To Ann, Bonnie, Clara, Michael, and Mildred
Contents in Brief

Part I
Introduction and Descriptive Statistics

1. Introduction: Defining the Role of Statistics in Business 3
2. Data Structures: Classifying the Various Types of Data Sets 19
3. Histograms: Looking at the Distribution of Data 35
4. Landmark Summaries: Interpreting Typical Values and Percentiles 65
5. Variability: Dealing with Diversity 95

Part II
Probability

6. Probability: Understanding Random Situations 125
7. Random Variables: Working with Uncertain Numbers 155

Part III
Statistical Inference

8. Random Sampling: Planning Ahead for Data Gathering 189
9. Confidence Intervals: Admitting That Estimates Are Not Exact 219
10. Hypothesis Testing: Deciding between Reality and Coincidence 249

Part IV
Regression and Time Series

11. Correlation and Regression: Measuring and Predicting Relationships 291
12. Multiple Regression: Predicting One Variable from Several Others 347
13. Report Writing: Communicating the Results of a Multiple Regression 417
14. Time Series: Understanding Changes over Time 429

Part V
Methods and Applications

15. ANOVA: Testing for Differences among Many Samples, and Much More 467
16. Nonparametrics: Testing with Ordinal Data or Nonnormal Distributions 491
17. Chi-Squared Analysis: Testing for Patterns in Qualitative Data 507
18. Quality Control: Recognizing and Managing Variation 523

Appendix A Employee Database 545
Appendix B Donations Database 547
Appendix C Self-Test: Solutions to Selected Problems and Database Exercises 551
Appendix D Statistical Tables 565
Glossary 597
Index 605
## Preface

Preface xv

## About the Author

About the Author xix

## Part I

### Introduction and Descriptive Statistics

#### 1. Introduction: Defining the Role of Statistics in Business

1.1 Why Statistics? 3
   - Why Should You Learn Statistics? 3
   - Is Statistics Difficult? 4
   - Does Learning Statistics Decrease Your Decision-Making Flexibility? 4

1.2 What Is Statistics? 4
   - Statistics Looks at the Big Picture
   - Statistics Doesn’t Ignore the Individual
   - Looking at Data
   - Statistics in Management

1.3 The Five Basic Activities of Statistics 5
   - Designing a Plan for Data Collection
   - Exploring the Data
   - Modeling the Data
   - Estimating an Unknown Quantity
   - Hypothesis Testing

1.4 Data Mining 8

1.5 What Is Probability? 14

1.6 General Advice 14

1.7 End-of-Chapter Materials 14
   - Summary
   - Key Words
   - Questions
   - Problems
   - Project

#### 2. Data Structures: Classifying the Various Types of Data Sets

2.1 How Many Variables? 19
   - Univariate Data

#### 2.2 Quantitative Data: Numbers

- Discrete Quantitative Data 21
- Continuous Quantitative Data 21
- Watch Out for Meaningless Numbers 22

#### 2.3 Qualitative Data: Categories

- Ordinal Qualitative Data 22
- Nominal Qualitative Data 23

#### 2.4 Time-Series and Cross-Sectional Data

#### 2.5 Sources of Data, Including the Internet

#### 2.6 End-of-Chapter Materials

- Summary 29
- Key Words 30
- Questions 30
- Problems 30
- Database Exercises 34
- Projects 34

#### 3. Histograms: Looking at the Distribution of Data

3.1 A List of Data 35
   - The Number Line 36

3.2 Using a Histogram to Display the Frequencies 37
   - Histograms and Bar Charts 39

3.3 Normal Distributions 40

3.4 Skewed Distributions and Data Transformation 43
   - The Trouble with Skewness 45
   - Transformation to the Rescue 45
   - Interpreting and Computing the Logarithm 46

3.5 Bimodal Distributions 47
   - Is It Really Bimodal? 47

3.6 Outliers 49
   - Dealing with Outliers 49

3.7 Data Mining with Histograms 53

3.8 Histograms by Hand: Stem-and-Leaf 54

3.9 End-of-Chapter Materials 55
   - Summary 55
   - Key Words 56
4. Landmark Summaries: Interpreting Typical Values and Percentiles

4.1 What Is the Most Typical Value? 65
The Average: A Typical Value for Quantitative Data 66
The Weighted Average: Adjusting for Importance 68
The Median: A Typical Value for Quantitative and Ordinal Data 70
The Mode: A Typical Value Even for Nominal Data 73
Which Summary Should You Use? 75
4.2 What Percentile Is It? 76
Extremes, Quartiles, and Box Plots 76
The Cumulative Distribution Function Displays the Percentiles 80
4.3 End-of-Chapter Materials 83
Summary 83
Key Words 84
Questions 84
Problems 85
Database Exercises 91
Projects 91
Case 92

5. Variability: Dealing with Diversity

5.1 The Standard Deviation: The Traditional Choice 96
Definition and Formula for the Standard Deviation and the Variance 97
Using a Calculator or a Computer 98
Interpreting the Standard Deviation 98
Interpreting the Standard Deviation for a Normal Distribution 99
The Sample and the Population Standard Deviations 106
5.2 The Range: Quick and Superficial 107
5.3 The Coefficient of Variation: A Relative Variability Measure 108
5.4 Effects of Adding to or Rescaling the Data 109
5.5 End-of-Chapter Materials 110
Summary 110
Key Words 112
Questions 112

6. Probability: Understanding Random Situations

6.1 An Example: Is It behind Door Number 1, Door Number 2, or Door Number 3? 126
6.2 How Can You Analyze Uncertainty? 127
The Random Experiment: A Precise Definition of a Random Situation 127
The Sample Space: A List of What Might Happen 127
The Outcome: What Actually Happens Events: Either They Happen or They Don't 128
6.3 How Likely Is an Event? 129
Every Event Has a Probability 129
Where Do Probabilities Come From? 130
Relative Frequency and the Law of Large Numbers 130
Theoretical Probability 131
The Equally Likely Rule 131
Subjective Probability 132
Bayesian and Non-Bayesian Analysis 132
6.4 How Can You Combine Information about More Than One Event? 133
Venn Diagrams Help You See All the Possibilities 133
Not an Event 133
The Complement (Not) Rule 134
One Event and Another 134
What If Both Events Can’t Happen at Once? 134
The Intersection (and) Rule for Mutually Exclusive Events 135
One Event or Another 135
The Union (or) Rule for Mutually Exclusive Events 135
Finding or from and and Vice Versa 135
One Event Given Another: Reflecting Current Information 136
The Rule for Finding a Conditional Probability Given Certain Information 137
Conditional Probabilities for Mutually Exclusive Events 138
Independent Events 138
### Part III: Statistical Inference

#### 8. Random Sampling: Planning Ahead for Data Gathering

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.1 Populations and Samples</td>
<td>190</td>
</tr>
<tr>
<td>What Is a Representative Sample?</td>
<td>190</td>
</tr>
<tr>
<td>A Sample Statistic and a Population Parameter</td>
<td>191</td>
</tr>
<tr>
<td>8.2 The Random Sample</td>
<td>192</td>
</tr>
<tr>
<td>Selecting a Random Sample</td>
<td>192</td>
</tr>
<tr>
<td>Sampling by Shuffling the Population</td>
<td>195</td>
</tr>
<tr>
<td>8.3 The Sampling Distribution and the Central Limit Theorem</td>
<td>196</td>
</tr>
<tr>
<td>8.4 A Standard Error Is an Estimated Standard Deviation</td>
<td>200</td>
</tr>
<tr>
<td>How Close Is the Sample Average to the Population Mean?</td>
<td>201</td>
</tr>
<tr>
<td>Correcting for Small Populations</td>
<td>203</td>
</tr>
<tr>
<td>The Standard Error of the Binomial Proportion</td>
<td>204</td>
</tr>
<tr>
<td>8.5 Other Sampling Methods</td>
<td>205</td>
</tr>
<tr>
<td>The Stratified Random Sample</td>
<td>205</td>
</tr>
<tr>
<td>The Systematic Sample</td>
<td>208</td>
</tr>
<tr>
<td>8.6 End-of-Chapter Materials</td>
<td>209</td>
</tr>
<tr>
<td>Summary</td>
<td>209</td>
</tr>
<tr>
<td>Key Words</td>
<td>210</td>
</tr>
<tr>
<td>Questions</td>
<td>211</td>
</tr>
<tr>
<td>Problems</td>
<td>211</td>
</tr>
<tr>
<td>Database Exercises</td>
<td>217</td>
</tr>
<tr>
<td>Projects</td>
<td>217</td>
</tr>
<tr>
<td>Case</td>
<td>218</td>
</tr>
</tbody>
</table>

#### 9. Confidence Intervals: Admitting That Estimates Are Not Exact

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.1 The Confidence Interval for a Population Mean or a Population Percentage</td>
<td>220</td>
</tr>
<tr>
<td>The t Table and the t Distribution</td>
<td>222</td>
</tr>
<tr>
<td>The Widely Used 95% Confidence Interval</td>
<td>225</td>
</tr>
<tr>
<td>Other Confidence Levels</td>
<td>228</td>
</tr>
<tr>
<td>9.2 Assumptions Needed for Validity</td>
<td>230</td>
</tr>
<tr>
<td>Random Sampling</td>
<td>230</td>
</tr>
<tr>
<td>Normal Distribution</td>
<td>232</td>
</tr>
<tr>
<td>9.3 Interpreting a Confidence Interval</td>
<td>234</td>
</tr>
<tr>
<td>Which Event Has a 95% Probability?</td>
<td>234</td>
</tr>
<tr>
<td>Your Lifetime Track Record</td>
<td>235</td>
</tr>
<tr>
<td>9.4 One-Sided Confidence Intervals</td>
<td>235</td>
</tr>
<tr>
<td>Be Careful! You Can't Always Use a One-Sided Interval</td>
<td>235</td>
</tr>
<tr>
<td>Computing the One-Sided Interval</td>
<td>235</td>
</tr>
<tr>
<td>9.5 Prediction Intervals</td>
<td>237</td>
</tr>
</tbody>
</table>
10. Hypothesis Testing: Deciding between Reality and Coincidence

10.1 Hypotheses Are Not Created Equal! 250
   The Null Hypothesis 250
   The Research Hypothesis 250
   What Will the Result Tell You? 250
   Examples of Hypotheses 251

10.2 Testing the Population Mean against a Known Reference Value 252
   Using Confidence Intervals: The Easy Way 252
   The t Statistic: Another Way, Same Answer 257

10.3 Interpreting a Hypothesis Test 258
   Errors: Type I and Type II 259
   Assumptions Needed for Validity 259
   Hypotheses Have No Probabilities of Being True or False 260
   Statistical Significance and Test Levels 260
   p-Values 261

10.4 One-Sided Testing 262
   How to Perform the Test 263

10.5 Testing Whether or Not a New Observation Comes from the Same Population 267

10.6 Testing Two Samples 268
   The Paired t Test 268
   The Unpaired t Test 270

10.7 End-of-Chapter Materials 274
   Summary 274
   Key Words 276
   Questions 276
   Problems 277
   Database Exercises 285
   Projects 286
   Case 286

11. Correlation and Regression: Measuring and Predicting Relationships

11.1 Exploring Relationships Using Scatterplots and Correlations 292
   The Scatterplot Shows You the Relationship 292
   Correlation Measures the Strength of the Relationship 295
   The Formula for the Correlation 296
   The Various Types of Relationships 297
   Linear Relationship 297
   No Relationship 299
   Nonlinear Relationship 301
   Unequal Variability 304
   Clustering 307
   Bivariate Outliers 308
   Correlation Is Not Causation 309

11.2 Regression: Prediction of One Variable from Another 310
   A Straight Line Summarizes a Linear Relationship 310
   Straight Lines 312
   Finding a Line Based on Data 312
   How Useful Is the Line? 315
   The Standard Error of Estimate: How Large Are the Prediction Errors? 315
   $R^2$: How Much Is Explained? 317
   Confidence Intervals and Hypothesis Tests for Regression 317
   The Linear Model Assumption Defines the Population 317
   Standard Errors for the Slope and Intercept 318
   Confidence Intervals for Regression Coefficients 319
   Testing Whether the Relationship Is Real or Coincidence 319
   Other Methods of Testing the Significance of a Relationship 320
   Computer Results for the Production Cost Data 320
   Other Tests of a Regression Coefficient 323
   A New Observation: Uncertainty and the Confidence Interval 324
   The Mean of $Y$: Uncertainty and the Confidence Interval 325
   Regression Can Be Misleading 326
   The Linear Model May Be Wrong 326
   Predicting Intervention from Observed Experience Is Difficult 327
   The Intercept May Not Be Meaningful 328
   Explaining Y from X versus Explaining X from Y 328
   A Hidden “Third Factor” May Be Helpful 328

11.3 End-of-Chapter Materials 329
   Summary 329
   Key Words 331
   Questions 331
   Problems 332
   Database Exercises 344
   Projects 345
   Case 345
12. Multiple Regression: Predicting One Variable from Several Others

12.1 Interpreting the Results of a Multiple Regression
- Regression Coefficients and the Regression Equation
- Interpreting the Regression Coefficients
- Predictions and Prediction Errors
- How Good Are the Predictions?
- Typical Prediction Error: Standard Error of Estimate
- Percent Variation Explained: \( R^2 \)
- Inference in Multiple Regression
- Assumptions
- Is the Model Significant? The \( F \) Test or \( R^2 \) Test
- Tables of Critical Values for Testing \( R^2 \)
- Which Variables Are Significant?
- A \( t \) Test for Each Coefficient
- Other Tests for a Regression Coefficient
- Which Variables Explain the Most?

12.2 Pitfalls and Problems in Multiple Regression
- Multicollinearity: Are the Explanatory Variables Too Similar?
- Variable Selection: Are You Using the Wrong Variables?
- Prioritizing the List of \( X \) Variables
- Automating the Variable Selection Process
- Model Misspecification: Does the Regression Equation Have the Wrong Form?
- Exploring the Data to See Nonlinearity or Unequal Variability
- Using the Diagnostic Plot to Decide If You Have a Problem
- Using Percent Changes to Model an Economic Time Series

12.3 Dealing with Nonlinear Relationships and Unequal Variability
- Transforming to a Linear Relationship: Interpreting the Results
- Fitting a Curve with Polynomial Regression
- Modeling Interaction between Two \( X \) Variables

12.4 Indicator Variables: Predicting from Categories
- Interpreting and Testing Regression Coefficients for Indicator Variables
- Separate Regressions

12.5 End-of-Chapter Materials
- Summary
- Key Words

13. Report Writing: Communicating the Results of a Multiple Regression

13.1 How to Organize Your Report
- The Executive Summary Paragraph
- The Introduction Section
- The Analysis and Methods Section
- The Conclusion and Summary Section
- Including References
- The Appendix Section

13.2 Hints and Tips
- Think about Your Audience
- What to Write First? Next? Last?
- Other Sources

13.3 Example: A Quick Pricing Formula for Customer Inquiries

13.4 End-of-Chapter Materials
- Summary
- Key Words
- Questions
- Problems
- Database Exercises
- Project

14. Time Series: Understanding Changes over Time

14.1 An Overview of Time-Series Analysis

14.2 Trend-Seasonal Analysis
- Trend and Cyclic: The Moving Average Seasonal Index: The Average Ratio-to-Moving-Average Indicates Seasonal Behavior
- Seasonal Adjustment: The Series Divided by the Seasonal Index
- Long-Term Trend and Seasonally Adjusted Forecast: The Regression Line Forecast: The Seasonalized Trend

14.3 Modeling Cyclic Behavior
- Using Box–Jenkins ARIMA Processes
- A Random Noise Process Has No Memory: The Starting Point
- An Autoregressive (AR) Process Remembers Where It Was
- A Moving-Average (MA) Process Has a Limited Memory
- The Autoregressive Moving-Average (ARMA) Process Combines AR and MA
Part V
Methods and Applications

15. Anova: Testing for Differences among Many Samples, and Much More
15.1 Using Box Plots to Look at Many Samples at Once 468
15.2 The F Test Tells You If the Averages Are Significantly Different 470
The Data Set and Sources of Variation 470
The Assumptions 470
The Hypotheses 471
The F Statistic 471
The F Table 472
The Result of the F Test 477
Computer Output: The One-Way ANOVA Table 477
15.3 The Least-Significant-Difference Test: Which Pairs Are Different? 479
15.4 More Advanced ANOVA Designs 480
Variety Is the Spice of Life 480
Two-Way ANOVA 480
Three-Way and More 481
Analysis of Covariance (ANCOVA) 481
Multivariate Analysis of Variance (MANOVA) 481
How to Read an ANOVA Table 481
15.5 End-of-Chapter Materials 484
Summary 484
Key Words 485
Questions 485
Problems 485
Database Exercises 489
Projects 490

16. Nonparametrics: Testing with Ordinal Data or Nonnormal Distributions
16.1 Testing the Median against a Known Reference Value 492
The Sign Test 492
The Hypotheses 492
The Assumption 493
16.2 Testing for Differences in Paired Data 495
Using the Sign Test on the Differences 495
The Hypotheses 496
The Assumption 496
16.3 Testing to See If Two Unpaired Samples Are Significantly Different 496
The Procedure Is Based on the Ranks of All of the Data 497
The Hypotheses 497
The Assumptions 497
16.4 End-of-Chapter Materials 500
Summary 500
Key Words 502
Questions 502
Problems 502
Database Exercises 506
Projects 506

17. Chi-Squared Analysis: Testing for Patterns in Qualitative Data
17.1 Summarizing Qualitative Data by Using Counts and Percentages 508
17.2 Testing If Population Percentages Are Equal to Known Reference Values 509
The Chi-Squared Test for Equality of Percentages 509
17.3 Testing for Association between Two Qualitative Variables 512
The Meaning of Independence 512
The Chi-Squared Test for Independence 512
17.4 End-of-Chapter Materials 516
Summary 516
Key Words 517
Questions 517
Problems 518
Database Exercises 521
Projects 521

18. Quality Control: Recognizing and Managing Variation
18.1 Processes and Causes of Variation 524
The Pareto Diagram Shows Where to Focus Attention 525
18.2 Control Charts and How to Read Them 526
The Control Limits Show if a Single Observation Is Out of Control 527
How to Spot Trouble Even within the Control Limits 528
18.3 Charting a Quantitative Measurement with X and R Charts 528
<table>
<thead>
<tr>
<th>18.4 Charting the Percent Defective</th>
<th>Appendix A</th>
<th>Employee Database</th>
</tr>
</thead>
<tbody>
<tr>
<td>18.5 End-of-Chapter Materials</td>
<td>Appendix B</td>
<td>Donations Database</td>
</tr>
<tr>
<td>Summary</td>
<td>Appendix C</td>
<td>Self Test: Solutions to Selected Problems and Database Exercises</td>
</tr>
<tr>
<td>Key Words</td>
<td>Appendix D</td>
<td>Statistical Tables</td>
</tr>
<tr>
<td>Questions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Problems</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Projects</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Glossary</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Index</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>545</td>
</tr>
<tr>
<td></td>
<td></td>
<td>547</td>
</tr>
<tr>
<td></td>
<td></td>
<td>551</td>
</tr>
<tr>
<td></td>
<td></td>
<td>565</td>
</tr>
<tr>
<td></td>
<td></td>
<td>597</td>
</tr>
<tr>
<td></td>
<td></td>
<td>605</td>
</tr>
</tbody>
</table>
Statistical literacy has become a necessity for anyone in business, simply because your competition has already learned how to interpret numbers and how to measure many of the risks involved in this uncertain world. Can you afford to ignore the tons of data now available (to anyone) online when you are searching for a competitive, strategic advantage? We are not born with an intuitive ability to assess randomness or process massive data sets, but fortunately there are fundamental basic principles that let us compute, for example, the risk of a future payoff, the way in which the chances for success change as we continually receive new information, and the best information summaries from a data warehouse. This book will guide you through foundational activities, including how to collect data so that the results are useful, how to explore data to efficiently visualize its basic features, how to use mathematical models to help separate meaningful characteristics from noise, how to determine the quality of your summaries so that you are in a position to make judgments, and how to know when it would be better to ignore the set of data because it is indistinguishable from random noise.

**EXAMPLES**

Examples bring statistics to life, making each topic relevant and useful. There are many real-world examples used throughout *Practical Business Statistics*, chosen from a wide variety of business sources, and many of them of current interest as of 2010 (take a look at the status of Facebook relative to other top websites in Chapter 11). The donations database, which gives characteristics of 20,000 individuals together with the amount that they contributed in response to a mailing, is introduced in Chapter 1 and used in many chapters to illustrate how statistical methods can be used for data mining. The stock market is used in Chapter 5 to illustrate volatility, risk, and diversification as measured by the standard deviation, while the systematic component of market risk is summarized by the regression coefficient in Chapter 11. Because we are all curious about the salaries of others, I have used top executive compensation in several examples and, yes, Enron was an outlier even before the company filed for bankruptcy and the CEO resigned. Quality control is used throughout the book to illustrate individual topics and is also covered in its own chapter (18). Opinion surveys and election polls are used throughout the book (and especially in Chapter 9) because they represent a very pure kind of real-life statistical inference that we are all familiar with and use frequently in business. Using the Internet to locate data is featured in Chapter 2. Prices of magazine advertisements are used in Chapter 12 to show how multiple regression can uncover relationships in complex data sets, and we learn the value of a larger audience with a higher income simply by crunching the numbers. Microsoft’s revenues and U.S. unemployment rates are used in Chapter 14 to demonstrate what goes on behind the scenes in time-series forecasting. Students learn better through the use of motivating examples and applications. All numerical examples are included in the Excel® files on the companion website, with ranges named appropriately for easy analysis.

**STATISTICAL GRAPHICS**

To help show what is going on in the data sets, *Practical Business Statistics* includes over 200 figures to illustrate important features and relationships. The graphs are exact because they were initially drawn with the help of a computer. For example, the bell-shaped normal curves here are accurate, unlike those in many books, which are distorted because they appear to be an artist’s enhancement of a casual, hand-drawn sketch. There is no substitute for accuracy!

**EXTENSIVE DEVELOPMENT: REVIEWS AND CLASS TESTING**

This book began as a collection of readings I handed out to my students as a supplement to the assigned textbook. All of the available books seemed to make statistics seem unnecessarily difficult, and I wanted to develop and present straightforward ways to think about the subject. I also wanted to add more of a real-world business flavor to the topic. All of the helpful feedback I have received from students over the years has been acted upon and has improved the book. *Practical Business Statistics* has been through several stages of reviewing and classroom testing. Now that five editions have been used in colleges and universities across the country and around the world, preparing
the sixth edition has given me the chance to fine-tune the book, based on the additional reviews and all the helpful, encouraging comments that I have received.

**WRITING STYLE**

I enjoy writing. I have presented the “inside scoop” wherever possible, explaining how we statisticians really think about a topic, what it implies, and how it is useful. This approach helps bring some sorely needed life to a subject that unfortunately suffers from dreadful public relations. Of course, the traditional explanations are also given here so that you can see it both ways: here is what we say, and here is what it means, all the while maintaining technical rigor.

It thrilled me to hear even some of my more quantitative-phobic students tell me that the text is actually enjoyable to read! And this was after the final grades were in!

**CASES**

To show how statistical thinking can be useful as an integrated part of a larger business activity, cases are included at the end of each of Chapters 3–12. These cases provide extended and open-ended situations as an opportunity for thought and discussion, often with no single correct answer.

**ORGANIZATION**

The reader should always know why the current material is important. For this reason, each part begins with a brief look at the subject of that part and the chapters to come. Each chapter begins with an overview of its topic, showing why the subject is important to business, before proceeding to the details and examples.

Key words, the most important terms and phrases, are presented in bold in the sentence where they are defined. They are collected in the Key Words list at the end of each chapter and also included in the glossary at the back of the book (hint! this could be very useful!). This makes it easy to study by focusing attention on the main ideas. An extensive index helps you find main topics as well as small details. Try looking up “examples,” “correlation,” “unpaired t test,” or even “mortgage.”

Extensive end-of-chapter materials are included, beginning with a summary of the important material covered. Next is the list of key words. The questions provide a review of the main topics, indicating why they are important. The problems give the student a chance to apply statistics to new situations. The database exercises (included in most chapters) give further practice problems based on the employee database in Appendix A. The projects bring statistics closer to the students’ needs and interests by allowing them to help define the problem and choose the data set from their work experience or interests from sources including the Internet, current publications, or their company. Finally, the cases (one each for Chapters 3–12) provide extended and open-ended situations as an opportunity for thought and discussion, often with no single correct answer.

Several special topics are covered in addition to the foundations of statistics and their applications to business. Data mining is introduced in Chapter 1 and carried throughout the book. Because communication is so important in the business world, Chapter 13 shows how to gather and present statistical material in a report. Chapter 14 includes an intuitive discussion of the Box–Jenkins forecasting approach to time series using ARIMA models. Chapter 18 shows how statistical methods can help you achieve and improve quality; discussion of quality control techniques is also interspersed throughout the text.

**Practical Business Statistics** is organized into five parts, plus appendices, as follows:

- **Part I, Chapters 1 through 5**, is “Introduction and Descriptive Statistics.” Chapter 1 motivates by showing how the use of statistics provides a competitive edge in business and then outlines the basic activities of statistics and offers varied examples including data mining with large databases. Chapter 2 surveys the various types of data sets (quantitative, qualitative, ordinal, nominal, bivariate, time-series, etc.), the distinction between primary and secondary data, and use of the Internet. Chapter 3 shows how the histogram lets you see what’s in the data set, which would otherwise be difficult to determine just from staring at a list of numbers. Chapter 4 covers the basic landmark summaries, including the average, median, mode, and percentiles, which are displayed in the box plot and the cumulative distribution function. Chapter 5 discusses variability, which often translates to risk in business terms, featuring the standard deviation as well as the range and coefficient of variation.

- **Part II, including Chapters 6 and 7**, is “Probability.” Chapter 6 covers probabilities of events and their combinations, using probability trees both as a way of visualizing the situation and as an efficient method for computing probabilities. Conditional probabilities are interpreted as a way of making the best use of the information you have. Chapter 7 covers random variables (numerical outcomes), which often represent those numbers that are important to your business but are not yet available. Details are provided concerning general discrete distributions, the binomial distribution, the normal distribution, the Poisson distribution, and the exponential distribution.

- **Part III, Chapters 8 through 10**, is “Statistical Inference.” These chapters pull together the descriptive summaries of Part I and the formal probability assessments of Part II, allowing you to reach probability conclusions
Part V, Chapters 15 through 18, is “Methods and Applications,” a grab bag of optional, special topics that extend the basic material covered so far. Chapter 15 shows how the analysis of variance allows you to use hypothesis testing in more complex situations, especially involving categories along with numeric data. Chapter 16 covers nonparametric methods, which can be used when the basic assumptions for statistical inference are not satisfied, that is, for cases where the distributions might not be normal or the data set might be merely ordinal. Chapter 17 shows how chi-squared analysis can be used to test relationships among the categories of nominal data. Finally, Chapter 18 shows how quality control relies heavily on statistical methods such as Pareto diagrams and control charts.

- Appendix A is the “Employee Database,” consisting of information on salary, experience, age, gender, and training level for a number of administrative employees. This data set is used in the database exercises section at the end of most chapters. Appendix B describes the donations database on the companion website (giving characteristics of 20,000 individuals together with the amount that they contributed in response to a mailing) that is introduced in Chapter 1 and used in many chapters to illustrate how statistical methods can be used for data mining. Appendix C gives detailed solutions to selected parts of problems and database exercises (marked with an asterisk in the text). Appendix D collects all of the statistical tables used throughout the text.

POWERPOINT SLIDES
A complete set of PowerPoint slides, that I developed for my own classes, is available on the companion website.

EXCEL® GUIDE
The Excel® Guide, prepared by me (and I have enjoyed spreadsheet computing since its early days) provides examples of statistical analysis using Excel® using data taken chapter-by-chapter from Practical Business Statistics. It’s a convenient way for students to learn how to use computers if your class is using Excel®.

COMPANION WEBSITE
The companion website http://www.elsevierdirect.com includes the PowerPoint presentation slides, the Excel Guide, and Excel files with all quantitative examples and problem data.

INSTRUCTOR’S MANUAL
The instructor’s manual is designed to help save time in preparing lectures. A brief discussion of teaching objectives and how to motivate students is provided for each chapter. Also included are detailed solutions to questions, problems, and database exercises, as well as analysis and discussion material for each case. The instructor’s manual is available at the companion website.

ACKNOWLEDGMENTS
Many thanks to all of the reviewers and students who have read and commented on drafts and previous editions of Practical Business Statistics over the years. I have been
lucky to have dedicated, careful readers at a variety of institutions who were not afraid to say what it would take to meet their needs.

I am fortunate to have been able to work with my parents, Mildred and Armand Siegel, who provided many careful and detailed suggestions for the text.

Very special thanks go to Lauren Schultz Yuhasz, Lisa Lamenzo, Gavin Becker, and Jeff Freeland, who have been very helpful and encouraging with the development and production of this edition. Warm thanks go to Michael Antonucci, who started this whole thing when he stopped by my office to talk about computers and see what I was up to and encourage me to write it all down. I am also grateful to those who were involved with previous editions, including Scott Isenberg, Christina Sanders, Catherine Schultz, Richard T. Hercher, Carol Rose, Gail Korosa, Ann Granacki, Colleen Tuscher, Adam Rooke, Ted Tsukahara, and Margaret Haywood. It’s a big job producing a work like this, and I was lucky to have people with so much knowledge, dedication, and organizational skill.


A special mention is given to a distinguished group of colleagues who have provided helpful guidance, including Bruce Barrett, University of Alabama; Brian Goff, Western Kentucky University; Anthony Seraphin, University of Delaware; Abbott Packard, Hawkeye Community College; William Seaver, University of Tennessee–Knoxville; Nicholas Jewell, University of California–Berkeley; Howard Clayton, Auburn University; Giorgio Canarrello, California State University–Los Angeles; Lyle Brenner, University of Florida–Gainesville; P. S. Sundararaghavan, University of Toledo; Julien Bramel, Columbia University, Ronald Bremer, Texas Tech University; Stergios Fotopoulos, Washington State University; Michael Ghanen, Webster University; Phillip Musa, Texas Tech University; Thomas Obremski, University of Denver; Darrell Radson, University of Wisconsin, Milwaukee; Terrence Reilly, Babson College; Peter Schuhmann, University of Richmond; Bala Shetty, Texas A&M University; L. Dwight Sneathen Jr., University of Arizona; Ted Tsukahara, St. Mary’s College; Edward A. Wasil, American University; Michael Wegmann, Keller Graduate School of Management; Mustafa Yilmaz, Northeastern University; Gary Yoshimoto, St. Cloud State University; Sangit Chatterjee, Northeastern University; Jay Devore, California Polytechnic State University; Burt Holland, Temple University; Winston Lin, State University of New York at Buffalo; Herbert Spier, University of Connecticut; Donald Westerfield; Webster University; Wayne Winston, Indiana University; Jack Yurkiewicz, Pace University; Betty Thorne, Stetson University; Dennis Petruska, Youngstown State University; H. Karim, West Coast University; Martin Young, University of Michigan; Richard Spinetto, University of Colorado at Boulder; Paul Paschke, Oregon State University; Larry Ammann, University of Texas at Dallas; Donald Marx, University of Alaska; Kevin Ng, University of Ottawa; Rahmat Tavallali, Walsh University; David Auer, Western Washington University; Murray Cote, Texas A&M University; Peter Lakner, New York University; Donald Adolphson, Brigham Young University; and A. Rahulji Parsa, Drake University.

TO THE STUDENT

As you begin this course, you may have some preconceived notions of what statistics is all about. If you have positive notions, please keep them and share them with your classmates. But if you have negative notions, please set them aside and remain open-minded until you’ve given statistics another chance to prove its value in analyzing business risk and providing insight into piles of numbers.

In some ways, statistics is easier for your generation than for those of the past. Now that computers can do the messy numerical work, you are free to develop a deeper understanding of the concepts and how they can help you compete over the course of your business career.

Make good use of the introductory material so that you will always know why statistics is worth the effort. Focus on examples to help with understanding and motivation. Take advantage of the summary, key words, and other materials at the ends of the chapters. Don’t forget about the detailed problem solutions and the glossary at the back when you need a quick reminder! And don’t worry. Once you realize how much statistics can help you in business, the things you need to learn will fall into place much more easily.

Why not keep this book as a reference? You’ll be glad you did when the boss needs you to draft a memo immediately that requires a quick look at some data or a response to an adversary’s analysis. With the help of Practical Business Statistics on your bookshelf, you’ll be able to finish early and still go out to dinner. Bon appétit!

ANDREW F. SIEGEL
Andrew F. Siegel is Professor, Departments of ISOM (Information Systems and Operations Management) and Finance, at the Michael G. Foster School of Business, University of Washington, Seattle. He is also Adjunct Professor in the Department of Statistics. He has a Ph.D. in statistics from Stanford University (1977), an M.S. in mathematics from Stanford University (1975), and a B.A. in mathematics and physics summa cum laude with distinction from Boston University (1973). Before settling in Seattle, he held teaching and/or research positions at Harvard University, the University of Wisconsin, the RAND Corporation, the Smithsonian Institution, and Princeton University. He has also been a visiting professor at the University of Burgundy at Dijon, France, at the Sorbonne in Paris, and at HEC Business School near Paris. The very first time he taught statistics in a business school (University of Washington, 1983) he was granted the Professor of the Quarter award by the MBA students. He was named the Grant I. Butterbaugh Professor beginning in 1993; this endowed professorship was created by a highly successful executive in honor of Professor Butterbaugh, a business statistics teacher. (Students: Perhaps you will feel this way about your teacher 20 years from now.) Other honors and awards include Burlington Northern Foundation Faculty Achievement Awards, 1986 and 1992; Research Associate, Center for the Study of Futures Markets, Columbia University, 1988; Excellence in Teaching Awards, Executive MBA Program, University of Washington, 1986 and 1988; Research Opportunities in Auditing Award, Peat Marwick Foundation, 1987; and Phi Beta Kappa, 1973.