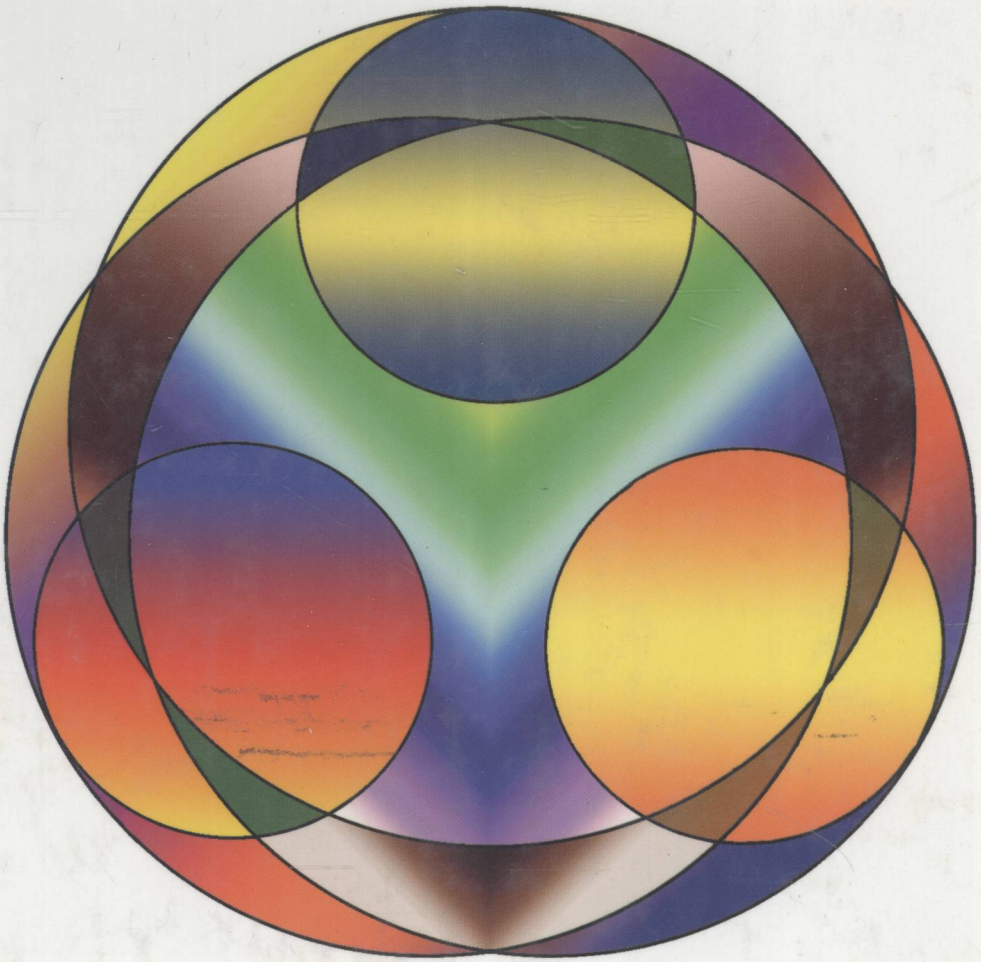
EIGHTH EDITION

ELEMENTARY STATISTICS



Robert Johnson • Patricia Kuby

ELEMENTARY STATISTICS

EIGHTH EDITION

Robert Johnson
Patricia Kuby

Monroe Community College

Duxbury
Thomson Learning™

Pacific Grove • Albany • Belmont • Boston • Cincinnati • Johannesburg • London • Madrid • Melbourne
Mexico City • New York • Scottsdale • Singapore • Tokyo • Toronto

Sponsoring Editor: *Carolyn Crockett*
Marketing Team: *Tom Ziolkowski, Beth Kroenke, Laura Hubrich*
Editorial Assistants: *Kimberly Raburn, Ann Day*
Production Editor: *Tessa McGlasson Avila*
Production Service: *Greg Hubit Bookworks*
Manuscript Editor: *Martha Ghent*
Permissions Editor: *Mary Kay Hancharick*
Interior Design: *Rita Naughton*
Cover Design and Illustration: *Laurie Albrecht*
Art Editor: *Lisa Torri*
Print Buyer: *Vena Dyer*
Typesetting: *Graphic World, Inc.*
Cover Printing: *Phoenix Color Corporation*
Printing and Binding: *R. R. Donnelley/Willard*

For more information, contact:
DUXBURY
511 Forest Lodge Road
Pacific Grove, CA 93950 USA
www.duxbury.com

Printed in the United States of America

10 9 8 7 6 5 4

Library of Congress Cataloging-in-Publication Data

Johnson, Robert Russell
Elementary statistics. — 8th ed. / Robert Johnson, Patricia Kuby.
p. cm.
Includes index.
ISBN 0-534-35676-1
1. Statistics. I. Kuby, Patricia. II. Title.
QA276.12.J64 2000
519.5—dc21 99-23303
CIP

ELEMENTARY STATISTICS

EIGHTH EDITION

Preface

Purpose and Prerequisites

This book was written for use as an introductory course for students who need a working knowledge of statistics but do not have a strong mathematical background. Statistics requires the use of many formulas and an occasional solution of a simple algebraic equation. Those students who have not completed intermediate algebra should complete at least one semester of college mathematics as a prerequisite before attempting this course.

Our Objectives

The primary objective of *Elementary Statistics, Eighth Edition,* is to present a truly readable textbook that will promote learning, understanding, and motivation by presenting statistics in a context that relates to personal experiences. Simply, our goal is a clear and interesting introductory statistics textbook.

Statistics is a practical discipline that evolves with the changing needs of our society. Today's student is the product of a particular cultural environment and is motivated differently from students of a few years ago. In this text we present statistics as a useful tool in learning about the world around us. While studying descriptive and inferential concepts, students will become aware of their real-world applications in such fields as the physical and social sciences, business, economics, and engineering.

Important Ongoing Features

This eighth edition continues to feature the following elements:

- A communication style that reflects current student culture;
- A strong computer flavor, with numerous annotated outputs, exercises, and instructions;
- A focus on interpreting computer output;
- Case Studies based on situations of interest and using real data;
- A Chapter 1 that introduces ideas of variability and data collection, as well as basic terms;
- An early descriptive presentation of linear correlation and linear regression in Chapter 3;

- The use of "word algebra" to express formulas in words;
- The use of margins for margin exercises and margin "conversation";
- The natural flow from sampling distribution, to estimation, to hypothesis testing, to *p*-values;
- Inclusion of both *p*-value and classical approaches to hypothesis testing;
- Two specially designed tables for determining *p*-values;
- Large exercise sets;
- Exercises that demonstrate statistical theory through simulation;
- Many exercises designed to be solved using a computer or calculator;
- All data sets of significant number of data are on the data disk, now called *StatSource CD;*
- Brief biographies of four prominent statisticians.

This Revision

In this eighth edition, we hope the users of the previous editions will appreciate the following improvements:

1. The presentation is more approachable, clearer, and more visual throughout. The presentation is *more approachable* because it is often less technical. For example, each chapter begins with an "everyday happening" type case study; it is *clearer* because the side-by-side presentation allows ease of comparison between the technology used in class and the technologies used elsewhere (see page 67); it is *more visual* as demonstrated by the side-by-side presentations of the *p*-value and classical approaches to hypothesis testing (see page 427).
2. The role of the Chapter Case Study has been expanded to be a more integral part of the chapter; it serves as an introduction and a wrap-up.
3. MINITAB (version 12), EXCEL (97), and TI-83 instructions are displayed side-by-side for ease of comparison (many have access to technology other than just the one used in class, so these three very popular means are shown; see page 423).
4. A rearranged model for the hypothesis test procedure is used to emphasize comparability of the *p*-value and classical approaches; these changes are reflected in Chapters 8 through 14.
5. Many new exercises with real data sets have been added.
6. Important terms have been highlighted for identification and ease of locating when referenced.
7. An Annotated Instructors Edition is being made available.
8. Chapter 1 was rewritten to include sample designs.
9. The sections making up Chapters 9 and 10 have been rearranged. Inferences about variance is now the last section in these chapters.
10. A CD, *StatSource,* has replaced the disk in the back of the textbook. The CD contains many helpful teaching and learning aids.

To the Instructor: The Text as a Teaching Tool

One primary objective of this book is to offer a truly readable presentation of elementary statistics. The chapters are designed to interest and involve students and to guide them step by step, in a logical manner, through the material.

The *Getting Started* section, page xxi, provides a brief explanation of the components of each chapter and makes suggestions about ways to use this book more effectively. It is suggested reading for both the instructor and the student.

The first three chapters are introductory in nature. Chapter 1 is an introduction to the language of statistics; Chapter 2 covers the descriptive presentation of single-variable data, while Chapter 3 is the descriptive presentation of bivariate data. The bivariate material is presented at this point in the book because students often ask about the relationship between two sets of data (such as heights and weights) while studying Chapter 2.

In the chapters on probability (4 and 5), the concepts of permutations and combinations are deliberately avoided. Instead, this material is contained in Appendix A, "Basic Principles of Counting," so that it may be included as the instructor wishes. The binomial coefficient is introduced in connection with the binomial probability distribution in Chapter 5.

The instructor has several options in the selection of topics to be studied in a given course. We consider Chapters 1 through 9 to be the basic core of a course (some sections of Chapters 2, 4, and 6, and all of Chapter 3 may be omitted without affecting continuity). Following the completion of Chapter 9, any combination of Chapters 3 and 10 through 14 may be studied. However, there are two restrictions: Chapter 3 must be studied prior to chapter 13, and Chapter 10 must precede Chapter 12.

To the Student: The Text as a Learning Tool

Statistics is different from other courses:

1. It has its own extensive technical vocabulary.
2. It is highly cumulative, in that many of the concepts you will be learning at each step become the basis for other concepts learned throughout the rest of the course; therefore, failure to master each concept as presented can cause great difficulty later on.
3. It requires very precise measurements and calculations (a seemingly minor error will often be magnified and lead to wrong answers in some procedures).
4. While it is an academic subject, statistics is also very real and touches each of us frequently in everyday life. Plain talk and emphasis on common sense are the book's main characteristics as a learning tool. This approach should allow you—provided you have the necessary basic mathematics skills—to work your way through the course with relative ease. Examples of this approach are: (a) Illustration 1-1 (p. 13), which is used to reemphasize the meaning of the eight basic definitions presented in Section 1.2, and (b) Chapter Objectives for Chapter 2 (p. 36), which use a familiar situation to motivate the topics of Chapter 2.

Our goal in writing this textbook is to motivate and involve you in the statistics that you are learning. Turn to page xxi and read *Getting Started* for a brief explanation of some of the ways you can use this book more effectively to succeed in learning statistics.

Supplements

The **StatSource CD** contains: Data Sets (263 files), Statistical Concept links, Videos (30- to 60-second film clips showing statistics in action), Video Tutorials (1- to 5-minute film clips demonstrating statistical concepts or calculations), Tutorials (dozens of "smart" questions for each chapter to help study), and PowerPoint slides (for lecture presentation or notes).

The **Statistical Tutor** is a student manual that:

a. Contains the complete solutions to all margin exercises and the odd-numbered exercises (the same exercises whose answers are in the back of the book).
b. Contains many helpful hints and suggestions to serve as a guide through the learning process. It includes many summaries and overviews.
c. Contains several review lessons to help refresh materials studied previously in other courses.

The **Instructor's Manual** is also intended to be uncommonly helpful. It contains:

a. Everything that is in the *Statistical Tutor.*
b. The complete solutions to all exercises.
c. Many helpful teaching suggestions that an instructor might incorporate. The notes specifically intended for the instructor are set in a type different from that for the student material.

The **Test Bank** contains a combination of true-false, multiple-choice, short-answer, matching, and computational test questions for each chapter in the text.

The **MINITAB Lab Manual** is a text-specific guide to the MINITAB statistical analysis system that is keyed to the text discussions and examples. Its purpose is to provide the instructor with a flexible means of integrating technology into their courses. It is also intended to integrate the use of technology to enhance presentation of concepts and to give students a feel for what statistics "is." The laboratory exercises are designed to motivate and involve the student in the statistics they are learning. The overriding goal is to strengthen the student's conceptual view without burdening the student with too many manual calculations.

The **EXCEL Manual** is a text-specific guide to Microsoft Excel, intended to demonstrate the statistical capabilities of Excel using examples and problems in *Elementary Statistics, Eighth Edition.* Microsoft Excel offers a complete repertoire of statistics functions for solving the problems in *Elementary Statistics.* This supplement contains detailed chapter-by-chapter solutions of problems in the textbook, including exercises, tips, and Excel demonstrations of key concepts from the text.

Excel features: 1) a large collection of graph types, 2) a rich assortment of random-number generating functions, and 3) instant connection between worksheet numbers and charts. Therefore, Excel supports interactive problem-solving and student demonstrations, and these are provided in the manual where appropriate. Because students majoring in business or related areas will likely be working with Excel in their careers, the supplement includes numerous business-oriented exercises using data from actual case studies.

The purpose of the **TI-83 Manual** is to explain how to use the Texas Instruments TI-83 Plus in order to solve typical statistics problems. The standard TI-83 Plus built-in statistics functions are discussed and several calculator programs are presented. Examples closely follow the text and are solved using the calculator. Keystrokes and screen illustrations are shown for easy reference.

Thomson World Class Learning Testing Tools™ is a fully integrated suite of test creation, delivery, and course management tools. It consists of three software programs—Test, Test On-line, and Manager—that enable you to quickly and easily assess student performance.

Creating practice exams or final exams on paper, via a local area network, or via the Internet, requires only pointing and clicking. This flexible, powerful software lets you create tests based on testbanks provided by Thomson Learning. Testing Tools is flexible, easy to learn, and provided at no cost to adopters of selected Thomson Learning textbooks.

Acknowledgments

We owe a debt to many other books. Many of the ideas, principles, examples, and developments that appear in this text stem from thoughts provoked by these sources.

It is a pleasure to acknowledge the aid and the encouragement we have received throughout the development of this text from students and colleagues at Monroe Community College. In addition, special thanks to all the reviewers who read and offered suggestions about this and previous editions:

Nancy Adcox
Mt. San Antonio College

Paul Alper
College of St. Thomas

William D. Bandes
San Diego Mesa College

Barbara Jean Blass
Oakland Community College

Austin Bonis
Rochester Institute of Technology

Nancy Bowers
Pennsylvania College of Technology

Louis F. Bush
San Diego City College

Ronnie Catipon
Franklin University

Rodney E. Chase
Oakland Community College

Pinyuen Chen
Syracuse University

David M. Crystal
Rochester Institute of Technology

Joyce Curry and Frank C. Denny
Chabot College

Larry Dorn
Fresno Community College

Shirley Dowdy
West Virginia University

Joan Garfield
University of Minnesota General College

Carol Hall
New Mexico State University

Silas Halperin
Syracuse University

Noal Harbertson
California State University, Fresno

Hank Harmeling
North Shore Community College

Bryan A. Haworth
California State College at Bakersfield

Harold Hayford
Pennsylvania State University, Altoona

Marty Hodges
Colorado Technical University

John C. Holahan
Xerox Corporation

James E. Holstein
University of Missouri

Soon B. Hong
Grand Valley State University

Robert Hoyt
Southwestern Montana State

Peter Intarapanach
Southern Connecticut State University

T. Henry Jablonski, Jr.
East Tennessee State University

Brian Jean
Bakersfield University

Jann-Huei Jinn
Grand Valley State University

Sherry Johnson

Meyer M. Kaplan
The William Patterson College of New Jersey

Michael Karelius
American River College

Anand S. Katiyar
McNeese State University

Jane Keller
Metropolitan Community College

Gayle S. Kent
Southern Florida College

Andrew Kim
Westfield State College

Amy Kimchuk
University of the Sciences in Philadelphia

Raymond Knodel
Bemidji State University

Larry Lesser
University of Northern Colorado

Robert O. Maier
El Camino College

Mark Anthony McComb
Mississippi College

Carolyn Meitler
Concordia University

John Meyer
Muhlenberg College

Jeffrey Mock
Diablo Valley College

David Naccarato
University of New Haven

Harold Nemer
Riverside Community College

Dennis O'Brien
University of Wisconsin, LaCrosse

Daniel Powers
University of Texas, Austin

Janet M. Rich
Miami-Dade Junior College

Larry J. Ringer
Texas A & M University

John T. Ritschdorff
Marist College

John Rogers
California State Polytechnic Institute at San Luis Obispo

Neil Rogness
Grand Valley State University

Thomas Rotolo
University of Arizona

Barbara F. Ryan and Thomas A. Ryan
Pennsylvania State University

Robert J. Salhany
Rhode Island College

Sherm Sowby
California State University, Fresno

Howard Stratton
State University of New York at Albany

Larry Stephens
University of Nebraska–Omaha

Paul Stephenson
Grand Valley State University

Thomas Sturm
College of St. Thomas

Edward A. Sylvestre
Eastman Kodak Co.

William Tomhave
University of Minnesota

Bruce Trumbo
California State University, Hayward

Richard Uschold
Canisius College

John C. Van Druff
Fort Steilacoom Community College

Philip A. Van Veidhuizen
University of Alaska

John Vincenzi
Saddleback College

Kenneth D. Wantling
Montgomery College

Mary Wheeler
Monroe Community College

Sharon Whitton
Hofstra University

Don Williams
Austin College

Thanks also to the many authors and publishers who so generously extended reproduction permissions for the news articles and tables used in the text. These acknowledgments are specified individually throughout the text.

And finally, the most significant acknowledgment of all—thank you to our spouses, Barbara and Joe, for your assistance and just "being there."

Robert Johnson
Patricia Kuby

GETTING STARTED

Your Guide to Getting the Most out of *Elementary Statistics*, Eighth Edition

Read this section to become familiar with the components of each chapter and their intended purposes.

First realize that studying statistics is more like studying a foreign language than like studying mathematics. Statistics is more than the mathematics of formulas and data. Statistics includes the processes of problem solving, statistical thinking, data collection, obtaining numerical and graphical results, and the follow-up questioning of those results. Sometimes statistics requires the use of mathematics, and sometimes it does not. Everybody (and everything) is an individual and uniquely different from all others. However, when it comes to a single trait, many values of a single variable taken from many individuals will generally form a pattern. Statistical methodologies are used to describe and help explain these patterns. Statistics uses mathematical techniques to quantify the ideas being investigated and to reduce the information to a numeric format, in which it can be treated graphically or algebraically. When concepts become quantified, they become an application of mathematics, not *the* mathematics.

As important as it is to be able to take a set of data and calculate a statistical value (mean, standard deviation, correlation coefficient, etc.) or draw a graphic display (histogram, scatter diagram, etc.), it is far more important that you understand the circumstances being investigated; that you understand the variables involved; that you understand why you are investigating the problem; and that you learn to question the data and the statistical results. Your life experience and understanding of real-life situations form the foundation for understanding statistics. Don't lose sight of the fact that statistics is about describing the world around us. You will see many statistical examples from business, the physical and social sciences, and many other fields and professions as you study from this textbook.

To get the most out of this book, become familiar with its many learning features. On the following pages are examples of several features this book contains and suggestions on how to make the best use of them. Take a moment to look them over and, then, please use them. Active involvement in learning about statistics is the single most important factor in determining success and satisfaction. Now open up your mind and let your imagination and your natural curiosity go to work.

GETTING STARTED

You cannot just get in your car and start driving and expect to arrive at the correct destination. You must know where you are going and what route to take before you start. Studying statistics is much the same. Every chapter in this textbook opens with three important study tools.

Chapter Outlines appear at the beginning of each chapter to give a schematic overview of what is to be presented. These outlines are annotated to give a first impression of key terms and concepts that appear in the chapter.

Chapter 2

Descriptive Analysis and Presentation of Single-Variable Data

CHAPTER OUTLINE

GRAPHIC PRESENTATION OF DATA

2.1 Graphs, Pareto Diagrams, and Stem-and-Leaf Displays
A **picture** is often worth a thousand words.

2.2 Frequency Distributions and Histograms
An **increase** in the amount of data requires us to modify our techniques.

NUMERICAL DESCRIPTIVE STATISTICS

2.3 Measures of Central Tendency
The four measures of central tendency—mean, median, mode, and midrange—are **average** values.

2.4 Measures of Dispersion
Measures of dispersion—range, variance, and standard deviation—assign numerical values to the **amount of spread** in a set of data.

2.5 Mean and Standard Deviation of Frequency Distribution
The **frequency distribution** is an aid in calculating mean and standard deviation.

2.6 Measures of Position
Measures of position allow us to **compare** one piece of data to the set of data.

2.7 Interpreting and Understanding Standard Deviation
A standard deviation is the length of a **standardized yardstick.**

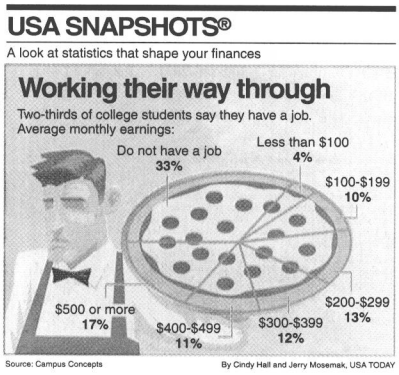
2.8 The Art of Statistical Deception
How the unwitting or the unscrupulous can use **"tricky" graphs** and **insufficient information** to mislead the unwary.

The **Chapter Case Studies** are mostly newspaper articles about everyday kind of phenomenon, with questions that ask the student about the methods to be presented in that chapter; an integral part of the chapter that also serves as an "example introduction" to the Chapter Objectives.

CHAPTER CASE STUDY

Paying Their Own Way

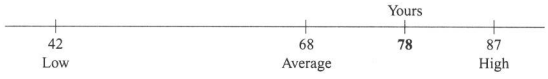
Many students work full-time or part-time, during the academic year or during the summer, or some combination, to earn part of (maybe even all of) their college expenses. Do you work to pay for some part of your college expenses? Do your friends work to pay for part of their college expenses? How much did you or your friends each earn last month? The USA Snapshot®, "Working their way through," that appeared in *USA Today*, March 17, 1998, describes the average monthly earnings of American college students.



Chapter Objectives prepare the way for new material by describing why the material is important and how it relates to previously studied topics. Use Chapter Objectives to understand the motivation for learning this material.

CHAPTER OBJECTIVES

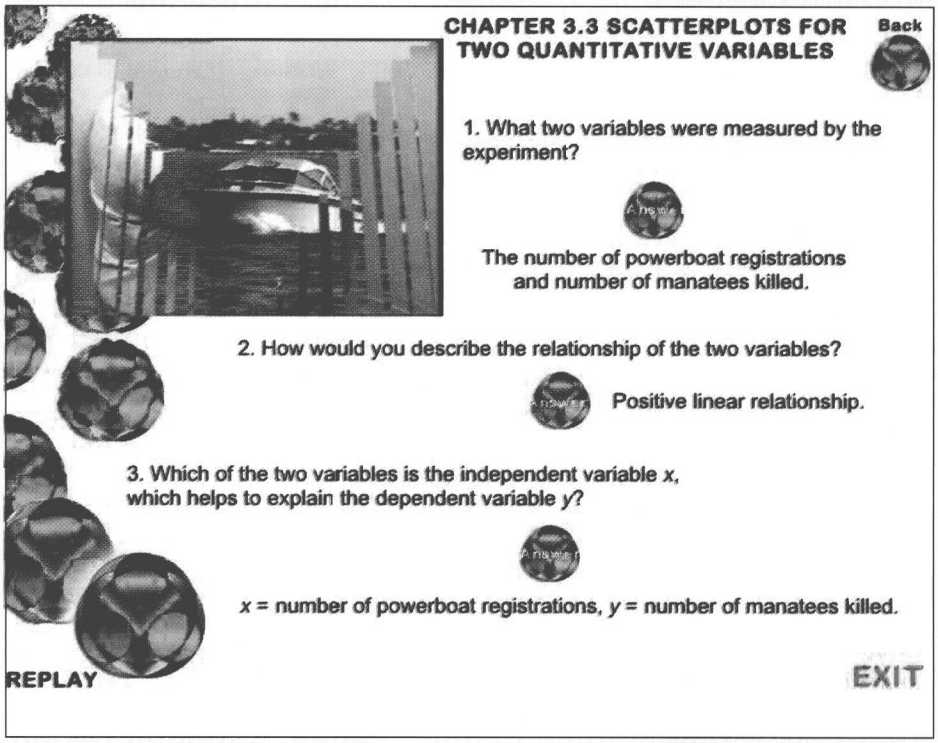
Imagine that you took an exam at the last meeting of your favorite class. Today your instructor returns your exam paper and it has a grade of 78. If you are like most students in this situation, as soon as you see your grade you want to know how your grade compares to those of the rest of the class and you immediately ask, "What was the average exam grade?" Your instructor replies, "The class average was 68." Since 78 is 10 points above the average, you ask, "How close to the top is my grade?" Your instructor replies that the grades ranged from 42 to 87 points. The accompanying figure summarizes the information we have so far.



A third question that is sometimes asked is "How are the grades distributed?" Your instructor replies that half the class had grades between 65 and 75. With this information you conclude that your grade is fairly good.

MAKING STATISTICS COME ALIVE

As you begin each chapter, view the corresponding short **Videos** on the CD (located inside the back cover) for a glimpse at an interesting illustration of statistics at work in a real-life situation. Answer the questions about the video clip.



Case Studies are designed to teach important concepts and to demonstrate how statistics work in the everyday world. Throughout the text, case studies can be found that incorporate statistical concepts as these concepts are presented. Margin exercises also accompany these case studies and provide an excellent way to try out your new knowledge in a real-world setting.

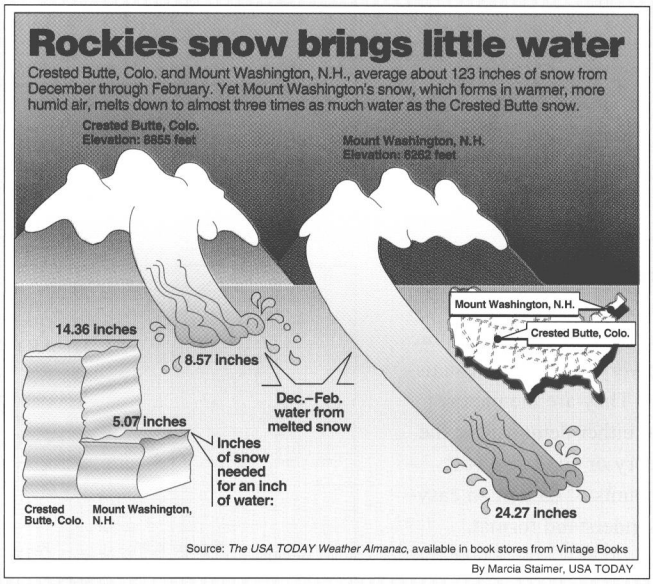
Case Study 8.2

Rockies Snow Brings Little Water

When snow melts it becomes water, sometimes more water than at other times. This newspaper article compares the water content of snow from two areas in the United States that typically get about the same amount of snow annually. However, the water content is very different. There are several point estimates for the average included in the USA Today article.

Exercise 8.15

"Rockies snow brings little water" lists "14.36 inches" and "5.07 inches" as statistics and uses them as point estimates. Describe why these numbers are statistics and why they are also point estimates.



UNDERSTANDING STATISTICS

All the important **definitions** and **key terms** appear in boldface; all indexed words are highlighted as they occur. This will assist the reader in identifying important items, and in locating referenced information.

Two Quantitative Variables

When the bivariate data are the result of two quantitative variables, it is customary to express the data mathematically as **ordered pairs** (x, y), where x is the **input variable** (sometimes called the **independent variable**) and y is the **output variable** (sometimes called the **dependent variable**). The data are said to be ordered because one value, x, is always written first. They are called *paired* because for each x value there is a corresponding y value from the same source. For example, if x is height and y is weight, a height and corresponding weight are recorded for each person. The input variable x is

Completely **worked-out examples** present the step-by-step solution process.

78 Chapter 2 Descriptive Analysis and Presentation of Single-Variable Data

The variance of our sample 6, 3, 8, 5, 3 is found in Table 2.10 using formula (2.6).

TABLE 2.10 Calculating Variance Using Formula (2.6)

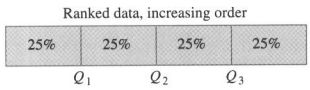
Step 1. Find $\sum x$	Step 2. Find $\bar{x}$	Step 3. Find each $x - \bar{x}$	Step 4. Find $\sum(x - \bar{x})^2$	Step 5. Sample variance
6	$\bar{x} = \frac{\sum x}{n}$	$6 - 5 = 1$	$(1)^2 = 1$	$s^2 = \frac{\sum(x - \bar{x})^2}{n - 1}$
3		$3 - 5 = -2$	$(-2)^2 = 4$	
8	$\bar{x} = \frac{25}{5}$	$8 - 5 = 3$	$(3)^2 = 9$	$s^2 = \frac{18}{4}$
5		$5 - 5 = 0$	$(0)^2 = 0$	
3	$\bar{x} = 5$	$3 - 5 = -2$	$(-2)^2 = 4$	$s^2 = \mathbf{4.5}$
$\sum x = 25$		$\sum(x - \bar{x}) = 0$ (ck)	$\sum(x - \bar{x})^2 = 18$	

Graphical displays can be found in the form of *charts, graphs,* and *tables.* Graphics are very important in statistics. They are the pictures that either demonstrate the theory or condense vast amounts of data in an easy-to-understand format.

QUARTILES

Values of the variable that divide the ranked data into quarters; each set of data has three quartiles. The first quartile, Q_1, is a number such that at most 25% of the data are smaller in value than Q_1 and at most 75% are larger. The second quartile is the median. The third quartile, Q_3, is a number such that at most 75% of the data are smaller in value than Q_3 and at most 25% are larger (see Figure 2.24).

Figure 2.24 Quartiles

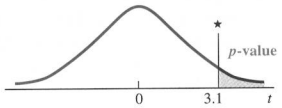


The procedure for determining the value of the quartiles is the same as that for percentiles and is shown in the following description of *percentiles.*

STEP 4 The Probability Distribution:

***P*-VALUE:**

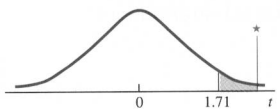
a. Calculate the *p*-value.
Use the right-hand tail since the H_a expresses concern for values related to "greater than."
$\mathbf{P} = P(t > 3.14$, with df = 25) as shown in the figure.



To find the *p*-value, use one of the three methods:
1. Use Table 6 (Appendix B) to place bounds on the *p*-value:
   $\mathbf{P < 0.005}$

CLASSICAL:

OR **a. Determine the critical region and critical value(s).**
The critical region is the right-hand tail since the H_a expresses concern for values related to "greater than." The critical value is obtained from Table 6:
$t_{(25,\,0.05)} = \mathbf{1.71}$.

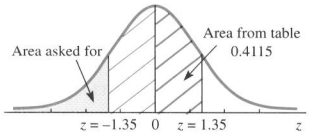


Some exercises are located in the book's margin; these **margin exercises** have been placed there for an initial practice on the adjacent concept. Solving these exercises before reading further is an excellent way to begin homework assignments.

ILLUSTRATION 6.5 ▼

The area to the left of $z = -1.35$ is found by subtracting 0.4115 from 0.5000.

Exercise 6.5
Find the area to the left of $z = -1.53$, $P(z < -1.53)$.



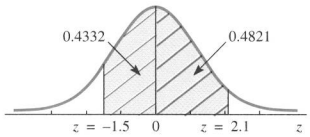
Therefore, we obtain

$$P(z < -1.35) = P(z < 0) - P(-1.35 < z < 0) = 0.5000 - 0.4115 = \mathbf{0.0885}$$

ILLUSTRATION 6.6 ▼

The area between $z = -1.5$ and $z = 2.1$, $P(-1.5 < z < 2.1)$, is found by adding the two areas together. Both probabilities are read directly from Table 3.

Exercise 6.6
Find the area between $z = -1.83$ and $z = 1.23$, $P(-1.83 < z < 1.23)$.



Therefore, we obtain

$$P(-1.5 < z < 2.1) = P(-1.5 < z < 0) + P(0 < z < 2.1)$$
$$= 0.4332 + 0.4821 = \mathbf{0.9153}$$