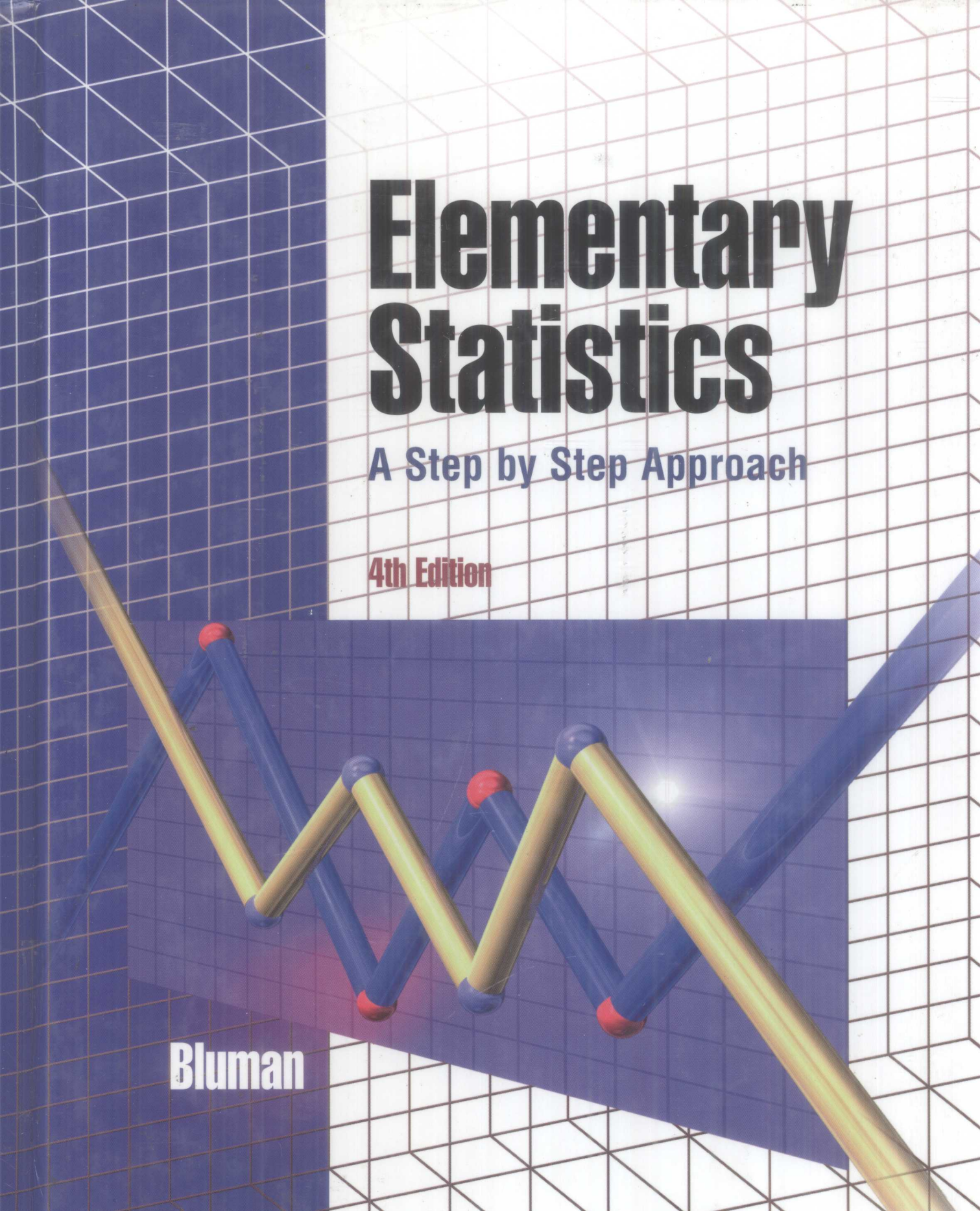# Elementary Statistics

A Step by Step Approach

4th Edition

Bluman

Elementary Statistics

A Step by Step Approach

Fourth Edition

Allan G. Bluman
Community College of Allegheny County

McGraw-Hill Higher Education

*A Division of The **McGraw-Hill** Companies*

Preface

Approach

Elementary Statistics: A Step by Step Approach is a textbook for students in the beginning statistics course whose mathematical background is limited to basic algebra. The book uses a nontheoretical approach in which concepts are explained intuitively and supported by examples. There are no formal proofs in the book. The applications are general in nature, and the exercises include problems from business, economics, health, medicine, science, engineering, social science, education, and general interest.

About This Book

The learning system found in *Elementary Statistics* provides the student with a valuable framework in which to learn and apply concepts.

- Each chapter begins with an outline and a list of **learning objectives.** The objectives are repeated at the beginning of each section to help students focus on the concepts presented within that section.

2–3

Histograms, Frequency Polygons, and Ogives

Objective 2. Represent data in frequency distributions using histograms, frequency polygons, and ogives.

After the data have been organized into a frequency distribution, they can be presented in graphic forms. The purpose of graphs in statistics is to convey the data to the viewer in pictorial form. It is easier for most people to comprehend the meaning of data presented graphically than data presented numerically in tables or frequency distributions. This is especially true if they have little or no statistical knowledge.

Statistical graphs can be used to describe the data set or analyze it. Graphs are also useful in getting the audience's attention in a publication or a speaking presentation. They can be used to discuss an issue, reinforce a critical point, or summarize a data set. They can also be used to discover a trend or pattern in a situation over a period of time.

- The outline and learning objectives are followed by a feature titled **Statistics Today,** in which a real-life problem shows students the relevance of the material in the chapter. This problem is subsequently solved near the end of the chapter using the statistical techniques that were presented in the chapter.

Statistics Today

Why Are We Running Out of 800 Numbers?

Phone companies and other agencies that deal in numbers need to know how many phone numbers, ID tags, or license plates they can issue using certain combinations of various letters and numbers. The article shown below explains that the phone companies are running out of toll-free 800 numbers. The question is: How many phone numbers with the 800 prefix can be issued in the United States?

Toll-free call? Get ready to dial 888

By Becky Beyers
USA TODAY

Get ready to keep your finger on the 8 when you make a toll-free call.

Phone companies will run out of 800 numbers early next year and start issuing toll-free numbers beginning with 888.

Use of 800 numbers has grown so fast, "we're a victim of our own success," says Dennis Byrne of the U.S. Telephone Association.

Only about 1.7 million of the 7.6 million possible 800-prefix combinations are still available. Why so few are left:

▶ Demand has taken off since May 1993, when the government allowed users to keep 800 numbers if they changed long-distance carriers.

▶ 800 numbers aren't just for big companies anymore. Small businesses use them, as do residential customers so family members can call home more cheaply than collect.

Such customers may pay 25 cents a minute for each call plus a monthly fee of $5.

▶ Some toll-free numbers are hoarded for promotional value or occasional use.

The industry's numbering council — phone companies and associations that set phone-number policies — is asking that little-used numbers be returned so they can be reissued.

Setting up a new toll-free access code involves the entire phone industry, Byrne says.

All internal systems must upgrade switching equipment so they can handle 888 calls.

What happens when the 888s are used up? It's on to 877, 866, and all the way down to 822.

Source: *USA Today,* February 13, 1995. Copyright 1995 *USA TODAY.* Used with permission.

In this chapter, you will learn the rule for counting, the differences between permutations and combinations, and how to figure out how many different combinations for specific situations exist.

- Over 300 **examples** with detailed solutions are provided to help students learn to solve problems. Examples are solved by using a step by step explanation. Illustrations provide a clear display of results for students.

Example 3–23

Find the variance and standard deviation for the amount of European auto sales for a sample of six years shown. The data are in millions of dollars.

11.2, 11.9, 12.0, 12.8, 13.4, 14.3

Source: *USA Today,* March 22, 1999.

Solution

STEP 1 Find the sum of the values.

$$\Sigma X = 11.2 + 11.9 + 12.0 + 12.8 + 13.4 + 14.3 = 75.6$$

STEP 2 Square each value and find the sum.

$$\Sigma X^2 = 11.2^2 + 11.9^2 + 12.0^2 + 12.8^2 + 13.4^2 + 14.3^2 = 958.94$$

STEP 3 Substitute in the formulas and solve.

$$s^2 = \frac{\Sigma X^2 - [(\Sigma X)^2/n]}{n-1} = \frac{958.94 - [(75.6^2)/6]}{5}$$
$$s^2 = 1.28$$

The variance of the sample is 1.28

$$s = \sqrt{1.28} = 1.13$$

Hence, the sample standard deviation is 1.13.

- Numerous examples and exercises use **real data.**

36. The following data represent the attendance at seven Pittsburgh museums for 1997 and 1998. Draw two boxplots for the data and compare the distributions. The data are in thousands.

1997	1998
101	107
754	764
291	293
83	110
589	627
152	103
158	154

Source: *The Pittsburgh Tribune Review,* February 21, 1999.

- Numerous **Procedure Tables** summarize processes for the student. All use the step by step method.

Procedure Table

Finding the Sample Variance and Standard Deviation for Grouped Data

STEP 1 Make a table as shown, and find the midpoint of each class.

A	B	C	D	E
Class	Frequency	Midpoint	$f \cdot X_m$	$f \cdot X_m^2$

STEP 2 Multiply the frequency by the midpoint for each class, and place the products in column D.

STEP 3 Multiply the frequency by the square of the midpoint, and place the products in column E.

STEP 4 Find the sums of columns B, D, and E. (The sum of column B is n. The sum of column D is $\Sigma f \cdot X_m$. The sum of column E is $\Sigma f \cdot X_m^2$.)

STEP 5 Substitute in the formula and solve to get the variance.

$$s^2 = \frac{\Sigma f \cdot X_m^2 - [(\Sigma f \cdot X_m)^2/n]}{n - 1}$$

STEP 6 Take the square root to get the standard deviation.

- The **Speaking of Statistics** sections invite students to think about poll results and other statistics-related news stories.

Speaking of **STATISTICS**

Here is a study on the snacks consumed during the 1998 Super Bowl. Suggest several ways that the data might have been obtained.

Super snack stats

■ **Snack fests:** Super Sunday, New Year's Eve and July Fourth are America's top snacking days.

■ **Eat up:** When it comes to the total amount of food eaten per person in a day, Super Sunday is No. 2, behind only Thanksgiving.

■ **Party on:** The Super Bowl is the No. 1 at-home party event of the year, says Hallmark Cards; the average party has 17 guests.

■ **The morning after:** Antacid sales increase 20 percent the day after the Super Bowl, reports the convenience store chain 7-Eleven.

SUPER BOWL FAVORITES

Snack	Pounds eaten during '98 game
Potato chips	11.2 million
Tortilla chips	8.2 million
Pretzels	4.3 million
Popcorn	3.8 million
Snack nuts	2.5 million

Sources: Snack Food Association, National Potato Promotion Board, Super Bowl 1998
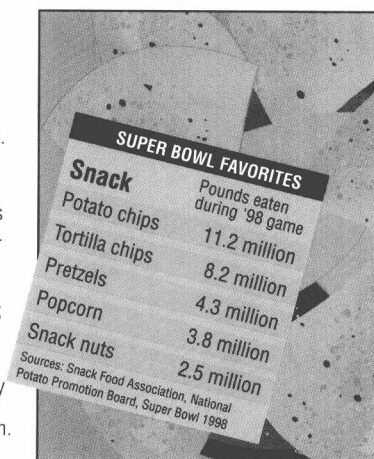
Source: *USA WEEKEND*, January 22–24, 1999. Used with permission.

- **Historical Notes, Unusual Stats,** and **Interesting Facts,** located in the margins, make statistics come alive for the reader.

Unusual Stats

According to the *Statistical Abstract of the United States,* 52% of Americans live within 50 miles of a coastal shoreline.

- **Rules and definitions** are set off for easy referencing by the student.

Objective 2. Find the total number of outcomes in a sequence of events using the multiplication rule.

Multiplication Rule

In a sequence of n events in which the first one has k_1 possibilities and the second event has k_2 and the third has k_3, and so forth, the total number of possibilities of the sequence will be

$$k_1 \cdot k_2 \cdot k_3 \cdot \cdots \cdot k_n$$

Note: "And" in this case means to multiply.

- Over 1,600 **exercises** are located at the end of major sections within each chapter.
- At the end of appropriate sections, **Technology Step by Step** boxes show students how to use MINITAB, the TI-83 graphing calculator, and Excel to solve the types of problems covered in the section. Instructions are presented in numbered steps, usually in the context of examples—including examples from the main part of the section. Numerous computer or calculator screens are displayed, showing intermediate steps as well as the final answer. This feature, **new to the fourth edition,** will be valuable to students using any of these tools.

Technology Step by Step

**MINITAB
Step by Step**

Finding the Mean and Standard Deviation

Example MT3–1

1. Type the data from Example 3–39 (in the following section) into C1 of MINITAB. Name the column **CARS-THEFT.**

> 52 58 75 79 57 65 62 77 56 59 51 53 51 66 55
> 68 63 78 50 53 67 65 69 66 69 57 73 72 75 55

2. Select **Stat>Basic Statistics>Display Descriptive Statistics.**
3. The cursor will be blinking in the Variables text box. Double-click C1.
4. Click [OK].

The results will be displayed in the Session Window as shown. The column label "CARS-THEFT" is truncated to 8 letters in the display. The standard deviation is the unbiased estimate, s. The trimmed mean or TrMean is the mean for the data after the lowest and highest 5% are discarded. If the trimmed mean is different from the mean, there may be outliers.

Session Window with Descriptive Statistics

Session

Descriptive Statistics: CARS-THEFT

Variable	N	Mean	Median	TrMean	StDev	SE Mean
CARS-THE	30	63.20	64.00	63.00	9.01	1.64

Variable	Minimum	Maximum	Q1	Q3
CARS-THE	50.00	79.00	55.00	69.75

- **Critical Thinking** sections at the end of each chapter challenge the students to apply what they have learned to new situations. The problems presented are designed to deepen conceptual understanding and/or to extend topical coverage.

Critical Thinking Challenges

1. A person decides to shake hands with six different people on a certain day. The next day, each of the six people will shake hands with six different people. The process continues until every person in the United States has shaken someone's hand. How many days will it take until everyone in the United States has shaken hands once? Assume that once a person shakes hands with six different people, he or she does not shake hands again. (*Hint:* The population of the United States is 248,709,873, according to the 1990 census.)

2. If it can be assumed that the maximum number of hairs on a human head is about 500,000, explain why at least two people living in Houston (population 1,629,902,

triangle is f... ...dding the two adja... ...mb
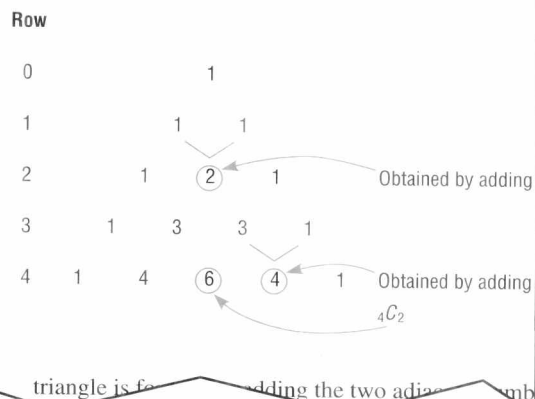
- **Data projects** further challenge students' understanding and application of the material presented in the chapter. Many of these require the student to gather, analyze, and report on real data. These projects, which appear at the end of each chapter, may include a World Wide Web icon [icon], indicating that websites are listed as possible sources of data.

Data Projects

Where appropriate, use MINITAB, the TI-83, Excel, or a computer program of your choice to complete the following exercises.

1. Select a categorical (nominal) variable, such as the colors of cars in the school's parking lot or the major fields of the students in statistics class, and collect data on this variable.
   a. State the purpose of the project.
   b. Define the population.
   c. State how the sample was selected.
   d. Show the raw data.
   e. Construct a frequency distribution for the variable.
   f. Draw some appropriate graphs (pie, Pareto, etc.) for the data.
   g. Analyze the results.

2. Using an almanac, select a variable that varies over a period of several years (e.g., silver production) and draw a time series graph for the data. Write a short paragraph interpreting the findings.

3. Select a variable (interval or ratio) and collect at least 30 values. For example, you may ask the students in your class how many hours they study per week or how old they are, etc..
   a. State the purpose of the project.
   b. Define the population.
   c. State how the sample was selected.
   d. Show the raw data.
   e. Construct a frequency distribution for the data.
   f. Draw a histogram, frequency polygon, and ogive for the data.
   g. Analyze the results.

You may use the following websites to obtain raw data:

http://www.mhhe.com/math/stat/bluman/
http://lib.stat.cmu.edu/DASL
http://www.oecd.org/statlist.htm
http://www.statcan.ca/english/

- **Hypothesis-Testing Summaries** are found at the end of Chapter 10 (z, t, χ^2, and F tests for testing means, proportions, and variances), Chapter 13 (correlation, chi-square, and ANOVA), and Chapter 14 (nonparametric tests) to show students the different types of hypotheses and the types of tests to use.

- A **Data Bank** listing various attributes (educational level, cholesterol level, gender, etc.) for 100 people is included and referenced in various exercises and projects throughout the book, including the projects presented in Data Projects sections.

- A **reference card** containing the formulas and the z, t, χ^2, and PPMC tables is included with this textbook.

- End-of-chapter **Summaries, Important Terms,** and **Important Formulas** give students a concise summary of the chapter topics and provide a good source for quiz or test preparation.

- **Review Exercises** are found at the end of each chapter.

- Special sections called **Data Analysis** require students to work with a data set to perform various statistical tests or procedures and then summarize the results. The data is included in the Data Bank following Appendix C and can be downloaded from the book's website at www.mhhe.com/math/stat/bluman.

- **Chapter quizzes,** found at the end of each chapter, include multiple choice, true–false, and completion questions along with exercises to test students' knowledge and comprehension of chapter content.

Content Changes for the Fourth Edition

To build on the success of the third edition, the content and format have been maintained in the fourth edition while changes based on suggestions of reviewers and the current thinking of those who are knowledgeable in the realm of statistical education have been integrated.

With this in mind, the major goals for this revision are to enable students to

1. Draw conclusions from data
2. Become more statistically literate
3. Have a better understanding of the "logic" of hypothesis testing
4. Know how to use the *P*-value method for hypothesis testing
5. Work with raw data as well as summary data

To achieve goal 1, questions based on interpreting the computational solutions for exercises have been included throughout the book. In addition, new problems that involve the comparisons of two data sets by frequency distributions, graphs, and summary statistics have been added. Examples and exercises comparing data sets using back-to-back stem and leaf plots and comparing data sets using two boxplots have been added. Also, students are asked to describe the nature of distributions with regard to their shape, spread, etc.

To achieve goal 2, topics in statistical literacy have been incorporated throughout the textbook where appropriate. These topics include

1. The nature of observational and experimental studies in Chapter 1
2. More material on misleading graphs in Chapter 2
3. Probability and risk-taking in Chapter 5
4. Expectation and gambling in Chapter 6
5. A new section on sampling techniques and questionnaire design in Chapter 15

To achieve goal 3, material has been added at the beginning of Chapter 9 on the logic of hypothesis testing. This information includes comparing two distributions, one when the null hypothesis is true and the other when the null hypothesis is false. Also, material has been added comparing the hypothesis testing situation to a jury trial.

To achieve goal 4, explanations, examples, and exercises on the *P*-value method of hypothesis testing have been included in Chapters 9 through 14. A five-step procedure for testing hypotheses using *P*-values has been added to Chapter 9. Finding intervals from tables for *P*-values have been shown for the t, χ^2, and F tests. To help students with this concept, the actual *P*-values obtained from a calculator have also been given with the interval values.

To achieve goal 5, examples and exercises using real data have been included in Chapters 2, 3, 8, 9, 10, 11, 12, 13, and 14.

Other changes have been made in the following chapters:

- Chapter 2: The purpose of relative frequency graphs has been explained. Note: Those wishing to teach scatter plots with the other graphs can teach Section 11–2 with this chapter. Coverage of ungrouped frequency distributions was streamlined.
- Chapter 3: Back-to-back stem and leaf plots have been added for the comparison of data sets. Quartiles instead of hinges are used in boxplots since most calculators and

computer programs and other textbooks use quartiles. An easy way to compute quartiles has been shown. Drawings for quartiles, deciles, and percentiles have been added. An explanation of how to compare boxplots has been added.

- Chapter 4: This chapter has been shortened considerably. There is only one multiplication rule, one permutation rule, and one combination rule. In addition, sections have been combined.
- Chapter 5: Bayes's theorem has been moved to an Appendix.
- Chapter 6: The title of this chapter has been changed from Probability Distributions to Discrete Probability Distributions.
- Chapter 7: The section on the finite population correction factor has been de-emphasized. The section on normal approximation to the binomial distribution has been rewritten and a procedure table has been added to clarify the concept.
- Chapter 8: The definition of *confidence level* now comes before the definition of *confidence interval.*
- Chapter 9: In addition to the changes mentioned previously, the formula

$$z = \frac{\hat{p} - p}{\sqrt{pq/n}} \quad \text{instead of} \quad z = \frac{X - \mu}{\sigma}$$

is used to test hypotheses for a proportion. Some of the section titles have been changed to more accurately describe the material contained in the sections. The material on the power of a test has been rewritten and simplified.
- Chapter 10: The formula for finding the confidence interval for the difference of two proportions has been changed.
- Chapter 11: As previously stated, Section 11–2 on drawing and analyzing scatter plots can be taught with the other graphs in Chapter 2. A subsection on prediction intervals using the *t* distribution has been added.
- Chapters 12 through 14: Examples and exercises for hypothesis testing using *P*-values and raw data values have been added.
- Chapter 15: A new section on surveys and questionnaire design has been added.
- Chapter 16, on quality control, has been removed.

Altogether, the pedagogical improvements discussed above have resulted in the following changes to the Fourth Edition:

5 new "Speaking of Statistics" boxes
Over 250 new or modified Exercises
17 new Critical Thinking Challenges
Many new Examples
24 additional Data Analysis Problems

Finally, the technology material (MINITAB, TI-83, and Excel) is covered in new "Technology Step by Step" subsections at the end of the appropriate chapter sections.

Supplements

The text is accompanied by an extensive set of supplements for use by you and your students, all of which are carefully coordinated with the text.

Website

The **website** for *Elementary Statistics: A Step by Step Approach,* Fourth Edition, provides the data sets used in examples and exercises in a variety of formats including

- MINITAB
- TI Graph Link files for TI-83
- Excel (for Windows and Macintosh)
- SPSS
- Comma-Delimited ASCII

This can save the student using a computer or calculator from having to enter data by hand, which takes up valuable time and increases the chances of error. The "Data Bank" statistics that are used in the Data Analysis assignments at the ends of chapters are provided for download as well.

The website also provides links to data sources referred to in many of the Data Projects and links to dozens of other statistics-related sites. A PowerPoint presentation also found on the website is available for use in lectures or as a student study aid.

The website address is www.mhhe.com/math/stat/bluman.

For the Instructor

- *Instructor's Solutions,* by Sally Robinson of South Plains College. This manual includes worked-out solutions to most of the exercises in the text.
- *Critical Thinking Workbook: Instructor's Edition,* contains solutions to the students' version of the *Critical Thinking Workbook* described below.
- The *Test Bank* contains a variety of questions, including true–false, multiple-choice, short answer, and short problems requiring analysis and written answers. The testing material is coded by type of question and level of difficulty.
- The computerized test bank enables you to efficiently select, add, and organize questions, such as by type of question or level of difficulty. It also allows for printing tests along with answer keys, as well as editing the original questions. The computerized test bank is available for Windows and Mac systems.
- Full-color lecture slides in PowerPoint format highlight chapter concepts, summarize main points, and illustrate examples. These files can be downloaded from the book's website at www.mhhe.com/math/stat/bluman. PowerPoint users can customize the slides to suit the specific needs of their course.
- *Against All Odds* and *Decisions through Data* are video series available to qualified adopters. Please contact your local sales representative for more information about these programs.

For the Student

- *Critical Thinking Workbook,* by James Condor of Manatee Community College, provides a number of additional challenging problems for students to solve that are drawn from real-world applications. Problems are keyed to each chapter and are designed to highlight and emphasize key concepts.
- *Student Study Guide* by Pat Foard of South Plains College will assist students in understanding and reviewing key concepts and preparing for exams. It emphasizes all important concepts contained in each chapter, includes explanations, and provides opportunities for students to test their understanding by completing related exercises and problems.
- *Student Solutions Manual,* by Sally Robinson of South Plains College, contains detailed solutions to all odd-numbered text problems.
- *MINITAB—Student Version.* This software and user manual provides the student with how-to information on data and file management, conducting various statistical analyses, and creating presentation-style graphics.

Acknowledgments

I would like to thank the following people and companies for granting permission to reprint their statistical tables and other material:

Addison-Wesley Publishing Company, Inc.
Benjamin/Cummings Publishing Company
CRC Press, Inc.
Institute of Mathematical Statistics
Prentice Hall, Inc.
Texas Instruments
Consumers Union (Copyright 1999 by Consumers Union of U.S., Yonkers, NY 10703-1057. Each data set used by permission from *Consumer Reports*. To subscribe, call 800-234-1645; www.ConsumerReports.org.)

I am also grateful to the many authors and publishers who granted me permission to use their articles and cartoons. I would like to acknowledge the cooperation of MINITAB, Inc., in the preparation of this textbook.

The most important and useful advice about *Elementary Statistics: A Step by Step Approach* comes from users of the book and other statistics instructors. I am grateful to the following reviewers for their recommendations.

William A. Ahroon, *Plattsburgh State University*
Anne G. Albert, *University of Findlay*
Abraham K. Biggs, *Broward Community College*
Cecil J. Coone, *State Technical Institute at Memphis*
Callie Harmon Daniels, *St. Charles County Community College*
Bill Dunn, *Las Positas College*
David J. French, *Tidewater Community College*
Kathleen Fritsch, *University of Tennessee at Martin*
James R. Fryxell, *College of Lake County*
Ashis K. Gangopadhyay, *Boston University*
Mark E. Glickman, *Boston University*
David R. Gurney, *Southeastern Louisiana University*
Barney Herron, *Muskegon Community College*
Robert L. Horvath, *El Camino College*
Rebecca M. Howard, *Roane State Community College*
Jane Keller, *Metropolitan Community College, Omaha*
Michael J. Keller, *St. Johns River Community College*
Michael Kelly, *State Technical Institute at Memphis*
Rhonda Magel, *North Dakota State University*
Donald K. Mason, *Elmhurst College*
Jeff Mock, *Diablo Valley College*
Carla Monticelli, *Camden County College*
Gerry Moultine, *Northwood University*
Sharon R. Neidert, *University of Tennessee*
Neal Rogness, *Grand Valley State University*
Susan C. Schott, *University of Central Florida*
Larry Snyder, *Ohio University*

David Stewart, *Community College of Baltimore County*

Richard H. Stockbridge, *University of Kentucky*

Donald B. White, *University of Toledo*

Laurie Sawyer Woodman, *University of New England*

I would also like to thank the more than 60 reviewers of the third edition, whose suggestions and insights have been a positive influence on every page of this book. They are:

Dan Abbey, *Broward Community College*

Randall Allbritton, *Daytona Beach Community College*

Michael S. Allen, *Glendale Community College*

Mostafa S. Aminzadeh, *Towson State University*

Raymond Badalian, *Los Angeles City College*

Carole Bernett, *William Rainey Harper College*

Rich Campbell, *Butte College*

Mark Carpenter, *Sam Houston State University*

Daniel Cherwien, *Cumberland County College*

James A. Condor, *Manatee Community College*

David T. Cooney, *Polk Community College*

James C. Curl, *Modesto Junior College*

Carol Curtis, *Fresno City College*

Steven Day, *Riverside Community College*

Nirmal Devi, *Embry-Riddle Aeronautical University*

Wayne Ehler, *Anne Arundel College*

Eugene Enneking, *Portland State University*

Michael Eurgubian, *Santa Rosa Junior College*

Ruby Evans, *Santa Fe Community College*

Jeff Gervasi, *Porterville College*

Dawit Getahew, *Chicago State University*

Gary Grimes, *Mt. Hood Community College*

Leslie Grunes, *Mercer County Community College*

Dianne Haber, *Westfield State College*

Ronald Hamill, *Community College of Rhode Island*

Mark Harbison, *Rio Hondo College*

Linda Harper, *Harrisburg Area Community College*

Susan Herring, *Sonoma State University*

Keith A. Hilmer, *Moorpark College*

Shu-ping Hodgson, *Central Michigan University*

Robert L. Horvath, *El Camino College*

K. G. Janardan, *Eastern Michigan University*

Steve Kahn, *Anne Arundel Community College*

David Kozlowski, *Triton College*

Don Krekel, *Southeastern Community College*

Marie Langston, *Palm Beach Community College*

Brief Contents