interaction Effects 447
 - Testing for Factor and Interaction Effects 448
 - Multiple Comparisons: The Tukey Procedure 452
 - Visualizing Interaction Effects: The Cell Means Plot 453
 - Interpreting Interaction Effects 454
- 11.3** The Randomized Block Design (*online*) 458
- 11.4** Fixed Effects, Random Effects, and Mixed Effects Models (*online*) 458

USING STATISTICS: The Means to Find Differences at Arlingtons Revisited 458**SUMMARY 458****REFERENCES 459****KEY EQUATIONS 459****KEY TERMS 460****CHECKING YOUR UNDERSTANDING 460****CHAPTER REVIEW PROBLEMS 460****CASES FOR CHAPTER 11 462**

- Managing Ashland MultiComm Services 462
- Digital Case 463
- Sure Value Convenience Stores 463
- CardioGood Fitness 463
- More Descriptive Choices Follow-Up 463
- Clear Mountain State Student Survey 463

CHAPTER 11 EXCEL GUIDE 464

- EG11.1 The Completely Randomized Design: One-Way ANOVA 464
- EG11.2 The Factorial Design: Two-Way ANOVA 466

CHAPTER 11 JMP GUIDE 467

- JG11.1 The Completely Randomized Design: One-Way ANOVA 467
- JG11.2 The Factorial Design: Two-Way ANOVA 468

CHAPTER 11 MINITAB GUIDE 469

- MG11.1 The Completely Randomized Design: One-Way ANOVA 469
- MG11.2 The Factorial Design: Two-Way ANOVA 470

12 Chi-Square and Nonparametric Tests 472

USING STATISTICS: Avoiding Guesswork About Resort Guests 472

- 12.1** Chi-Square Test for the Difference Between Two Proportions 473
- 12.2** Chi-Square Test for Differences Among More Than Two Proportions 480
 - The Marascuilo Procedure 483
 - The Analysis of Proportions (ANOP) 485
- 12.3** Chi-Square Test of Independence 486
- 12.4** Wilcoxon Rank Sum Test for Two Independent Populations 492
- 12.5** Kruskal-Wallis Rank Test for the One-Way ANOVA 498
 - Assumptions of the Kruskal-Wallis Rank Test 501
- 12.6** McNemar Test for the Difference Between Two Proportions (Related Samples) (*online*) 502
- 12.7** Chi-Square Test for the Variance or Standard Deviation (*online*) 502
- 12.8** Wilcoxon Signed Ranks Test for Two Related Populations (*online*) 503
- 12.9** Friedman Rank Test for the Randomized Block Design (*online*) 503

USING STATISTICS: Avoiding Guesswork..., Revisited 503**SUMMARY 503****REFERENCES 504****KEY EQUATIONS 504****KEY TERMS 505****CHECKING YOUR UNDERSTANDING 505****CHAPTER REVIEW PROBLEMS 505****CASES FOR CHAPTER 12 507**

- Managing Ashland MultiComm Services 507
- Digital Case 508
- Sure Value Convenience Stores 508
- CardioGood Fitness 508
- More Descriptive Choices Follow-Up 509
- Clear Mountain State Student Survey 509

CHAPTER 12 EXCEL GUIDE 510

- EG12.1 Chi-Square Test for the Difference Between Two Proportions 510
- EG12.2 Chi-Square Test for Differences Among More Than Two Proportions 510

- EG12.3 Chi-Square Test of Independence 511
- EG12.4 Wilcoxon Rank Sum Test: A Nonparametric Method for Two Independent Populations 511
- EG12.5 Kruskal-Wallis Rank Test: A Nonparametric Method for the One-Way ANOVA 512

CHAPTER 12 JMP GUIDE 513

- JG12.1 Chi-Square Test for the Difference Between Two Proportions 513
- JG12.2 Chi-Square Test for Difference Among More Than Two Proportions 513
- JG12.3 Chi-Square Test Of Independence 513
- JG12.4 Wilcoxon Rank Sum Test for Two Independent Populations 513
- JG12.5 Kruskal-Wallis Rank Test for the One-Way ANOVA 514

CHAPTER 12 MINITAB GUIDE 514

- MG12.1 Chi-Square Test for the Difference Between Two Proportions 514
- MG12.2 Chi-Square Test for Differences Among More Than Two Proportions 515
- MG12.3 Chi-Square Test of Independence 515
- MG12.4 Wilcoxon Rank Sum Test: A Nonparametric Method for Two Independent Populations 515
- MG12.5 Kruskal-Wallis Rank Test: A Nonparametric Method for the One-Way ANOVA 515

13 Simple Linear Regression 516

USING STATISTICS: Knowing Customers at Sunflowers Apparel 516

Preliminary Analysis 517

- 13.1** Simple Linear Regression Models 518
- 13.2** Determining the Simple Linear Regression Equation 519
 - The Least-Squares Method 519
 - Predictions in Regression Analysis: Interpolation Versus Extrapolation 522
 - Computing the Y Intercept, b_0 and the Slope, b^1 523
 - VISUAL EXPLORATIONS: Exploring Simple Linear Regression Coefficients 525
- 13.3** Measures of Variation 527
 - Computing the Sum of Squares 527
 - The Coefficient of Determination 528
 - Standard Error of the Estimate 530
- 13.4** Assumptions of Regression 532
- 13.5** Residual Analysis 532
 - Evaluating the Assumptions 532
- 13.6** Measuring Autocorrelation: The Durbin-Watson Statistic 536
 - Residual Plots to Detect Autocorrelation 536
 - The Durbin-Watson Statistic 537
- 13.7** Inferences About the Slope and Correlation Coefficient 540
 - t Test for the Slope 540
 - F Test for the Slope 541
 - Confidence Interval Estimate for the Slope 543
 - t Test for the Correlation Coefficient 543

- 13.8** Estimation of Mean Values and Prediction of Individual Values 546
 The Confidence Interval Estimate for the Mean Response 547
 The Prediction Interval for an Individual Response 548
- 13.9** Potential Pitfalls in Regression 550
EXHIBIT: Seven Steps for Avoiding the Potential Pitfalls 550
- USING STATISTICS: Knowing Customers..., Revisited 552**
- SUMMARY 553**
- REFERENCES 554**
- KEY EQUATIONS 554**
- KEY TERMS 555**
- CHECKING YOUR UNDERSTANDING 555**
- CHAPTER REVIEW PROBLEMS 556**
- CASES FOR CHAPTER 13 559**
 Managing Ashland MultiComm Services 559
 Digital Case 559
 Rye Packaging 560
- CHAPTER 13 EXCEL GUIDE 561**
 EG13.2 Determining the Simple Linear Regression Equation 561
 EG13.3 Measures of Variation 562
 EG13.4 Assumptions of Regression 562
 EG13.5 Residual Analysis 562
 EG13.6 Measuring Autocorrelation: the Durbin-Watson Statistic 563
 EG13.7 Inferences about the Slope and Correlation Coefficient 563
 EG13.8 Estimation of Mean Values and Prediction of Individual Values 563
- CHAPTER 13 JMP GUIDE 564**
 JG13.2 Determining the Simple Linear Regression Equation 564
 JG13.3 Measures of Variation 564
 JG13.4 Assumptions of Regression 564
 JG13.5 Residual Analysis 564
 JG13.6 Measuring Autocorrelation: the Durbin-Watson Statistic 564
 JG13.7 Inferences about the Slope and Correlation Coefficient 564
 JG13.8 Estimation of Mean Values and Prediction of Individual Values 565
- CHAPTER 13 MINITAB GUIDE 566**
 MG13.2 Determining the Simple Linear Regression Equation 566
 MG13.3 Measures of Variation 567
 MG13.4 Assumptions of Regression 567
 MG13.5 Residual Analysis 567
 MG13.6 Measuring Autocorrelation: the Durbin-Watson Statistic 567
 MG13.7 Inferences about the Slope and Correlation Coefficient 567
 MG13.8 Estimation of Mean Values and Prediction of Individual Values 567

14 Introduction to Multiple Regression 568

- USING STATISTICS: The Multiple Effects of OmniPower Bars 568**
- 14.1** Developing a Multiple Regression Model 569
 Interpreting the Regression Coefficients 570
 Predicting the Dependent Variable Y 572
- 14.2** r^2 , Adjusted r^2 , and the Overall F Test 574
 Coefficient of Multiple Determination 574
 Adjusted r^2 575
 Test for the Significance of the Overall Multiple Regression Model 575

- 14.3** Multiple Regression Residual Analysis 578
- 14.4** Inferences About the Population Regression Coefficients 579
 Tests of Hypothesis 580
 Confidence Interval Estimation 581
- 14.5** Testing Portions of the Multiple Regression Model 583
 Coefficients of Partial Determination 587
- 14.6** Using Dummy Variables and Interaction Terms 589
 Interactions 592
- 14.7** Logistic Regression 601
- 14.8** Influence Analysis (*online*) 607
- USING STATISTICS: The Multiple Effects..., Revisited 607**
- SUMMARY 607**
- REFERENCES 609**
- KEY EQUATIONS 609**
- KEY TERMS 610**
- CHECKING YOUR UNDERSTANDING 610**
- CHAPTER REVIEW PROBLEMS 610**
- CASES FOR CHAPTER 14 613**
 Managing Ashland MultiComm Services 613
 Digital Case 613
- CHAPTER 14 EXCEL GUIDE 614**
 EG14.1 Developing a Multiple Regression Model 614
 EG14.2 r^2 , Adjusted r^2 , and the Overall F Test 615
 EG14.3 Multiple Regression Residual Analysis 615
 EG14.4 Inferences about the Population Regression Coefficients 616
 EG14.5 Testing Portions of the Multiple Regression Model 616
 EG14.6 Using Dummy Variables and Interaction Terms 616
 EG14.7 Logistic Regression 617
- CHAPTER 14 JMP GUIDE 617**
 JG14.1 Developing a Multiple Regression Model 617
 JG14.2 r^2 , Adjusted r^2 , and the Overall F Test Measures of Variation 618
 JG14.3 Multiple Regression Residual Analysis 618
 JG14.4 Inferences about the Population 618
 JG14.5 Testing Portions of the Multiple Regression Model 619
 JG14.6 Using Dummy Variables and Interaction Terms 619
 JG14.7 Logistic Regression 619
- CHAPTER 14 MINITAB GUIDE 620**
 MG14.1 Developing a Multiple Regression Model 620
 MG14.2 r^2 , Adjusted r^2 , and the Overall F Test 621
 MG14.3 Multiple Regression Residual Analysis 621
 MG14.4 Inferences about the Population Regression Coefficients 621
 MG14.5 Testing Portions of the Multiple Regression Model 621
 MG14.6 Using Dummy Variables and Interaction Terms in Regression Models 621
 MG14.7 Logistic Regression 622
 MG14.8 Influence Analysis 623

15 Multiple Regression Model Building 624

- USING STATISTICS: Valuing Parsimony at WSTA-TV 624**
- 15.1** Quadratic Regression Model 625
 Finding the Regression Coefficients and Predicting Y 626
 Testing for the Significance of the Quadratic Model 628

- Testing the Quadratic Effect 629
- The Coefficient of Multiple Determination 631
- 15.2 Using Transformations in Regression Models 633**
 - The Square-Root Transformation 633
 - The Log Transformation 635
- 15.3 Collinearity 637**
- 15.4 Model Building 639**
 - EXHIBIT: Successful Model Building 639**
 - The Stepwise Regression Approach to Model Building 641
 - The Best Subsets Approach to Model Building 642
 - Model Validation 645
- 15.5 Pitfalls in Multiple Regression and Ethical Issues 647**
 - Pitfalls in Multiple Regression 647
 - Ethical Issues 648

USING STATISTICS: Valuing Parsimony..., Revisited 648

SUMMARY 649

REFERENCES 650

KEY EQUATIONS 650

KEY TERMS 650

CHECKING YOUR UNDERSTANDING 650

CHAPTER REVIEW PROBLEMS 650

CASES FOR CHAPTER 15 652

- The Mountain States Potato Company 652
- Sure Value Convenience Stores 653
- Digital Case 653
- The Craybill Instrumentation Company Case 653
- More Descriptive Choices Follow-Up 654

CHAPTER 15 EXCEL GUIDE 655

- EG15.1 The Quadratic Regression Model 655
- EG15.2 Using Transformations in Regression Models 655
- EG15.3 Collinearity 656
- EG15.4 Model Building 656

CHAPTER 15 JMP GUIDE 657

- JG15.1 The Quadratic Regression Model 657
- JG15.2 Using Transformations in Regression Models 657
- JG15.3 Collinearity 657
- JG15.4 Model Building 657

CHAPTER 15 MINITAB GUIDE 658

- MG15.1 The Quadratic Regression Model 658
- MG15.2 Using Transformations in Regression Models 659
- MG15.3 Collinearity 659
- MG15.4 Model Building 659

16 Time-Series Forecasting 661

USING STATISTICS: Is the ByYourDoor Service Trending? 661

- 16.1 Time Series Component Factors 662**
- 16.2 Smoothing an Annual Time Series 664**
 - Moving Averages 665
 - Exponential Smoothing 667
- 16.3 Least-Squares Trend Fitting and Forecasting 669**
 - The Linear Trend Model 669
 - The Quadratic Trend Model 671
 - The Exponential Trend Model 672
 - Model Selection Using First, Second, and Percentage Differences 674

EXHIBIT: Model Selection Using First, Second, and Percentage Differences 674

- 16.4 Autoregressive Modeling for Trend Fitting and Forecasting 679**
 - Selecting an Appropriate Autoregressive Model 680
 - Determining the Appropriateness of a Selected Model 681
 - EXHIBIT: Autoregressive Modeling Steps 683**
- 16.5 Choosing an Appropriate Forecasting Model 687**
 - Residual Analysis 687
 - The Magnitude of the Residuals Through Squared or Absolute Differences 688
 - The Principle of Parsimony 688
 - A Comparison of Four Forecasting Methods 689
- 16.6 Time-Series Forecasting of Seasonal Data 691**
 - Least-Squares Forecasting with Monthly or Quarterly Data 691
- 16.7 Index Numbers (online) 697**
- CONSIDER THIS: Let the Model User Beware 697**
- USING STATISTICS: Is the ByYourDoor..., Revisited 697**
- SUMMARY 697**
- REFERENCES 698**
- KEY EQUATIONS 698**
- KEY TERMS 699**
- CHECKING YOUR UNDERSTANDING 700**
- CHAPTER REVIEW PROBLEMS 700**
- CASES FOR CHAPTER 16 701**
 - Managing Ashland MultiComm Services 701
 - Digital Case 701
- CHAPTER 16 EXCEL GUIDE 702**
 - EG16.2 Smoothing an Annual Time Series 702
 - EG16.3 Least-Squares Trend Fitting and Forecasting 703
 - EG16.4 Autoregressive Modeling for Trend Fitting and Forecasting 703
 - EG16.5 Choosing an Appropriate Forecasting Model 704
 - EG16.6 Time-Series Forecasting of Seasonal Data 704
- CHAPTER 16 JMP GUIDE 705**
 - JG16.2 Smoothing an Annual Time Series 705
 - JG16.3 Least-Squares Trend Fitting and Forecasting 706
 - JG16.4 Autoregressive Modeling for Trend Fitting and Forecasting 706
 - JG16.5 Choosing an Appropriate Forecasting Model 707
 - JG16.6 Time-Series Forecasting of Seasonal Data 707
- CHAPTER 16 MINITAB GUIDE 707**
 - MG16.2 Smoothing an Annual Time Series 707
 - MG16.3 Least-Squares Trend Fitting and Forecasting 708
 - MG16.4 Autoregressive Modeling for Trend Fitting and Forecasting 709
 - MG16.5 Choosing an Appropriate Forecasting Model 709
 - MG16.6 Time-Series Forecasting of Seasonal Data 709

17 Business Analytics 710

USING STATISTICS: Back to Arlingtons for the Future 710

- 17.1 Business Analytics Categories 711**
 - Inferential Statistics and Predictive Analytics 712
 - Supervised and Unsupervised Methods 712
- CONSIDER THIS: What's My Major if I Want to be a Data Miner? 713**

17.2 Descriptive Analytics 714
Dashboards 714
Data Dimensionality and Descriptive Analytics 715
17.3 Predictive Analytics for Prediction 716
17.4 Predictive Analytics for Classification 719
17.5 Predictive Analytics for Clustering 720
17.6 Predictive Analytics for Association 723
Multidimensional scaling (MDS) 724
17.7 Text Analytics 725
17.8 Prescriptive Analytics 726
USING STATISTICS: Back to Arlingtons..., Revisited 727
REFERENCES 727
KEY EQUATIONS 728
KEY TERMS 728
CHECKING YOUR UNDERSTANDING 728
CHAPTER REVIEW PROBLEMS 728
CASES FOR CHAPTER 17 730
The Mountain States Potato Company 730
The Craybill Instrumentation Company 730
CHAPTER 17 SOFTWARE GUIDE 731
Introduction 731
SG17.2 Descriptive Analytics 731
SG17.3 Predictive Analytics for Prediction 733
SG17.4 Predictive Analytics for Classification 733
SG17.5 Predictive Analytics for Clustering 734
SG17.6 Predictive Analytics for Association 734

18 Getting Ready to Analyze Data in the Future 736

USING STATISTICS: Mounting Future Analyses 736

18.1 Analyzing Numerical Variables 737
EXHIBIT: Questions to Ask When Analyzing Numerical Variables 737
Describe the Characteristics of a Numerical Variable? 737
Reach Conclusions About the Population Mean or the Standard Deviation? 737
Determine Whether the Mean and/or Standard Deviation Differs Depending on the Group? 738
Determine Which Factors Affect the Value of a Variable? 738
Predict the Value of a Variable Based on the Values of Other Variables? 739
Classify or Associate Items 739
Determine Whether the Values of a Variable Are Stable Over Time? 739
18.2 Analyzing Categorical Variables 739
EXHIBIT: Questions to Ask When Analyzing Categorical Variables 739
Describe the Proportion of Items of Interest in Each Category? 739
Reach Conclusions About the Proportion of Items of Interest? 740
Determine Whether the Proportion of Items of Interest Differs Depending on the Group? 740

Predict the Proportion of Items of Interest Based on the Values of Other Variables? 740
Classify or Associate Items 740
Determine Whether the Proportion of Items of Interest Is Stable Over Time? 740

USING STATISTICS: The Future to Be Visited 741

CHAPTER REVIEW PROBLEMS 741

19 Statistical Applications in Quality Management (online) 19-1

USING STATISTICS: Finding Quality at the Beachcomber 19-1

19.1 The Theory of Control Charts 19-2
19.2 Control Chart for the Proportion: The p Chart 19-4
19.3 The Red Bead Experiment: Understanding Process Variability 19-10
19.4 Control Chart for an Area of Opportunity: The c Chart 19-12
19.5 Control Charts for the Range and the Mean 19-15
The R Chart 19-16
The X Chart 19-18
19.6 Process Capability 19-15
Customer Satisfaction and Specification Limits 19-21
Capability Indices 19-23
CPL , CPU , and C_{pk} 19-24
19.7 Total Quality Management 19-26
19.8 Six Sigma 19-28
The DMAIC Model 19-29
Roles in a Six Sigma Organization 19-30
Lean Six Sigma 19-30

USING STATISTICS: Finding Quality at the Beachcomber, Revisited 19-31

SUMMARY 19-31

REFERENCES 19-32

KEY EQUATIONS 19-32

KEY TERMS 19-33

CHAPTER REVIEW PROBLEMS 19-34

CASES FOR CHAPTER 19 19-36

The Harnswell Sewing Machine Company Case 19-36

Managing Ashland Multicomm Services 19-38

CHAPTER 19 EXCEL GUIDE 19-39

EG19.2 Control Chart for the Proportion: The p Chart 19-39
EG19.4 Control Chart for an Area of Opportunity: The c Chart 19-40
EG19.5 Control Charts for the Range and the Mean 19-41
EG19.6 Process Capability 19-42

CHAPTER 19 JMP GUIDE 19-43

JG19.2 Control Chart for the Proportion: The p Chart 19-43
JG19.4 Control Chart for an Area of Opportunity: The c Chart 19-44
JG19.5 Control Charts for the Range and the Mean 19-45
JG19.6 Process Capability 19-46

CHAPTER 19 MINITAB GUIDE 19-47

- MG19.2 Control Chart for the Proportion: The p Chart 19-48
- MG19.4 Control Chart for an Area of Opportunity:
The c Chart 19-49
- MG19.5 Control Charts for the Range and the Mean 19-50
- MG19.6 Process Capability 19-51

20 Decision Making (online) 20-1**USING STATISTICS: Reliable Decision Making 20-1****20.1 Payoff Tables and Decision Trees 20-2****20.2 Criteria for Decision Making 20-2**

- Maximax Payoff 20-6
- Maximin Payoff 20-7
- Expected Monetary Value 20-7
- Expected Opportunity Loss 20-9
- Return-to-Risk Ratio 20-11

20.3 Decision Making with Sample Information 20-16**20.4 Utility 20-21****CONSIDER THIS: Risky Business 20-22****USING STATISTICS: Reliable Decision-Making,
Revisited 20-22**

SUMMARY 20-23

REFERENCES 20-23

KEY EQUATIONS 20-23

KEY TERMS 20-23

CHAPTER REVIEW PROBLEMS 20-23

CASES FOR CHAPTER 20 20-26

- Digital Case 20-26

CHAPTER 20 EXCEL GUIDE 20-27

- EG20.1 Payoff Tables and Decision Trees 20-27
- EG20.2 Criteria for Decision Making 20-27

Appendices 743**A. Basic Math Concepts and Symbols 744**

- A.1 Operators 744
- A.2 Rules for Arithmetic Operations 744
- A.3 Rules for Algebra: Exponents and Square Roots 744
- A.4 Rules for Logarithms 745
- A.5 Summation Notation 746
- A.6 Greek Alphabet 749

B. Important Software Skills and Concepts 750

- B.1 Identifying the Software Version 750
- B.2 Formulas 750
- B.3 Excel Cell References 752
- B.4 Excel Worksheet Formatting 753
- B.5E Excel Chart Formatting 754
- B.5J JMP Chart Formatting 755
- B.5M Minitab Chart Formatting 756

- B.6 Creating Histograms for Discrete Probability Distributions (Excel) 756

- B.7 Deleting the "Extra" Histogram Bar (Excel) 757

C. Online Resources 758

- C.1 About the Online Resources for This Book 758
- C.2 Data Files 758
- C.3 Files Integrated With Microsoft Excel 765
- C.4 Supplemental Files 765

D. Configuring Software 766

- D.1 Microsoft Excel Configuration 766
- D.2 JMP Configuration 768
- D.3 Minitab Configuration 768

E. Table 769

- E.1 Table of Random Numbers 769
- E.2 The Cumulative Standardized Normal Distribution 771
- E.3 Critical Values of t 773
- E.4 Critical Values of χ^2 775
- E.5 Critical Values of F 776
- E.6 Lower and Upper Critical Values, T_j , of the Wilcoxon Rank Sum Test 780
- E.7 Critical Values of the Studentized Range, Q 781
- E.8 Critical Values, d_L and d_U , of the Durbin-Watson Statistic, D (Critical Values Are One-Sided) 783
- E.9 Control Chart Factors 784
- E.10 The Standardized Normal Distribution 785

F. Useful Knowledge 786

- F.1 Keyboard Shortcuts 786
- F.2 Understanding the Nonstatistical Functions 786

G. Software FAQs 788

- G.1 Microsoft Excel FAQs 788
- G.2 PHStat FAQs 788
- G.3 JMP FAQs 789
- G.4 Minitab FAQs 789

H. All About PHStat 790

- H.1 What is PHStat? 790
- H.2 Obtaining and Setting Up PHStat 791
- H.3 Using PHStat 791
- H.4 PHStat Procedures, by Category 792

Self-Test Solutions and Answers to Selected Even-Numbered Problems 793**Index 825****Credits 837**