ALGEBRA

Exploring Symbols

G. BURRILL, M. CLIFFORD, R. SCHEAFFER

DATA-DRIVEN MATHEMATICS

DALE SEYMOUR PUBLICATIONS®

Exploring Symbols: An Introduction to Expressions and Functions

DATA-DRIVEN MATHEMATICS

Gail F. Burrill, Miriam Clifford, and Richard Scheaffer

Dale Seymour Publications®
Orangeburg, New York

This material was produced as a part of the American Statistical Association's Project "A Data-Driven Curriculum Strand for High School" with funding through the National Science Foundation, Grant #MDR-9054648. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Managing Editor: Cathy Anderson, Alan MacDonell

Senior Math Editor: Carol Zacny

Project Editor: Nancy Anderson

Production/Manufacturing Director: Janet Yearian

Senior Production Coordinator: Alan Noyes

Design Manager: Jeff Kelly

Text and Cover Design: Christy Butterfield

Composition: Joan Olson

Art: Carl Yoshihira

Cover Photo: Kirk Yarnell/Bald Eagle

This book is published by Dale Seymour Publications®, an imprint of Addison Wesley Longman, Inc.

Dale Seymour Publications
125 Greenbush Road South
Orangeburg, NY 10962
Customer Service: 800-872-1100

Printed in the United States of America.

Order number DS21173

ISBN 1-57232-230-6

1 2 3 4 5 6 7 8 9 10-ML-01 00 99 98 97

This Book Is Printed On Recycled Paper

DALE SEYMOUR PUBLICATIONS®

Authors

Gail F. Burrill
Whitnall High School
Greenfield, Wisconsin
University of Wisconsin-Madison
Madison, Wisconsin

Miriam Clifford
Nicolet High School
Glendale, Wisconsin

Richard Scheaffer
University of Florida
Gainesville, Florida

Consultants

Jack Burrill
Whitnall High School
Greenfield, Wisconsin
University of Wisconsin-Madison
Madison, Wisconsin

Emily Errthum
Homestead High School
Mequon, Wisconsin

Henry Kranendonk
Rufus King High School
Milwaukee, Wisconsin

Vince O'Connor
Milwaukee Public Schools
Milwaukee, Wisconsin

Maria Mastromatteo
Brown Middle School
Ravenna, Ohio

Jeffrey Witmer
Oberlin College
Oberlin, Ohio

Murray Seigel
Marietta City Schools
Marietta, Georgia

Data-Driven Mathematics Leadership Team

Miriam Clifford
Nicolet High School
Glendale, Wisconsin

James M. Landwehr
Bell Laboratories
Lucent Technologies
Murray Hill, New Jersey

Kenneth Sherrick
Berlin High School
Berlin, Connecticut

Gail F. Burrill
Whitnall High School
Greenfield, Wisconsin
University of Wisconsin-Madison
Madison, Wisconsin

Richard Scheaffer
University of Florida
Gainesville, Florida

Acknowledgments

The authors thank the following people for their assistance during the preparation of this module:

- The many teachers who reviewed drafts and participated in field tests of the manuscripts
- The members of the *Data-Driven Mathematics* leadership team, the consultants, and the writers
- Nancy Kinard, Ron Moreland, Peggy Layton, and Kay Williams for their advice and suggestions in the early stages of the writing
- Richard Crowe, Barbara Perry, and Tom Hyde for their thoughtful and careful review of the early drafts
- Kathryn Rowe and Wayne Jones for their help in organizing the field-test process and the Leadership Workshops
- Barbara Shannon for many hours of word processing and secretarial services
- Jean Moon for her advice on how to improve the field-test process
- Beth Cole, Bryan Cole, and Maria Mastromatteo for writing answers for the Teacher's Edition
- The many students from Nicolet High School, Whitnall High School, and the University of Florida, Gainesville, who helped shape the ideas as they were being developed

Table of Contents

About *Data-Driven Mathematics*

Historically, the purposes of secondary-school mathematics have been to provide students with opportunities to acquire the mathematical knowledge needed for daily life and effective citizenship, to prepare students for the workforce, and to prepare students for postsecondary education. In order to accomplish these purposes today, students must be able to analyze, interpret, and communicate information from data.

Data-Driven Mathematics is a series of modules meant to complement a mathematics curriculum in the process of reform. The modules offer materials that integrate data analysis with high-school mathematics courses. Using these materials will help teachers motivate, develop, and reinforce concepts taught in current texts. The materials incorporate major concepts from data analysis to provide realistic situations for the development of mathematical knowledge and realistic opportunities for practice. The extensive use of real data provides opportunities for students to engage in meaningful mathematics. The use of real-world examples increases student motivation and provides opportunities to apply the mathematics taught in secondary school.

The project, funded by the National Science Foundation, included writing and field testing the modules, and holding conferences for teachers to introduce them to the materials and to seek their input on the form and direction of the modules. The modules are the result of a collaboration between statisticians and teachers who have agreed on statistical concepts most important for students to know and the relationship of these concepts to the secondary mathematics curriculum.

Using This Module

You can solve many problems using symbols. Mathematical reasoning requires you to think abstractly so that a solution to one problem can be generalized to solve other similar problems. This abstraction begins by first expressing the problem using symbols. It is also efficient and helpful to use symbols to communicate with others about a given problem. Symbols are even necessary to interact with technology. Symbolic expression is a key that helps unlock the power of mathematics.

In this module, lessons are built using symbols to describe real data and to solve problems. The data are often organized in a spreadsheet, and a graphing calculator or computer can be used for performing the necessary calculations. Some standard formulas are investigated; you will have the opportunity to create your own formulas in other situations. A special mathematical rule called a *function* will be explored in Unit II.

Content

Mathematics

- Variables to represent data in tables and graphs
- Graphical representations of data
- Translations of statements into symbols
- Functions as ordered pairs, graphs, and formulas
- Mathematical formulas
- Evaluation of expressions and formulas

Statistics

- Calculation and interpretation of summary statistics
- Symbolic expressions for statistical summaries
- Data transformation and summary statistics
- Relationships between summary statistics and features of a graph
- Plots over time

Unit I

Variables, Expressions, and Formulas

INTRODUCTORY ACTIVITY

Evaluating with Formulas

Report Cards

Your teacher has to assemble information about your understanding and progress in class every marking period.

1. Make a list of the factors you think should be considered by your teacher when evaluating your work.
2. Decide how you think these factors should be combined to produce your grade. Write a formula that your teacher could use to find grades for your class.
3. Do you think a formula is necessary? Why or why not?
4. Compare your answers for the first three questions to those of your classmates. How were they alike? How were they different?
5. If your teacher uses a formula, how does that compare to the one you created?

LESSON 1

Variables and Formulas

Who is the best hitter on your school softball team?

Who gets on base nearly every time or hits many home runs?

How can you keep track of the "hits" a ball player has? Does getting on base always count as a hit?

Why would a major league baseball team offer a slugger a great deal of money to play for them?

OBJECTIVES

Identify and use variables in formulas.

Evaluate expressions.

Use a spreadsheet for a series of calculations.

Baseball has been called "the great American pastime." Thousands of Americans follow the game and can quote statistics on players throughout the years. Major league baseball hitters are of various types. Some players often hit the ball accurately but not very far and, as a result, reach first base safely. Such a hit is called a *single*. Other players hit the ball farther and reach second base for a *double;* reach third base, a *triple;* or make it all around the bases and back to home base for a *home run*. The players who hit many doubles, triples, home runs, and not many singles are called *sluggers*. Generally, sluggers are among the highest paid players in baseball. What are some factors you could use to measure a player's slugging ability?

INVESTIGATE

Slugging It Out!

Babe Ruth, who played for the New York Yankees in the 1920s, has been called one of the greatest baseball players of all time. How does he compare to Mickey Mantle, who played for the Yankees in the 1950s, or to Barry Bonds, an outstanding hitter for the San Francisco Giants in 1993? Baseball almanacs have records that go back to 1901 with a variety of variables related to a player's ability to hit the ball.

Among the variables recorded for baseball players are:

the number of times at bat (AB)
the number of hits (H)
the number of doubles (D)
the number of triples (T)
the number of home runs (HR)

Variables are descriptions of quantities that may change as situations change, such as the number of home runs a player hits. A variable is determined by the context of the problem. Sometimes variables are described by words, but often they are described by symbols or letters. Variables, often recorded in a table or a spreadsheet format, can be used to write a general rule, or *formula.*

Table 1.1
Baseball Sluggers

Players	*AB*	*H*	*D*	*T*	*HR*
Bonds (1993)	539	181	38	4	46
Ruth (1920)	458	172	36	9	54
Mantle (1956)	533	188	22	5	52

Source: *Sports Illustrated Baseball Almanac,* 1993

Discussion and Practice

1. Consider Table 1.1.

a. Give two reasons for using the letters *AB, H, D,* and so on, rather than words to describe each variable.

b. Why did this record book not report the number of singles?

c. Find the number of singles for each player in the table.

d. Use the information in the table to write a formula to compute the number of singles. Did you use words or symbols to write your formula? Explain why you chose the method you used.

2. The total number of bases a player reaches on hits could be a factor in deciding whether or not that player is a slugger.

Another way to say this mathematically is that slugging average is a *function* of the total number of bases. From the data in Table 1.1, you can compute the total number of bases *TB* for each player.

a. A single counts as one base. How many bases are counted for a triple? A home run?

b. Use the variables defined in the table to write a formula to compute a player's total number of bases *TB*.

c. Suppose a baseball player had two singles and a triple in four at-bats during a game. What is the player's *TB* for that particular game?

3. Baseball players are often rated by their slugging average *SA*.

   a. How would you calculate a slugging average for a player if you were given the information in the Table 1.1? Describe your process either by writing a formula or by using words.

   b. What does your method give for each of the three players?

   c. How do your results compare to those of your classmates?

   d. Do you think using slugging average to rate a player's batting ability is appropriate? Why or why not?

4. Think about your slugging-average method.

   a. Does your method take into account the number of singles? Strikeouts?

   b. Baseball leagues actually calculate the slugging averages with the formula

   $$\frac{TB}{AB}$$

   where *AB* is the number of at-bats. This quotient is usually rounded to the nearest thousandth, or to three decimal places. How does your method compare to the one used by the baseball leagues?

5. Use the league formula and the information in Table 1.1 for the following problems.

   a. In 1920, Babe Ruth set the major-league record for slugging average. Compute the Babe's 1920 slugging average if you have not already done so.

b. Between 1942 and 1992, Mickey Mantle was the only player whose slugging average was high enough to be in the top ten all-time major league slugging averages. What was his 1956 slugging average?

c. In 1993, Barry Bonds of the San Francisco Giants was voted the Most Valuable Player in the National League. This was the third time in four years that Bonds had won the award. Compute his 1993 slugging average. Compare Bonds's average to Babe Ruth's 1920 average.

d. Rank the three players by the slugging averages you found in parts a–c. How does their actual ranking for slugging average compare with the ranking you found with your method?

6. During the 1993 season, Barry Bonds received 126 walks. A walk does not count as an at-bat.

a. Suppose Bonds had been less selective in his batting and 40 of those walks had been strikeouts, which *do* count as at-bats. Compute his slugging average in this situation, leaving all other data the same.

b. Suppose the 126 walks counted as singles, since in each case the player reaches first base. Compute Barry Bonds's slugging average in this situation.

7. Barry Bonds is a slugger, and he is paid a very large salary to play baseball. In fact, he was paid more money to play baseball for one week in 1993 than Babe Ruth was paid for the entire 1920 season. Give some possible reasons for this great difference in salaries.

8. Suppose a player has 500 at-bats and gets 175 hits.

a. What is the minimum slugging average this player could have? The maximum?

b. Find a combination of doubles, triples, and home runs that gives a slugging average greater than 1.000.

c. If 60 home runs is the maximum for this player, what is the greatest slugging average he could have?

Using Technology

You can use either a graphing calculator or a computer spreadsheet program to find the slugging average for a great number of players. If you have a graphing calculator, follow the steps below. These steps are appropriate for a TI-83 but can be easily modified for another calculator with list capabilities.

Select the STAT edit menu and enter the variables in these five lists:

L1 = singles
L2 = doubles
L3 = triples
L4 = home runs
L5 = at-bats

Return to the home screen. Type the formula you produced for the slugging average, but use the list name L5 in place of at-bats *AB*, the list name L1 in place of singles *S*, and so on. Store the results of the formula in an unused list: L6 (L1 + 2L2 + 3L3 + 4L4) ÷ L5. Use ENTER to apply the formula, return to the STAT edit menu, and find List 6. The slugging average for the players should be in that list.

With a computer spreadsheet program, you can quickly and easily perform calculations. In order to use a spreadsheet, you need a name for each variable, and you also need formulas to tell the computer what to do with the variables.

Use your calculator or spreadsheet to do the following problem.

9. In Connecticut, high-school football teams are rated on the strength of their schedules, their win-loss records, and their opponents' records. Table 1.2 lists seven high schools and their team stats.

Table 1.2
High-School Football Ranking

School	Wins *WC*	Ties *TC*	Wins *WH*	Ties *TH*	Opponents' Wins *OW*
Berlin	5	0	1	0	10
Darien	1	0	0	0	7
East Lyme	2	0	3	0	25
New London	1	1	1	0	25
Plainfield	2	1	0	1	20
Redding	4	0	2	0	21
Wolcott	6	0	0	1	22

Each school is placed in a class according to enrollment. The rating formula uses the following variables:

the number of wins in class *WC*, each worth 100 points
the number of ties in class *TC*, each worth 50 points
the number of wins in a higher class *WH*, each worth 110 points
the number of ties in higher class *TH*, each worth 55 points
the total number of opponents' wins *OW*, each worth 10 points
the number of games *N*

Every team plays 9 games, so the formula for the season point value is

$$R = \frac{(100 \cdot WC + 50 \cdot TC + 110 \cdot WH + 55 \cdot TH + 10 \cdot OW)}{9}.$$

Write the formula in a language appropriate for your calculator.

a. Find the season point value for each team and rank the teams according to their point value.

b. Which is the greater advantage, playing a tough schedule, that is, teams who win most of their games, or winning most of your games? Explain your reasoning based on the formula.

Rating Quarterbacks

The slugging average for baseball seems to be a fairly natural way to look at a batter's power. The numbers in the formula for the total number of bases *TB* seem logical. However, methods for rating performers in other sports are not always so clear.

10. What are some variables that might be associated with

a. figure skaters?

b. spikers in volleyball?

c. backs in soccer?

11. Table 1.3 shows the all-time leading passers in the National Football League (NFL) as of the start of the 1994 season, along with the variables used by the NFL to rate its quarterbacks on passing ability. Len Dawson has a rating of 82.56. The rating is a function of the number of completions, attempts, yards gained, touchdowns, and interceptions.

The formula used by the NFL is

$$R = \frac{50 + 2000\frac{C}{A} + 8000\frac{T}{A} - 10{,}000\frac{I}{A} + 100\frac{Y}{A}}{24}, \text{ where}$$

R = rating
A = attempts
C = number of completions
T = number of touchdowns
I = number of interceptions
Y = yards gained by passing

Table 1.3
NFL All-Time Passing Leaders After 1993–94

Player	Attempts *A*	Completions *C*	Touchdowns *T*	Interceptions *I*	Yards Gained *Y*	Rating *R*
Ken Anderson	4475	2654	197	160	32,838	
Len Dawson	3741	2136	239	183	28,711	
Brett Favre*	1580	983	70	53	10,412	
Sonny Jurgensen	4262	2433	255	189	32,224	
Jim Kelly *	3942	2397	201	143	29,527	
Bernie Kosar *	3225	1896	120	82	22,394	
David Kreig*	4390	2562	231	166	32,114	
Neil Lomax	3153	1817	136	90	22,771	
Dan Marino *	6049	3604	328	185	45,173	
Joe Montana	5391	3409	273	139	40,551	
Danny White	2950	1761	155	132	21,959	
Roger Staubach	2958	1685	153	109	22,700	
Steve Young*	2429	1546	140	68	19,869	

Source: *The Sporting News 1995 Football Register*

a. Find the rating for Ken Anderson.

b. Link the data into your calculator or use a spreadsheet to find the ratings *R* for the other quarterbacks on the list. Record the results on *Activity Sheet 1*. Who are the top five quarterbacks?

c. The starred quarterbacks were still active in 1995. If there were only one variable in which they could improve to become the leading quarterback, which variable should it be? Why?

12. Consider the numbers used in the formula for quarterback ratings.

a. Why does one variable have a minus sign in front of it?

b. How many times as important is a touchdown as a simple completion?

c. Does a touchdown have a greater impact on the rating than an interception?

d. Why do you think the formula uses yards per attempt rather than just total yards passing?

Summary

Variables are quantities that change for different situations. The variables involved in a situation depend on the context and on changes within that context. It is often more efficient to represent a variable by using a letter rather than writing out a description. Variables can be combined and used to write formulas that describe a series of calculations for a given process. You can evaluate a formula by using given values for the variables. A spreadsheet or graphing calculator can generate a large set of values for a formula very quickly.

Practice and Applications

13. The records for the 1993 NFL leading quarterbacks are given in Table 1.4.

a. Use the formula on page 10 to determine how they ranked and record the results on *Activity Sheet 1*. Overall, how do the 1993 quarterbacks compare to the all-time leading passers?

Table 1.4
1993 National Football League Passing Leaders

Player	Attempts *A*	Completions *C*	Touchdowns *T*	Interceptions *I*	Yards Gained *Y*	Rating *R*
Troy Aikman	392	271	15	6	3100	
Steve Beuerlein	418	258	18	17	3164	
Buddy Brister	309	181	14	5	1905	
Brett Favre	471	302	18	13	3227	
Jim Harbaugh	325	200	7	11	2002	
Bobby Hebert	430	263	24	17	2978	
Jim McMahon	331	200	9	8	1968	
Phil Simms	400	247	15	9	3038	
Wade Wilson	388	221	12	15	2457	
Steve Young	462	314	29	16	4023	

Source: *The Universal Almanac*, 1995

b. Find the passing statistics for your favorite NFL quarterback from last season. How does he compare to the all-time leaders?

14. The National College Athletic Association (NCAA) uses this slightly different formula to rate quarterbacks:

$$R = 100\frac{C}{A} + 330\frac{T}{A} - 200\frac{I}{A} + 8.4\frac{Y}{A}$$

a. Does the NCAA use the same set of variables as the NFL?

b. Use the NCAA formula to find the ratings for the NFL quarterbacks in Table 1.4. Then compare the ratings to those calculated by the NFL formula.

15. Variables are used in many other formulas. Some formulas you have studied in mathematics are given below. Indicate what the variables represent in each of the other formulas and describe how the formula is used.

a. $A = l \cdot w$

b. $A = \frac{1}{2}(b \cdot h)$

c. $V = \pi r^2 h$

d. $a^2 + b^2 = c^2$

Extension

16. The original NFL rating formula was given as

$$R = \frac{50 + 2000\frac{C}{A} + 8000\frac{T}{A} - 10{,}000\frac{I}{A} + 100\frac{Y}{A}}{24}.$$

Write an equivalent expression with fewer divisions that might be simpler to calculate.

17. Obtain the data for your school's softball and baseball teams. Compute the slugging averages for last season for the players on these teams. Was there a slugger at your school?

18. Obtain the formula for rating divers in a diving competition. Collect the data from a competition and show how the formula is used.

LESSON 2

Formulas That Manage Money

How do storeowners keep track of the salary their employees earn?

Do all employees get paid the same wages?

Are some hours more difficult to work than others? Do people working the more difficult times deserve a higher salary?

What are some variables storeowners need to consider?

The school is opening a store to sell school supplies and T-shirts bearing the school emblem. Your class is assigned to organize the management of the employees who will work in the store.

OBJECTIVES

Identify and use variables in formulas.

Evaluate expressions.

Use a spreadsheet for a series of calculations.

INVESTIGATE

Managing Payrolls

If you are the store manager, you have to think carefully about all of the factors in setting up and running a store. Variables could describe quantities such as the number of hours an employee works. Recall that a variable is determined by the context of the problem and that it may be described by words, but more often by a symbol or a letter.

There are many variables to record for a weekly payroll. A convenient way to do this is to use a spreadsheet. Recall that spreadsheets are software programs that allow you to quickly and easily perform calculations. If you use a spreadsheet, you need a name for each variable, and you also need formulas to tell the computer what to do with the variables. Try to write a general rule or formula to summarize and record payroll information.

Discussion and Practice

1. As a class, make a list of some of the factors involved in running a store and paying employees.
2. Come to some agreement on each of the following:
   a. How many hours per week should the store be open?
   b. How many employees will be needed to service the customers while the store is open?
   c. How many hours per week should employees work?
   d. What is an appropriate hourly pay rate?
   e. Should supervisors be paid more than other employees?

The symbols R, H, and P can be used as a shorthand for the variables hourly rate, hours worked, and total pay, respectively.

Table 2.1
Payroll Records

Worker's Name	Hourly Pay Rate R	Hours Worked H	Total Pay P

3. Based on your decisions in Problem 2, set up a spreadsheet similar to Table 2.1. You might have different column headings, depending on your choice of variables.
   a. Make up a roster of four employees for the store, and fill in the hourly rate you agree to pay each one, remembering that rates may differ for employees with different responsibilities; the number of hours each employee is likely to work in a given week; and the other variables you selected.
   b. An employee's pay each week depends on the number of hours worked and the hourly pay rate. You can say that

the total pay P for the week for each employee is a function of the variables R and H. What are some reasonable limits on the size of R and H?

c. Find a formula for total pay per employee and write the formula in the last column on your spreadsheet.

If the work is done on a graphing calculator, use the STAT edit menu. Enter the hourly pay rate in List 1 and the number of hours worked in List 2. The formula you created above can be used with the calculator to construct the values for P from the other two columns. Return to the home screen (2nd QUIT), write the formula for total pay, and store the result in List 3. For example: LI * L2 → L3. When you press ENTER, the values for P will be in List 3.

4. Find the total amount of employee wages paid during one week.

a. With your calculator, use the STAT calc menu. Calculate the one-variable statistics for the list you want to total. The total amount will be displayed as $\sum x$.

b. Suppose you paid each employee \$20 plus an hourly rate that was \$0.50 less than your original rate. Write a formula for the amount each employee would earn under this plan.

c. Would you pay out more or less for the week under the new plan?

d. List the advantages and disadvantages of each plan.

Employees who are successful at selling the products a store has to offer are often rewarded for their hard work. One way to reward them is to give them a *commission* in addition to their regular pay. A commission is pay based on how much an employee sells. A simple way to give a commission is to add a certain percent of the total sales the employee made during the week to the employee's pay. For example, if a salesperson's total sales were \$250, with a 10% commission she would earn an extra \$25 (10% of \$250).

5. Decide on a plan to reward good employees with a commission. To calculate the amount of the commission, estimate the employee's total sales for a week. The store makes a profit of about 20% on most items it sells.

a. What variable(s) should be added to the spreadsheet to keep track of each employee's total pay for the week?

b. Write a formula for the commission with the agreed-on percent, and use it to define the total commission.

The final salary, all pay received by an employee both in hourly wages and commissions, is sometimes called the *gross pay G*. A spreadsheet with all of the information might look like Table 2.2.

Table 2.2
Wages and Commissions

Worker's Name	Hourly Pay Rate *R*	Hours Worked *H*	Base Pay *P*	Amount of Sales *S*	Total Commission *C*	Gross Pay *G*
________	________	________	________	________	________	________
________	________	________	________	________	________	________
________	________	________	________	________	________	________
________	________	________	________	________	________	________

6. a. Write G as a formula involving C and P. Fill in this column on the spreadsheet. You may use *Activity Sheet 2* to record your results.

b. Express G in terms of other variables on the spreadsheet *without* using C and P.

c. Does this spreadsheet include everything a store manager would want to know about variables affecting the weekly payroll? Explain your answer.

Taxes and More Taxes

A weekly paycheck is often calculated using spreadsheet software. State income tax, federal income tax, and Social Security (FICA) are all deducted from your gross salary. The result is your take-home pay. Table 2.3 shows the data necessary for a payroll spreadsheet for seven workers.

Table 2.3
Gross Salary

Employee	Hourly Pay Rate *R*	Hours Worked *H*	Gross Pay *G*	Federal Income Tax *FI*	State Income Tax *SI*	Social Security *SS*	Total Pay *P*
Emma	\$5.65	40	______	______	______	______	______
Bond	\$4.70	30	______	______	______	______	______
Latisha	\$5.80	20	______	______	______	______	______
Matt	\$6.70	30	______	______	______	______	______
Terine	\$9.80	35	______	______	______	______	______
Sal	\$8.35	40	______	______	______	______	______
Myrna	\$7.10	20	______	______	______	______	______

7. a. Suppose the federal income tax rate for all of these employees is 15%, the state income tax rate is 2%, and Social Security is 7.65%. Write a formula to find the weekly pay for each employee. Record your results on *Activity Sheet 2*.

b. How much did each employee earn?

c. Compare your formula and results with those of a classmate. Were your formulas different? If so, describe the differences.

8. Suppose each employee worked overtime at a rate of \$10 per hour for 6 hours during a particularly busy week.

a. Explain how to adjust the table to accommodate overtime pay.

b. Write a formula to find the take-home pay for the employees for that week.

c. How much did each employee earn that week?

Summary

Variables are quantities that change for different situations. The variables involved in a situation depend on the context and what changes within that context. It is often more efficient to represent a variable by using a letter rather than writing out a description. Variables can be combined in a formula that describes a standard series of calculations for a given process. Graphing calculators or spreadsheets can apply a formula to a very large and complicated set of data and find a result very quickly.

Practice and Applications

9. The payroll information for the school store for a week is given below. Use *Activity Sheet* 2 to record your information.

Table 2.4
Weekly Payroll

Name	Hourly Pay Rate *R*	Hours Worked *H*	Base Pay *P*	Amount of Sales *S*	Total Commission *C*	Gross Pay *G*
Lucinda	$4.25	4	______	$75.99	______	______
Frank	$4.25	6	______	$88.54	______	______
Kordell	$5.15	8	______	$72.00	______	______
Brittany	$5.15	8	______	$74.95	______	______
Mikhail	$4.25	3.5	______	$68.35	______	______
Erin	$4.25	10	______	$91.50	______	______

a. Find the gross pay for each employee using a 5% commission.

b. How much did it cost to staff the store during the week?

c. For which columns of the spreadsheet does it make sense to find the sum? Find those sums and explain what each sum represents.

d. For which columns would the mean, or average, of the column be a useful number? Justify your answer by explaining how each mean could be used.

10. What advice do you have for the store managers based on your work?

11. Suppose an extra one hour of pay were added to the gross pay of all employees for the week represented in Table 2.4. Use a spreadsheet, and

a. write a formula to find the adjusted gross pay.

b. add a new column to show the adjusted gross pay.

12. By how much will the total payroll in problem 11 be increased?

13. Many people rent cars—when a person's car breaks down, when someone flies into a city and needs a car for traveling around the area, or when someone does not own a car and needs one for a specific occasion.

a. Table 2.5 contains some of the variables identified by car-rental agencies. It gives the cost at each agency of renting a midsize car in Phoenix, Arizona, for a seven-day period from October 3 to October 9, 1994, or for a three-day period from October 7 to October 9, 1994. List any other variables you think might be important to consider when renting a car.

b. Write a formula to calculate the cost of renting the car for a three-day period. Include insurance, taxes, and surcharge using the information in Table 2.5.

Table 2.5
Weekly Car Rentals

Rental Agency	Weekly Rate *W*	Daily Rate *D*	Insurance per Day *I*	Daily Surcharge *S*	Tax Rate *T*	Total Cost *C*
Avis	$155.54	$25.19	$12.99	$2.50	13.8%	
Dollar	$157.00	$36.95	$11.95	$2.50	13.8%	
Hertz	$182.99	$34.19	$12.99	$2.50	19%	
Budget	$158.65	$27.95	$11.95	$2.50	13.7%	
Alamo	$143.09	$27.99	$12.99	$2.50	13.8%	
Thrifty	$153.95	$27.95	$9.95	$2.50	12.0%	
National	$149.40	$26.00	$12.00	$2.50	13.8%	

Source: Phone survey, 1994

c. Enter the data using a calculator or a spreadsheet. Which agency is the least expensive for a three-day rental?

d. Modify your formula to calculate the cost of renting a car for a three-day period using the weekly rate, again including insurance, taxes, and surcharge. Is the agency from part c still the least expensive?

14. Return to problems 9–12 about the school store. The rules for compensating employees have changed. Each employee is guaranteed a pay rate of $5 per hour. However, the commission rate agreed upon is now doubled. If the total commission in any one week is greater than the base pay, the employee receives only the commission for that week. If the commission is less than the base pay, the employee receives only the base pay.

a. For the data on your spreadsheet, add a column for gross pay using the new rules. What happens to the store's total payroll under this new plan?

b. Suppose Erin always works 20 hours a week. Write an expression that shows how her gross pay, under the new rules, is related to her sales.

c. Erin sells the following amounts per week for an eight-week period: \$30, \$35, \$40, \$45, \$50, \$55, \$60, \$65. Find her total gross pay for that time period.

d. Make a graph, plotting Erin's gross pay along the vertical scale and amount of sales along the horizontal scale. How much does Erin need to sell before her gross pay becomes all commission rather than the guaranteed base?

e. Do you think the store is better off under the new pay rules or the old pay rules? Explain your answer.

Project

15. Make a list of the categories in which you spend money. Keep track of every penny you spend in each category for a week.

a. Set up a spreadsheet and use the information you collected to estimate how much money you would spend in each category throughout an entire year. Consider the week for which you collected data a typical week.

b. In what category did you spend the most?

c. Based on your total spending for this week, estimate your total spending for one year.

LESSON 3

Creating Your Own Formulas

What factors do you use to rate movies?

What is the difference between rating movies and ranking movies?

How is class rank determined in your school?

How would you rate the football players in your school this year?

Is your football team ranked in the top ten of your state?

How high would you rate your family car?

Ratings and ranks are regularly used to express the way people think and feel about such things as performances, achievements, products, and events. Cars are rated on various performance criteria and then ranked against one another. Athletes are rated on their performance and then ranked in comparison to others. In Olympic figure skating, for example, a skater with a high rating earns a rank close to the top of the scale. The skater with the highest rating is ranked number one and is the winner.

OBJECTIVES

Use formulas written in words.

Use variables to develop formulas for ranking.

Translate sentences into symbols.

INVESTIGATE

Ratings and Ranks

In Lesson 1, you looked at ratings and some standard formulas used by different sports to rate performances. In this lesson, you will generate your own formulas.

Discussion and Practice

1. Consider the dialogue in the box below.

> Peter Travers of *Rolling Stone* says "'Quiz Show' is the best American movie this year." Mike Clark of *USA Today* rates the movie as " * * * *..." PG-13

Source: *The New York Times*, September 25, 1994

a. What does PG mean? Four stars? Best American movie?

b. Identify each of the descriptors—PG, four stars, and best American movie—as a rating or a rank.

2. Florida regulators ranked 36 auto manufacturers for the period from January 1992 to January 1994 according to complaints received from consumers about chronic defects. The car manufacturer ranked first, the one with fewest complaints, was Oldsmobile. Buick was second and Toyota was third. The lowest rank went to Porsche.

a. Why would anyone be interested in such a ranking of cars?

b. What other variables, besides the number of complaints, should regulators consider?

3. Ratings and ranks are related, but different.

a. Describe the difference between a rating and a rank.

b. The A.C. Nielsen Company reports that the TV show *Frasier* was watched by 15.0% of households with television sets during its time slot for the week of August 28, 1994. Is this information in the form of a rating or a rank? Explain your answer.

c. The Nielsen data also indicated that *Frasier* had the third highest percent of the viewing audience for its time slot. Is this information a rating or a rank? Explain.

d. Name at least two other situations in which

i. something is ranked.

ii. something is rated.

In each case, explain why, and to whom, the results are of interest.

A widely used source of ratings and ranks is the *Places Rated Almanac* (by David Savageau and Richard Boyer, Prentice Hall

Travel, 1994). Here, all metropolitan areas of the United States are rated in 10 categories, including housing costs, jobs, education, crime, and recreational facilities. A brief description of the variables in each category and how the ratings are established is provided below.

- Housing costs *H*—the annual payments on a 15-year, 8% mortgage for an average-priced home after making a 20% down payment
- Jobs *J*—the product of the estimated percent increase in new jobs from 1993 to 1998 and the number of new jobs created between 1993 and 1998, added to a base score of 2000
- Education *E*—enrollment in two-year colleges divided by 100, plus enrollment in private colleges divided by 75, plus enrollment in public colleges divided by 50
- Crime C—the violent crime rate plus one-tenth of the property crime rate; crime rates are reported as the number of crimes per 100,000 residents.
- Recreation *R*—a score directly proportional to the number of public golf courses, good restaurants, zoos, aquariums, professional sports teams, miles of coastline on oceans or the Great Lakes, national forests, parks and wildlife refuges, and state parks

4. According to the *Places Rated Almanac*, Pittsburgh, Pennsylvania, is predicted to lose 1,763 blue-collar jobs by 1998 and to gain 36,053 white-collar jobs. The projected job growth rate is 2.75%, which means that the number of jobs is expected to grow by 2.75%.

   a. What variables are used to define the Job category?

   b. What is the job rating for Pittsburgh?

   c. How do you think projections are made about population growth? About job growth?

5. The college enrollment in Austin, Texas, is summarized in Table 3.1:

Table 3.1
Austin College Enrollment

College/University	Kind	Enrollment
Two-year	Public	23,067
Southwest Texas State	Public	20,800
University of Texas, Austin	Public	50,245
Concordia Lutheran College	Private	603
Houston-Tillotson College	Private	695
Saint Edward's University	Private	2,964
Southwestern University	Private	1,239

Source: *Places Rated Almanac*, 1994

a. What variables are used to define the Education category?

b. Find the education rating for Austin.

c. Write a formula for education ratings in general.

d. Which would have the greater effect on this rating, an increase of 1,000 students in two-year colleges or an increase of 1,000 students in public colleges? Explain your answer.

6. Review the descriptions of the five variables defined for rating metropolitan areas.

a. For which of these categories are high ratings good and low ratings bad?

b. Which of these categories might tend to produce better ratings in larger cities?

7. Refer to the crime rating.

a. Why do you think a multiplier of $\frac{1}{10}$ is used in the formula for crime rating?

b. Does the $\frac{1}{10}$ seem reasonable? Explain your answer.

8. Think about the quality of recreation in various metropolitan areas.

a. Name a city in the United States that might have a relatively good rating for recreation. Explain your answer.

b. Name a city that might have a relatively poor rating for recreation. Explain your answer.

Understanding Ratings and Ranks

Ratings in the five categories for a selection of eight cities are provided in Table 3.2.

Table 3.2
City Ratings

City	Housing	Jobs	Education	Crime	Recreation
Boston, MA	18,903	3,456	4,176	1,051	2,278
Washington, DC	15,466	16,288	3,764	1,028	1,857
Atlanta, GA	8,676	16,777	1,692	1,474	1,822
San Diego, CA	20,322	14,772	2,335	1,266	3,800
Terra Haute, IN	4,116	2,028	290	823	1,100
Lincoln, NE	6,362	2,457	554	993	1,486
Greenville, NC	6,911	3,477	377	882	900
Salem, OR	6,226	2,787	237	869	1,784

Source: *Places Rated Almanac*, 1994

9. Review your answer to problem 6b. Do the data suggest that you were correct?

10. **a.** For which category are the ratings most variable? Explain.

b. For which are the ratings least variable? Does this seem reasonable? Explain your answer.

11. Use 1 to stand for the most desirable and 8 the least desirable to rank the cities with regard to housing cost *H*. What features of the original data are lost when you go from the actual ratings to the ranks?

12. Now consider the other categories.

a. Rank the cities within each of the other four categories, with 1 as the most desirable. Write the ranks on the table next to the ratings. Compare the rankings across the categories. What observations can you make?

b. What are some advantages to reporting the ranks rather than the ratings?

13. Consider the ratings for housing cost.

a. Suppose a person is moving from Washington, DC, to San Diego. What is the percentage change in housing cost this person can expect to pay?

b. Suppose a person is moving from Washington, DC, to Greenville. What is the percent change in housing cost this person can expect to pay?

c. Write an expression for the percentage change that a person can expect when moving from City A to City B.

Combining Ranks

14. Think about the formulas you have studied.

a. You want to find an overall rating of the cities by combining the five categories. How might you combine the ranks across the categories?

b. Do you think that all of the categories are of equal importance? What category would be more important to your grandparents? To your family? To you?

One of the ways that you can count one category more than another is to assign a weight to each category in order of importance.

On Table 3.2, divide ten pennies among the five categories according to the relative importance you assign to the categories. Actually place the pennies on the table. For example, if you think that all five categories are equally important, place two pennies on each column. If you think housing and jobs are equally important and nothing else matters, then place five pennies on housing and five on jobs. The number of pennies assigned to a category divided by 10 (which makes a convenient decimal) is your personal weighting for that category.

15. Find a rank by using the weights you have established.

a. Use your weights to write a score for Boston that is the weighted average of the ranks for the five categories.

b. Write a formula for the score for each city that is the weighted average of the ranks for the five categories.

c. Use your calculator or computer and the formula you wrote to find your weighted rank for each city.

16. Consider all the cities in Table 3.2.

a. Rank the cities according to your weighted ranks from Problem 15.

b. Compare the overall ranking with the ranking within each category. Which categories have similar rankings to your overall ranking?

Summary

Comparisons among objects can be based upon a number of different variables. A rating assigns a numerical outcome to each variable that reflects the worth of the object according to that variable. A ranking is the result of placing the outcomes for one variable in numerical order. When combining variables to form a single ranking, different weights may be used on the respective variables.

Practice and Applications

17. An article entitled "The Best States for High School Football" appeared in the August 16, 1994, *USA Today.* The article investigated the history of high-school football teams.

- **a.** List some variables that might be factors in rating states on high-school football.
- **b.** One of the characteristics reported in the article was the number of players that go on from high school to play football in college and in the National Football League. Use the data in Table 3.3 to create a rating for each state. Use your ratings to rank the states.

Table 3.3
Rating High-School Football

State	NFL	College	Rating	Ranking
Alabama	38	70		
California	169	272		
Florida	106	316		
Georgia	71	46		
Illinois	38	61		
Maryland	23	27		
Michigan	35	55		
Texas	125	218		
Virginia	43	69		

Source: *USA Today,* August 16, 1994

18. Think of another way that a composite ranking of the cities in previous problems could be obtained. What are the advantages and disadvantages of your new method?

19. What questions would be good to ask when someone reports a new study that ranks desirable spots for a vacation?

20. The September, 1994, issue of *Money* magazine contains data on the "top-ten" cities in the United States. The categories used and the ratings given for the top-ten cities are in Table 3.4.

Table 3.4
Rating Cities

City	Health	Crime	Economy	Housing	Education	Transportation	Climate	Leisure	Arts
Albuquerque, NM	61	13	63	94	36	42	43	23	17
Gainesville, FL	45	4	92	49	45	45	79	5	22
Provo/Orem, UT	59	58	79	61	41	58	29	37	22
Raleigh/ Durham, NC	88	20	93	88	95	41	38	4	28
Rochester, MN	96	63	71	43	97	81	14	25	26
Salt Lake City/ Ogden, UT	72	27	81	75	51	40	26	35	22
San Jose, CA	82	35	37	25	40	27	83	93	91
Seattle, WA	79	20	67	28	62	28	47	94	52
Sioux Falls, SD	71	54	96	43	6	75	9	2	14
Stamford/ Norwalk, CT	86	53	74	40	14	18	28	88	100

a. Write a formula you would use to select the top-ten cities. Explain what the variables represent.

b. The final rankings given by *Money* magazine are in Table 3.5. What do you think might have been done to arrive at these rankings?

Table 3.5
Final Cities Ranking

City	Rank
Albuquerque, NM	10
Gainesville, FL	7
Provo/Orem, UT	3
Raleigh/Durham, NC	1
Rochester, MN	2
Salt Lake City/Ogden, UT	4
San Jose, CA	5
Seattle, WA	8
Sioux Falls, SD	9
Stamford/Norwalk, CT	6

Extension

21. Find a copy of the *Places Rated Almanac.*

a. What method is used to produce a composite ranking of the metropolitan areas in the United States?

b. Select cities of interest to you and find combined ratings and ranks for those cities. You may want to use different categories, according to your interests.

22. Select a topic, determine what categories are important, and the criteria for rating each category. Collect the data, find the rankings, and prepare a report on your work. Some possible categories are top-ten movies, best quarterbacks, best rock groups, best cars, and best teachers.

LESSON 4

Expressions and Rates

What variables do police study when investigating automobile accidents?

How is safety a factor in driving an automobile?

How do city officials decide whether or not to repair a road?

OBJECTIVES

Understand rates in numerical and symbolic form.

Understand the importance of units of measurement.

Combine rates by using weighted averages.

The first motor-vehicle death in the United States was reported in New York City on September 13, 1899. Since then, according to *The Universal Almanac,* 1994, more than 2,850,000 people have died in motor-vehicle accidents.

INVESTIGATE

Motor-Vehicle Safety

The United States government and state governments keep careful records on motor-vehicle travel and safety. Among other things, these data are used to help make decisions on issues such as how tax dollars are used for road construction and repairs and ways to improve safety. How do you think driving safety might be improved?

Discussion and Practice

1. Think about ways to improve motor-vehicle safety.
   a. What variables do you think should be measured to describe and improve motor-vehicle safety?
   b. How is driving safety monitored and improved in your community?

2. Go to the library to find information about motor-vehicle safety in resources such as an almanac or the *Statistical Abstract of the United States.*

   a. What variables are represented in the data they found? What information was given about motor-vehicle travel and safety?

   b. How do you think the data were collected?

Suppose you are taking a trip that involves driving in five different states. How do these states rank in terms of traffic safety? The 1994 data on three variables for each state are in Table 4.1.

R = number of motor-vehicle registrations
M = vehicle miles of travel in the state
D = motor-vehicle deaths

Table 4.1
Safe Driving

State	*R* (thousands)	*M* (billions)	*D*
Alaska	486	3.8	106
California	22,202	262.5	3,816
Florida	10,232	114.3	2,480
New York	9,780	109.9	1,800
Rhode Island	622	7.7	79

Source: *Statistical Abstract of the United States,* 1994

3. Refer to Table 4.1.

   a. How many motor vehicles were registered in Florida in 1994?

   b. How many miles were driven in California in 1994? Write the answer in millions.

   c. How did the number of deaths per registered vehicle in Rhode Island compare to that number in Alaska?

4. How do you think the measurements R, M, and D were determined?

5. You wish to rank the states from best to worst in terms of traffic safety. Can this be done on the basis of R, M, or D alone? Explain why or why not.

Consider the ratio $\frac{D}{M}$ as a measure of driving safety. In Alaska, $\frac{106}{3.8}$ = 27.9 deaths per billion miles driven. Such a measure is called a *rate.* In this case, the rate is a kind of average, or the number of occurrences for a given amount. $\frac{D}{M}$ as a rate represents the average number of deaths to be expected for each billion vehicle miles driven.

6. The following problems deal with rates and how they are used.

   a. Express $\frac{D}{M}$ for Alaska *without* using the word *billion.*

   b. What does the expression $\frac{M}{D}$ measure? Find $\frac{M}{D}$ for Alaska.

   c. Based on the data in the table, in which state is it safest to drive? How did you use the data to decide?

7. Is there any other information you would find helpful in making your decision about the safest state for driving?

What's the Average?

Phil is interested in buying a three-year-old car. The car, which has been driven 35,000 miles, has been advertised as having low mileage.

8. To determine "low mileage," you first need to consider typical or average mileage.

   a. How could you use the data in Table 4.1 to find an estimate of the typical number of miles driven per vehicle per year for all five states combined?

   b. What is a typical number of miles? Compare your method and answer to those of your classmates.

9. Think carefully about the process you used to find a typical number of miles.

   a. What is the average number of miles per vehicle for Alaska? Write an expression showing how you found your answer.

   b. What are your estimates of a typical value for the number of miles driven for each of the five states?

   c. Are your estimates in part b a good measure of the typical distance a vehicle is driven in a year for each of the states?

10. Suppose you wanted to think about a typical distance over all of the five states. Phil and Jerene each have a different way to find an average rate for the five states. This is Phil's formula:

$$\frac{\frac{M_A}{R_A} + \frac{M_C}{R_C} + \frac{M_F}{R_F} + \frac{M_N}{R_N} + \frac{M_R}{R_R}}{5}$$

a. What do you think $\frac{M_A}{R_A}$ represents?

b. What does $\dfrac{\frac{M_A}{R_A} + \frac{M_C}{R_C} + \frac{M_F}{R_F} + \frac{M_N}{R_N} + \frac{M_R}{R_R}}{5}$ represent? Find a numerical value for the expression.

c. Describe in words what Phil's method will give you. Why did Phil divide by 5?

11. Jerene used this expression:

$$\frac{M_A + M_C + M_F + M_N + M_R}{R_A + R_C + R_F + R_N + R_R}$$

a. What does $M_A + M_C + M_F + M_N + M_R$ represent? Find the numerical value of this expression.

b. Describe in words how Jerene is computing a typical number of miles driven.

c. Which expression, Phil's or Jerene's, do you think gives a better estimate for the typical number of miles driven in the five states? Why?

d. What are the units for the typical number of miles driven by both methods? Calculate the answer by the method you prefer. Change your answer so the units are miles per vehicle.

12. Should the car Phil was thinking about buying have been advertised as having low mileage? Why or why not?

Summary

A rate is a ratio of two measurements, often used to put data in a standard form for purposes of comparison. Rates are expressed as one measurement per a unit of another. Gas mileage can be measured in miles per gallon, speed in miles per hour or feet per second, and population density in people per square mile. We can use rates to compare cars traveling different distances with respect to gas mileage and speed. Rates can also be used to compare population density of countries of different sizes. Care must be taken in choosing an average rate, because there are two basically different ways to construct this average.

Practice and Applications

13. José has taken four trips to a neighboring town. He wanted to know whether he was getting good gas mileage on the car, and so he recorded the amount of gas and the number of miles for each trip. His data are in Table 4.2.

Table 4.2
Gas Mileage

Date	Miles	Gas (gallons)
6/19	126.48	5.6
6/22	230.74	8.3
6/28	115.20	4.8
7/1	188.70	7.4

a. Gas mileage is usually measured in miles per gallon. Find the gas mileage for each trip.

b. Find the gas mileage for the four trips combined.

14. In some colleges and universities, grades of students are recorded on a four-point system. The grading system takes into account the fact that courses meet for different amounts of time. A course that meets for three hours a week is worth three credit hours. In a four-point system, grades are usually awarded the following values for each credit hour:

A 4 points
B 3 points
C 2 points
D 1 point
F 0 points

If a student has a C in a three credit hour class, that grade is worth 6 points. The grade-point average of a student is actually the average number of points per credit hour. This is a weighted average in which the hours are the weights.

a. Recy got a B in a course that met for five hours a week, an A in a course that met for three hours a week, and a C in a course that met for four hours a week. What was Recy's grade-point average?

b. Recy wants to improve his grade-point average to at least 3.0. He thinks he can get an A in a summer course if he signs up for one that does not involve many hours. How many credit hours of A work does he need to meet his goal?

c. Can Recy ever bring his grade-point average up to 3.0 if he gets a B in all the courses he takes in the future? Explain your answer.

15. Do this problem only if you have a computer or a graphing calculator. The data on car safety in Table 4.1 has been expanded for all of the states in Table 4.3. Enter the data, and use the formulas you have developed to answer the questions below for all of the states. Record your results on *Activity Sheet 3*. In each case, write a paragraph explaining the method you used, why this seems to be a reasonable method, and the conclusion you reach.

Table 4.3
States and Safe Driving

State	Total Reg. Vehicles (thousands)	Vehicle Miles (billions)	Deaths
Alabama	3,304	45.8	1,001
Alaska	486	3.8	106
Arizona	2,801	35.0	810
Arkansas	1,502	23.0	587
California	22,202	262.5	3,816
Colorado	2,915	28.9	519
Connecticut	2,429	26.5	296
Delaware	545	6.9	140
Washington, DC	256	3.6	NA*
Florida	10,950	114.3	2,480
Georgia	5,899	77.9	1,323
Hawaii	774	8.0	128
Idaho	1,034	10.8	243
Illinois	7,982	87.6	1,375
Indiana	4,516	57.0	902
Iowa	2,706	23.9	437
Kansas	1,921	24.2	387
Kentucky	2,983	38.1	819
Louisiana	3,094	33.9	871
Maine	978	12.2	213
Maryland	3,689	41.9	664
Massachusetts	3,663	47.3	485
Michigan	7,311	84.2	1,295
Minnesota	3,484	41.2	581
Mississippi	1,954	26.2	604
Missouri	4,004	53.3	985
Montana	907	8.5	190
Nebraska	1,355	14.6	270
Nevada	921	10.9	251
New Hampshire	894	10.1	123
New Jersey	5,591	59.4	766
New Mexico	1,352	18.5	461
New York	9,780	109.9	1,800
North Carolina	5,307	67.5	1,262
North Dakota	655	6.1	88
Ohio	9,030	95.2	1,440
Oklahoma	2,737	35.1	619
Oregon	2,583	27.9	464
Pennsylvania	8,179	89.2	1,545
Rhode Island	622	7.7	79

Table 4.3 (continued)
States and Safe Driving

State	Total Reg. Vehicles (thousands)	Vehicle Miles (billions)	Deaths
South Carolina	2,601	35.0	807
South Dakota	702	7.2	161
Tennessee	4,645	50.0	1,155
Texas	12,697	163.3	3,057
Utah	1,252	16.3	269
Vermont	465	6.0	96
Virginia	5,239	63.4	839
Washington	4,466	49.4	651
West Virginia	1,273	16.5	420
Wisconsin	3,735	47.6	644
Wyoming	483	6.2	118

*NA = Not Available
Source: *The American Almanac*, 1995

a. In which state is it safest to drive?

b. What are the top five safest states based on the given data? What is the typical number of miles driven per vehicle in the United States?

16. **a.** What is the value of using a rate (like the number of deaths per vehicle mile) rather than an absolute number (like the number of deaths) for making comparisons?

b. How does a weighted average differ from a simple average?

Extension

17. The data in the lesson examined above suggests that Americans seem to drive a lot! Choose one or more of the following questions and use the data in Table 4.4 to answer them. Show how you used the data and expressions to answer the questions.

a. How much does a typical family drive?

b. What does driving cost the typical American family?

c. Are families cutting down on the amount of driving they do each year?

d. Is the fuel efficiency of automobiles improving?

Table 4.4
Driving Data

Year	1983	1988
Households with vehicles (millions)	73	81
Vehicles (millions)	130	148
Vehicle miles traveled (billions)	1,219	1,510
Motor fuel consumed (billion gallons)	81	82
Motor fuel expenditure (billion $)	95	81

Source: *The World Almanac and Book of Facts,* 1992

18. **a.** For which of the variables in Table 4.4 does it make sense to look at the total across the two years?

b. For which of the variables does it make sense to look at the average across the two years?

LESSON 5

Rates, Frequencies, and Percents

What vital statistics are used to determine whether a person is allowed to drive?

Do you know drivers who speed up when a traffic light turns yellow?

What statistics do car manufacturers study to determine automobile safety?

What measurable factors can be used to determine a person's cause of death?

More Teenagers Are Killed (headline)

Records on vital statistics, such as deaths, are collected and reported by the National Center for Health Statistics, an agency of the U. S. government. These records show the death rates for each year (and sometimes by months within years) in categories such as cause of death, age group, and year.

OBJECTIVES

Apply rates in practical situations.

Consider time as a variable and to look at time plots.

Change from rates to counts and percents.

INVESTIGATE

Leading Causes of Death

Data on causes of death provide researchers with an opportunity to study patterns in death rates over the years, with the possibility of using this information to pinpoint major concerns and to make decisions that will improve chances of survival.

A death rate is defined as the number of deaths for a given number of people in the population being studied. Suppose a city of 600,000 people has a motor-vehicle death rate of 35 people per 100,000. To find the actual number of people killed in motor-vehicle accidents, you can solve the following proportion:

$$\frac{35}{100{,}000} = \frac{n}{600{,}000}$$

$$100{,}000n = 35 \times 600{,}000$$

$$n = \frac{35 \times 6}{1} = 210$$

There were actually 210 deaths in the population in one year.

Discussion and Practice

1. Table 5.1 provides data from the National Center for Health Statistics based on all death certificates filed in the 50 states plus the District of Columbia. According to the National Center, the causes of death are compiled in accordance with standards set by the World Health Organization.

   a. Why do you think rates, rather than the number of deaths, are used to report data on deaths?

   b. Write a formula to find the number of deaths in a city with a given death rate.

Table 5.1
Death Rates for Leading Causes of Death for All Ages (per 100,000 people)

Cause	1979	1989	1990	1991
All Causes *AC*	852.2	871.3	863.8	858.5
Accidents *A*	46.9	38.5	37.0	36.4
Motor-Vehicle Accidents* *MV*	23.8	19.3	18.8	17.9
Suicide *S*	12.1	12.2	12.4	12.0
Homicide *H*	10.0	9.3	10.0	10.9
Heart Disease *HD*	326.5	297.3	289.5	284.8
Malignancies *M*	179.6	201.0	203.2	204.0
HIV *HIV*	—	8.9	10.1	11.7

*Motor-vehicle accidents are a subset of all accidents.
Source: *Statistical Abstract of the United States*, 1995

Interpreting Data

2. Consider the data in the table above.

   a. What does the 852.2 for all causes in 1979 represent?

   b. In 1991, did more people die from-motor vehicle accidents than from homicide?

   c. Why do you think there is a blank in the 1979 column?

3. Refer again to Table 5.1.

   a. Do the rates for specific causes of death add to the rate for all causes? Explain why or why not.

   b. Describe three trends you can observe from the table.

4. According to the Census Bureau, the approximate numbers of people in the United States for the years listed in Table 5.1 were:

   1979: 225 million
   1989: 247 million
   1990: 249 million
   1991: 252 million

   a. Approximately how many people died from all causes in 1979?

   b. How many died from accidents in 1991?

   c. How many died from motor-vehicle accidents in 1990?

5. Use a spreadsheet to do the following problems with the data in Table 5.1.

   a. Add the population figures to the table. Should they be added in a row or a column?

   b. Calculate the number of deaths for any given year for each cause. Explain what your variables represent.

6. The report from the National Center for Health Statistics states that the increases in death rates from 1990 to 1991 are led by HIV infections and homicides.

   a. Find the amount of increase and percent of increase for each rate.

   b. Was the statement on increases in the report correct? Why or why not?

Looking at Trends

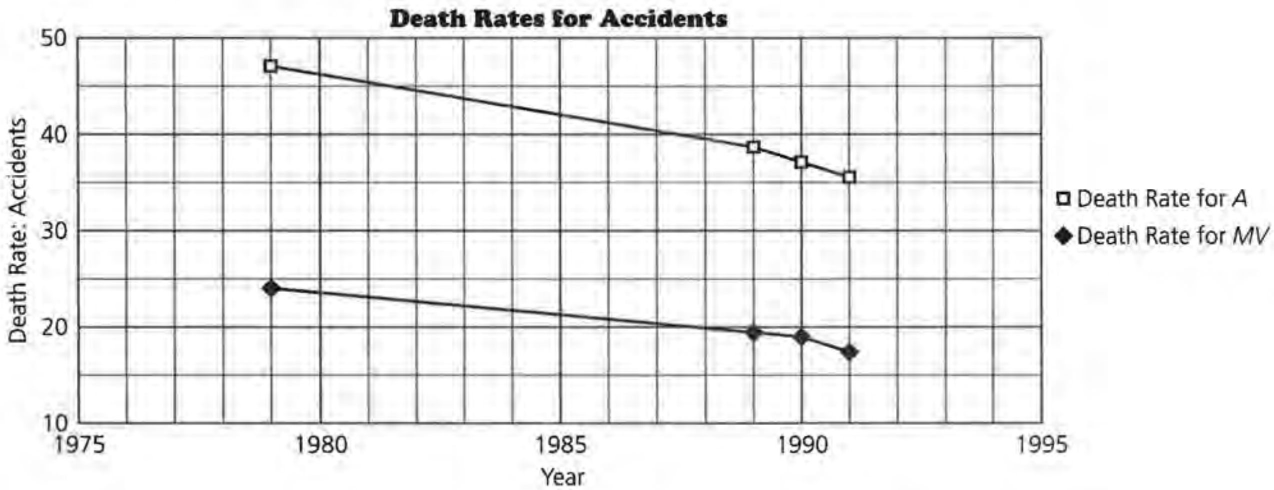
7. One important consideration in data analysis is to look for trends over time.

   a. What might an increasing trend in a death rate over time indicate?

   b. What might a decreasing trend in a death rate over time indicate?

c. Does a decreasing trend in death rates necessarily imply that there is a corresponding decreasing trend in the number of deaths? Explain your answer.

8. A plot can be used to show trends over time. The line graph, or time plot, shown below gives the death rates for all accidents (*A*) and for motor-vehicle (*MV*) accidents using the data from Table 5.1.



a. Will the graph for motor-vehicle accidents ever be above the graph for accidents? Why or why not?

b. Describe the trends in the plots.

c. Construct line graphs similar to those in the plots above for death rates from heart disease and from malignancies. Use different symbols or colors for each variable. Comment on any trends across time that you observe for these graphs.

Death Rates for the Age Group 15–24

The death rates for residents of the United States are computed separately for different age groups. Death rates per 100,000 people for the causes discussed in Table 5.1 are presented in Table 5.2 for the age group 15–24.

Table 5.2
Death Rates for the Age Group 15–24
(per 100,000 people aged 15–24)

Cause	1979	1989	1990	1991
All Causes *AC*	114.8	97.6	99.2	100.1
Accidents *A*	62.6	44.8	43.9	42.0
Motor-Vehicle Accidents* *MV*	45.6	34.6	34.1	32.0
Suicide *S*	12.4	13.0	13.2	13.1
Homicide *H*	14.5	16.5	19.9	22.4
Heart Disease *HD*	2.6	2.5	2.5	2.7
Malignancies *M*	6.1	5.0	4.9	5.0
HIV *HIV*	—	1.6	1.5	1.7

*Motor-vehicle accidents are a subset of all accidents.
Source: *Statistical Abstract of the United States,* 1995

Interpreting the Data

9. Refer to Table 5.2. Write the meaning of each phrase.

a. 114.8 in 1979

b. 34.1 in 1990

c. 1.7 in 1991

10. Consider your answers to problem 9.

a. What information is needed in order to transform these rates into a number of cases?

b. Compare motor-vehicle death rate for 15- to 24-year-olds to that for the population as a whole.

Looking for Trends

11. You should have discovered that the death rates for motor-vehicle accidents among 15- to 24-year-olds are much greater than those for all ages from Table 5.1.

a. Sketch plots for people ages 15–24 using the same categories as in problem 8.

b. According to your plots for the 15–24 age group, which is decreasing faster, death rates due to accidents or death rates due to motor-vehicle accidents?

12. Compare the trend over time in suicide rates to the trend over time in homicide rates for 15- to 24-year-olds. Comment on the comparison.

13. Comment on any other over-time trends that appear to be interesting.

Changing from Rates to Fractions and Percents

14. Sometimes it is useful to look at one cause of death as a fraction of another.

- **a.** Use Table 5.1 to tell what fraction of all accidental deaths were due to motor-vehicle accidents in 1979? Explain how you found your answer.
- **b.** How consistent has this fraction been over the years?

15. If D_M is the number of deaths from motor-vehicle accidents and D_A the number of deaths from all accidents, can the fraction $\frac{D_M}{D_A}$ ever be greater than 1? Explain why or why not.

16. Change the fraction you found in problem 14 to a percent. What does this percent represent?

17. Look again at Table 5.2.

- **a.** Write an expression for D_A, the number of deaths due to accidents for that age group, using P for the population size for 15- to 24-year-olds in a given year and A for the death rate due to accidents in that year.
- **b.** What does $\frac{A \cdot P}{100{,}000}$ represent?
- **c.** Explain what the equation $\frac{\frac{MV \cdot P}{100{,}000}}{\frac{A \cdot P}{100{,}000}} = \frac{MV}{A}$ represents.
- **d.** Is it necessary to know population size to compare rates as in part c?

18. Use your results from problem 17.

- **a.** Calculate the percent of all accidental deaths that are due to motor vehicles for each year in Table 5.2.
- **b.** Plot these percents over time and observe the trend.

c. Compare this trend to the ones observed for the rates in problem 11. Summarize your observations.

19. Refer again to the date for the 15–24 age group.

  a. List several questions that cannot be answered from these data.

  b. Summarize the main results of your study of the data for this age group.

Summary

Rates are used to compare data sets of different sizes. Death rates are often given as a number out of 100,000, although it could be a rate per 1,000 or some other number. You can use rates to understand how frequency changes as population changes. A plot over time can be useful in understanding trends. Rates calculated from the same base can be compared directly; percents are useful in making such comparisons.

Practice and Applications

20. What information is needed to transform a rate into the number of cases?

21. For either Table 5.1 or Table 5.2, suppose you want to reduce the number of categories. To do this, combine the heart-disease and malignancy rates into a single rate for "disease." Can this be done by simply adding the rates in the two rows? Use symbols to help you justify your answer.

Extension

22. Deaths from HIV infections are of particular interest now. Table 5.3 expands the HIV death-rate data given in Table 5.1 to more age categories.

  a. How is Table 5.3 different from Table 5.1?

  b. Use the data from Table 5.3 in conjunction with the two Tables 5.1 and 5.2 to write a summary of important comparisons and trends involving HIV infections. The summary should contain at least one graph. Why is this information important?

c. Is it possible to combine age categories in the table below by simply adding the rates for categories to be combined? That is, is the 1991 rate for 15- to 34-year-olds equal to 1.7 + 22.1? Use symbols to justify your answer.

Table 5.3
Death Rates from HIV Infections (per 100,000 people)

Year	1989	1990	1991
15–24 years	1.6	1.5	1.7
25–34 years	17.9	19.7	22.1
35–44 years	23.5	27.4	31.2
45–54 years	13.3	15.2	18.4
55–64 years	5.4	6.2	7.4

Source: *Statistical Abstract of the United States,* 1995

LESSON 6

Formulas That Summarize Typical Values in Data

What is the typical age of students in your class?

How long does it take you to get to school?

How many people live in your house?

Think about the ways numbers describe your life and the lives of those around you. Most people have about 13 years of formal schooling. Most workers are on the job for about 40 years before they retire. The typical American family has four people and often owns two cars. In Columbus, Ohio, on a typical summer day, the temperature will be around 80° Fahrenheit and on a typical winter day around 35°. A winning football team may score around 21 points, a winning basketball team around 90 points, and a winning baseball team around 9 runs. A typical newborn baby weighs around 8 pounds. These are examples of the many "typical" values by which we make judgments every day. What are some of the typical moral or social values that are important to you and your family?

OBJECTIVES

Use formulas to summarize typical values.

Use symbolic notation as an efficient method of communication.

INVESTIGATE

Summary Numbers

Many animals are in danger of extinction. To see the extent of the problem, study the numbers of animals on an endangered species list compiled by the Fish and Wildlife Service of the United States Department of the Interior as of June 3, 1993. The data in Table 6.1 show the numbers of endangered species for various groups of animals. The count is the total number of endangered species within the group. There are 37 different kinds of mammals on the endangered species list within just the United States, 19 others in both the United States and foreign countries, and 249 others in the rest of the world excluding the United States.

Discussion and Practice

Table 6.1
Endangered Species

Group	U.S. Only *US*	U.S. and Foreign *USF*	Foreign Only *F*
Mammals	37	19	249
Birds	57	16	153
Reptiles	8	8	64
Amphibians	6	0	8
Fishes	55	3	11
Snails	12	0	1
Clams	50	0	2
Crustaceans	10	0	0
Insects	13	2	4
Arachnids	3	0	0

Source: *The World Almanac and Book of Facts,* 1994

1. Refer to Table 6.1.

   a. How many birds are on the endangered list altogether?

   b. Which group of animals is most endangered? How did you decide?

2. Study the data in Table 6.1.

   a. Use a single number to summarize the data in the U.S. Only column. What number did you use? Why?

   b. If you wanted to summarize the situation for endangered mammals worldwide, what number would be most meaningful? Why?

   c. Does the average of the numbers in the U.S. Only column have a useful interpretation? Explain your answer.

   d. Write a paragraph summarizing the information in the table. Include the numbers you chose in parts a–c.

The Greek capital letter sigma, Σ, is commonly used when dealing with totals.

$\sum_{i=1}^{9} x_i$ tells you to add the values of the variable x_i from the value of i below Σ (1, in this case) to the value of i above the symbol (9, in this case).

Table 6.2 contains the number of *threatened* species. Threatened species are those in which the population size indicates a need for concern but which is not yet small enough to be in danger of extinction.

Table 6.2
Threatened Species

Group	U.S. Only *US*	U.S. and Foreign *USF*	Foreign Only *F*
1. Mammals	6	3	22
2. Birds	8	9	0
3. Reptiles	14	4	14
4. Amphibians	4	1	0
5. Fish	31	6	0
6. Snails	7	0	0
7. Clams	5	0	0
8. Crustaceans	2	0	0
9. Insects	9	0	0
10. Arachnids	0	0	0

Source: *The World Almanac and Book of Facts*, 1994

Note that the animal groups in the United States are labeled *US*, those in both the United States and foreign countries *USF*, and those in foreign countries *F*. The rows in the table are numbered, and these numbers can help you locate an individual animal group without writing the entire name. Together the letter and the number indicate the group and category. Thus, US_1 represents the number of United States threatened species in the first row, mammals. $US_1 = 6$. US_5 represents the number of threatened species in the United States in the animal group in the fifth row, fish. $US_5 = 31$. The number below *US* is called a *subscript*. The total number of threatened species for the entire United States can be labeled as US_T, where T indicates total, and expressed symbolically as:

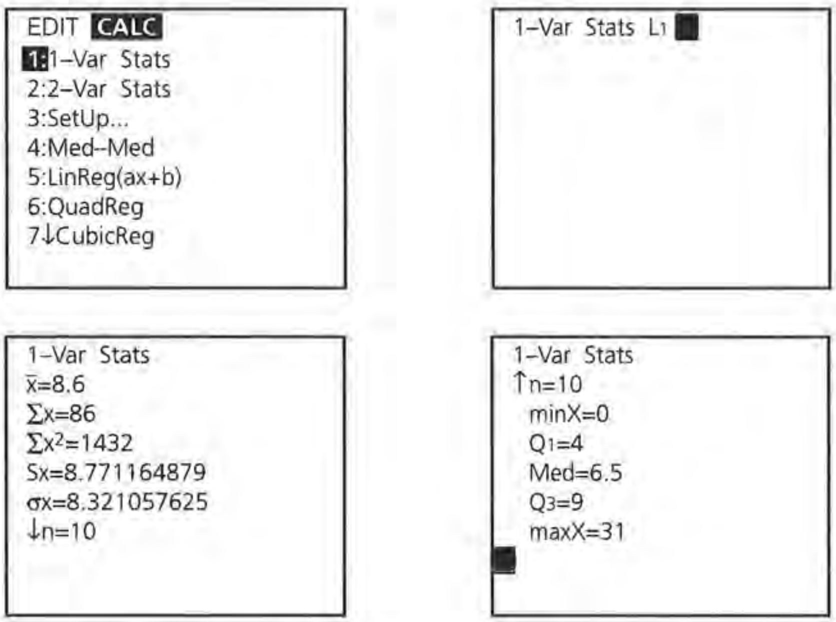
$$US_T = \sum_{i=1}^{10} US_i$$

3. a. Find the value of US_T.

   b. What does F_T represent? Find its value.

   c. What does $\sum_{i=1}^{6} USF_i$ represent? Find its value.

4. You can find summary numbers using a graphing calculator. Enter the numbers of threatened species in the United States *US* in one list, those from both the United States and foreign countries *USF* in a second list, and those from foreign countries only *F* in a third list. Use the statistics calculations menu to find the one-variable statistics for a given list. The diagrams show this for a TI-82.



Note that $\bar{x}$ is the mean of the data in List 1, Σx is the sum of the values in List 1, and n is the number of the values in List 1. If you use the down arrow, you will also find the quartiles and the median.

   a. If x represents the values in List 1, what does Σx^2 represent? Does Σx^2 make sense for the data?

   b. How do $(\Sigma x)^2$ and Σx^2 differ?

   c. Use your calculators to find $\Sigma US + \Sigma USF$. What does this sum represent?

   d. Use your calculator to find ΣUSF and ΣF.

Sometimes it is very important to be able to summarize a set of data using summary numbers. The total, the mean, and the median are examples of different summary numbers you can use. The context of the problem can help you decide which is appropriate. In some cases, more than one of them might be useful in helping you understand the data. In other cases, some of them are clearly inappropriate.

5. Suppose you have to describe different brands of sports drinks in terms of cost and calories. The data in Table 6.3 show this information for the leading sports drinks. One 8-fluid-ounce serving is the basic unit for both calories and cost.

Table 6.3
Sports Drinks

Brand	Calories per 8-oz. Serving *C*	Cost in Dollars per Serving *D*
1. 10-K	60	.22
2. All Sport	70	.24
3. Daily's 1st Ade	60	.26
4. Exceed	70	.34
5. Gatorade	50	.26
6. Hydra Fuel	66	.52
7. Nautilus Plus	60	.22
8. PowerAde	67	.24
9. Snapple Snap-Up	80	.35

Source: *Consumer Reports,* August 1993

a. What is a good summary number for calories per serving for these drinks? For cost per serving for these drinks?

b. Does the total of the calories column provide useful information? What about the total of the cost column?

c. Suppose Hydra Fuel is eliminated from the list. What impact does that have on the mean number of calories per serving? What impact does that have on the mean cost per serving?

d. Which drink has greatest influence on the mean number of calories per serving? Give a reason for your choice.

6. Some of the sports drinks come in powdered form. Exceed powder comes in a 32-serving size for $9.43 and has 70 calories per serving. Gatorade powder has a 32-serving size for $3.59 and has 60 calories per serving. Suppose the powdered drinks are put on the same list as the liquid drinks. Is it fair to include their values in the means? Why or why not?

7. Some of the drinks come in light versions. All Sport Lite has 2 calories per serving, at a cost of 24 cents per serving. Gatorade Light has 25 calories per serving, at a cost of 26 cents per serving.

   a. What effect will adding the light varieties to the list have on the mean cost per serving? On the mean number of calories per serving?

   b. Find the medians for the number of calories and for the cost per serving from Table 6.3.

   c. Add data for the light versions to the table. How does the new median compare to the new mean in each case?

Summary Numbers and Symbols

Symbols can help you communicate mathematical operations and generalize results. The arithmetic average, or mean, of the calorie values for the nine drinks on the table can be symbolized as:

$$\bar{C} = \frac{1}{9}\left(\sum_{i=1}^{9} C_i\right)$$

8. Refer again to Table 6.3.

   a. Use Σ notation to write symbolic expressions for the total and the average of the cost per serving for the drinks listed in the table.

   b. What would $\sum_{i=1}^{8} D_i$ give you?

   c. If you add Exceed and Gatorade powder to the table, how do you have to change the notation $\frac{1}{9}\left(\sum_{i=1}^{9} D_i\right)$ to find the mean of all 11 sports drinks? How can you find this new mean quickly?

   d. Which of the sports drinks has the greatest effect on the mean cost? Explain your answer.

9. Write a paragraph summarizing the cost per serving of the sports drinks by using the total, the mean, and the median.

Be sure to think about each of the values in terms of the data and what the summary numbers will convey or misrepresent to someone who might see only that number.

10. Sometimes the Σ notation can be used with shortcuts to find results.

$$\sum_{i=1}^{2} 3x_i = 3x_1 + 3x_2$$

a. Is the following true or false for every set of data? Would it make a difference if the 3 were a 4? Explain your answer.

$$\sum_{i=1}^{n} 3x_i = 3\sum_{i=1}^{n} x_i$$

b. Suppose there were a 5% tax on the drink prices. Describe the difference between ($0.05\ D_1 + 0.05\ D_2 + 0.05\ D_3 + \ldots + 0.05\ D_9$) and $0.05\ (D_1 + D_2 + D_3 + \ldots + D_9)$. Remember that D_1 is the cost of the first drink, 10-K.

11. When data points are repeated in a data set, the results are often given in a frequency table like Table 6.4. For example, a quiz given to a class was graded on a four-point scale (0, 1, 2, 3) with 3 a perfect score. Here are the results:

Table 6.4
Quiz Scores

Score X	Frequency F	Proportion P
3	16	0.64
2	4	0.16
1	2	0.08
0	3	0.12

a. How many students were in the class? How many earned a score of 2?

b. What does the 0.64 represent?

c. What was the mean score on the quiz? Write a formula showing how to find it.

d. Use a formula to show that the average score can be calculated using the proportions in the column.

12. Tong found answers to problem 11 using a graphing calculator. He entered the scores in List 1 and the frequencies in List 2. He defined List 3 as L1 × L2.

L1	L2	L3
3	16	-------
2	4	
1	2	
0	3	
-------	-------	
L3 = L1*L2		

He found the mean by selecting 1-VAR STATS L3. Reproduce his work in your calculator. Was his procedure correct? Explain your answer.

Comparing Data

Often you need a typical or summary value for a variable in order to make comparisons. You might, in part, base your career choice on the typical salary paid in various jobs. You probably choose the stores in which you buy clothes at least partially on the basis of costs.

13. Consider the price and gasoline mileage for certain small and mid-size cars. The number of miles per gallon is figured on a mixture of city, country, and expressway driving. The size variable is coded *S* for a small car and *M* for a mid-size car.

Table 6.5
Car Prices and MPG

Model	Price Range in Dollars *D*	Miles per Gallon *G*	Size
1. Geo Prizm	10,730–11,500	33	S
2. Saturn	9,995–13,395	27	S
3. Honda Civic	9,400–16,490	29	S
4. Toyota Corolla	12,098–16,328	30	S
5. Subaru Impreza	11,200–19,100	29	S
6. Nissan Sentra	10,199–14,819	28	S
7. Ford Escort	9,135–12,400	27	S
8. Dodge Colt	9,120–12,181	34	S
9. Mitsubishi Mirage	8,989–14,529	34	S
10. Hyundai Elantra	9,799–11,924	25	S
11. Toyota Camry	16,428–23,978	24	M
12. Honda Accord	14,130–21,550	26	M
13. Ford Taurus	16,240–20,500	20	M
14. Nissan Maxima	22,429–23,529	21	M
15. Pontiac Grand Prix	16,254–18,375	19	M
16. Buick Regal	18,324–20,624	20	M
17. Oldsmobile Cutlass	16,670–25,470	20	M

Source: *Consumer Reports,* April 1994

a. Write a formula that might have been used to find the miles-per-gallon figure in the table.

b. Find $\sum_{i=1}^{10} G_i$. What does this represent?

c. Find $\bar{G}$ for both small and mid-size cars.

d. Write a statement summarizing the change in gas mileage you could expect if you decided to buy a mid-size car to replace your small car.

14. Since a range of prices is given for each car, price comparisons are more difficult to make; but such comparisons are important if you are the buyer. Find a way to summarize price comparisons among the small cars and among the mid-size cars.

Summary Numbers and Percents

15. What is the mean family size in the United States? The United States Bureau of the Census defines a family as two or more persons related by birth, marriage, or adoption who reside together in a household. Family sizes, according to the 1990 census, are in Table 6.6.

Table 6.6
Family Size

Number of Persons *N*	Percent of Families *P*
2	42
3	23
4	21
5	9
6	3
7 or more *	2

(* The percent of families with more than 7 members is very small.)
Source: *Statistical Abstract of the United States,* 1994

a. Write a sentence describing what the 2 and the 42 in the first row represent.

b. If you selected a random sample of 2,000 households, in how many of those would you expect to have four people in the family?

c. Take a survey of your class and see how the data on persons per family from your class compare to the data from the *Statistical Abstract.* What might explain any differences?

16. Refer to the data in Table 6.6.

a. Find the mean family size.

b. The mean family size calculated above is really an approximation. Why is this so?

c. Will the approximate value be too high, too low, or close to the true value? Explain your answer.

Summary

In this lesson, you learned how to summarize data with a single typical value and to write a formula for the mean. The summary numbers you studied were the total, the mean, and the median. The context for the data help you decide which of these numbers make sense. You also learned to write formulas using the Σ symbol to represent addition and using subscripts with variables, such as D_i.

Practice and Applications

Track Points

17. The track coach at Old Trier High School awards points for various events to the members of his track team. The number of points is based on the athlete's performance in track meets. The data for the 1996 season are summarized in Table 6.7. Notice that some of the team are juniors and some are seniors. P_i stands for the total number of points earned by each athlete.

Table 6.7
1996 Track Points

Grade	Athlete	Points P_i
Junior	1	67
Junior	2	70
Junior	3	62
Junior	4	80
Senior	5	60
Senior	6	62
Senior	7	68
Senior	8	71
Senior	9	78
Senior	10	65
Senior	11	70
Senior	12	76

a. How many points are associated with P_3?

b. Find $P_1 + P_2 + P_3 + P_4$. Write the sum using Σ notation.

18. Find each and explain what it represents.

a. $\sum_{i=1}^{4} P_i$

b. $\sum_{i=5}^{12} P_i$

c. $\sum_{i=1}^{12} P_i$

d. $(\sum_{i=1}^{12} P_i) \div n$

19. The data for the 1997 track team were summarized as

$$\left(\sum_{i=1}^{12} P_i\right)/12 = 77.$$

One of the students who had earned 52 points was later disqualified. As a result, which track team had the better record? How can you tell?

20. Suppose each senior on the 1996 team had actually earned twice as many points as given in Table 6.7.

a. Think of two expressions for the new total number of points earned by the seniors. Write the expressions two ways using symbols.

b. The coach decided to give all of the juniors on the 1996 team a bonus of 10 points. Which of the following expressions represents this situation? Explain your answer.

$$\sum_{i=5}^{12} (P_i + 10) \text{ or } \sum_{i=5}^{12} (P_i) + 10$$

21. Suppose every member of the 1996 team had earned exactly 70 points.

a. How could this simplify the calculation of the team total

$\sum_{i=1}^{n} P_i$ for any number n of students?

b. Develop a general rule for finding $\sum_{i=1}^{n} c_i$ for any constant c.

22. Write three rules for using the Σ symbol.

LESSON 7

Formulas That Summarize Variation in Data

What is your life expectancy?

Is your life expectancy longer than that of your parents'?

Is it longer than someone your age of the opposite sex?

What factors might have an impact on life expectancy?

Do people of some countries have much lower life expectancies than people of other countries?

Life expectancy is the length of time people born in a given year can expect to live. Suppose you knew that the average life expectancies E for countries in the Americas were between 46 to 78 years for those born in 1993. Sketch a possible histogram of the life span for people in countries of North, Central, and South America. If you know that the median life expectancy of people in countries in the Americas is 70.5 years, how will this change your sketch? Suppose you know that people in half of the countries have life expectancies between 65 and 73 years. Make a new sketch that might represent the distribution.

OBJECTIVES

Use formulas to describe the spread or variability in a set of data.

Use inequalities to describe unusual values.

Compare two formulas graphically and numerically.

INVESTIGATE

Variability and the Median

The data (the life expectancies of people born in 1993 in selected countries in the Americas and Europe) for creating the sketches in the previous paragraph are given in Table 7.1. You can see the variability in the life expectancies by making a number-line plot. The interval that contains the middle half of the data marks the *interquartile range* or IQR. The interquartile range is the distance between the first and third *quartile,* that is, $IQR = Q_3 - Q_1$ where Q_3 is the third quartile, and Q_1 the first quartile.

Discussion and Practice

Table 7.1
Life Expectancy

European Countries	Life Expectancy	Americas	Life Expectancy
Austria	76	United States	76
Belgium	77	Canada	78
Bulgaria	73	Argentina	71
Czech Republic	76	Bolivia	63
Finland	76	Brazil	63
France	78	Chile	74
Germany	76	Colombia	72
Greece	78	Cuba	77
Hungary	71	Dominican Republic	68
Italy	77	Ecuador	70
Netherlands	78	El Salvador	66
Poland	72	Guatemala	64
Portugal	75	Haiti	46
Romania	71	Honduras	67
Spain	78	Mexico	73
Sweden	78	Paraguay	73
Switzerland	78	Peru	65
United Kingdom	76	Venezuela	73

Source: *Statistical Abstract of the United States,* 1994

1. Refer to Table 7.1.

a. Make a number-line plot of the life expectancies *E* for the Americas. Explain how to find Q_3 and Q_1 and why this interval gives you the middle half of the data.

b. Mark Q_1 and Q_3 on the plot of the life expectancies and determine the IQR. What countries have life expectancies greater than Q_3? What fraction of all the countries is this?

c. Do any countries seem to have an exceptionally short or long life expectancies? Which ones?

d. Q_2 is the second quartile, or median. Where is Q_2 on your plot?

In the data on life expectancies for the Americas, it is clear that one country stands out as unusual. Do you think there are any other unusually high or low expectancies in the Americas? It is not always visually clear, however, that a value is unusually large or small compared to other data values. A mathematical formula for determining an unusual value, or *outlier*, would be helpful in those cases.

2. Try to find outliers in the data in Table 7.1.

a. Create a rule that could be used to determine whether a country would be an outlier. Is Haiti an outlier according to your rule?

b. Why is it important to have such a rule?

c. Give an example of a situation where you would like to be an outlier and an example where you would not like to be an outlier.

3. Suppose x is a data point. If x satisfies the rule below, x can be called an outlier.

$x \leq Q_1 - (1.5)(Q_3 - Q_1)$

Use Q_1 and Q_3 and this rule to determine whether the life expectancy in Guatemala is an outlier.

4. Refer again to the date in Table 7.1.

a. Draw a number line and shade in the region specified by the rule in problem 3.

b. How does the rule in problem 3 compare to the rule $x \leq Q_1 - (1.5)$ IQR?

c. Show that the life expectancy of Haiti is an outlier by the IQR rule.

d. Do the life expectancies for any other countries qualify as outliers by this rule?

5. Two rules used together can give you a formula for all outliers. x is an outlier if

$$\underbrace{x \le Q_1 - (1.5)(Q_3 - Q_1)}_{\text{lower boundary}} \text{ or } \underbrace{x \ge Q_3 + (1.5)(Q_3 - Q_1)}_{\text{upper boundary}}.$$

a. Describe in your own words how you can tell if a data point is an outlier.

b. Shade in the regions described by the rules on your number-line plot. Could you replace the *or* between the two parts of the rule with an *and*? Why or why not?

c. Describe the set of data points that are not outliers using *and*.

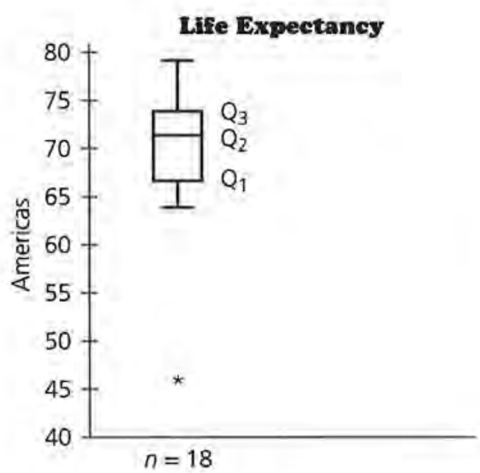
d. Is the distance between the least data point and the lower boundary described by the rule equal to the distance between the greatest data point and the upper boundary described by the rule? Explain your answer.

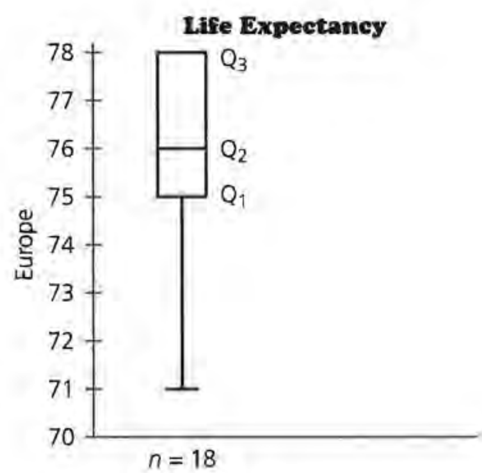
6. Jessica devised this rule for an outlier for small values: x is an outlier if $x \le 2.5\,(Q_1) - 1.5\,(Q_3)$.

a. How does Jessica's rule compare to the rule for small values in problem 5?

b. Find another way to write the rule for large outliers.

7. Box plots conveniently display the lower extreme, Q_1, Q_2, Q_3, and the upper extreme on a real-number line. Box plots for the life-expectancy data are shown in the figures below. The outlier is marked with a star, and the whisker extends only to to the next smaller data point.

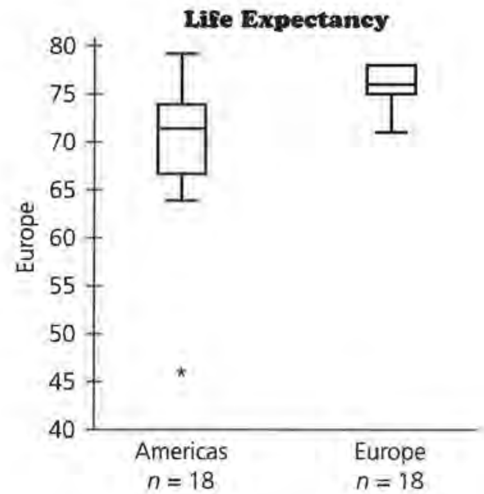




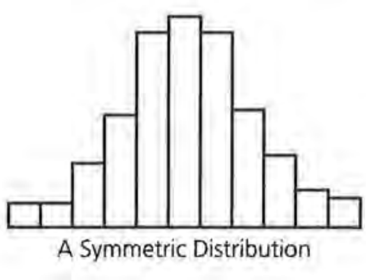
a. Describe the differences between the two box plots.

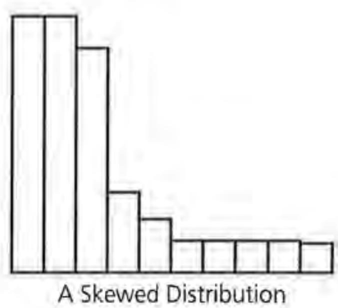
b. The box plot for Europe has no whisker on the upper end. What feature of the data set has caused this to happen?

c. The two sets of data are plotted on the same grid below. Does your impression of the differences change? Explain your answer.



Distributions are often described as *symmetric* or *skewed.* When displayed as a histogram, a skewed distribution has a long tail pointing in the direction of its skewness.





8. The life expectancies for the Americas, with Haiti included, are skewed toward the smaller values.

a. What happens to the skewness in the data for the Americas when Haiti is removed?

b. Is the distribution of the European life expectancies skewed? If, so, in which direction?

9. Based on the box-plot displays and the calculations made earlier, write a description of American and European life expectancies. Include a description of the key features of each data set as well as a discussion comparing the two groups.

Variability and the Mean

The interquartile range is a quartile-based measure of *variability*. There is also a mean-based measure for variability. Consider the following problem.

10. What are the dimensions of the desktops in your classroom? Independently, without talking or consulting with anyone else, measure the length and width of your desk to the nearest millimeter. Record the desk measurements for each group member in a chart like the one shown below.

Table 7.2
Dimensions of Desktops

Student	Length	Width
____________	____________	____________
____________	____________	____________
____________	____________	____________
____________	____________	____________

a. How do the desk measurements compare?

b. Use the measurements in the chart to describe the area of a desktop. How did you find the area? How reliable do you think your answer is?

11. As you may have recognized, it is difficult to find one answer for the area with so many different measurements. It is often useful to try to summarize a set of data by looking at key features of the data. Two questions to consider are:

- Where do the data seem to *center?*
- How do the data spread out to either side of the center?

a. Why are the center and spread key features of the data?

b. What do you lose by using only two key features to describe a set of data?

c. Why is it helpful to be able to describe a set of data with two numbers?

d. Would it be enough to describe the spread by using the greatest and least values? Explain your answer by giving an example.

You learned earlier in this lesson to summarize data by using the median and IQR. It is also possible to summarize data using the mean and a measure of spread called the *standard deviation*. This method involves finding the distance between a data point and the mean. You can find the standard deviation by using a spreadsheet where A_i denotes an individual area measurement, $\bar{A}$ is the mean, and n is the number of areas you are investigating.

Area A_i	$A_i - \bar{A}$	$(A_i - \bar{A})^2$	$\frac{(A_i - \bar{A})^2}{n}$
		$\Sigma(A_i - \bar{A})^2$	$\frac{\Sigma(A_i - \bar{A})^2}{n}$

$$\text{Standard deviation} = \sqrt{\frac{(\Sigma(A_i - \bar{A})^2)}{n}}$$

12. Refer to the table above.

a. Explain each expression.

b. Apply the process to find the standard deviation for the areas computed by your group. Explain in words what this number represents.

c. Compare your standard deviation to those found by other groups. What observations can you make about the variability for each group?

You can find the standard deviation on your calculator by selecting the statistics calculation menu and calculating the one-variable statistics for the list containing your data. The standard deviation will be the value for s_x.

13. Enter the area found by every member of your class into your calculator and find the standard deviation.

a. Write a sentence using standard deviation to describe the variability in the areas calculated by the class.

b. Sketch a histogram of the areas found by the class. Describe the variability in area you can see from the plot.

c. Add the standard deviation to the mean $\overline{x}$, and mark the result on a sketch of the histogram. Subtract the standard deviation from the mean, $\overline{x}$, and mark the result on a sketch of the histogram. Estimate the percent of the areas A that satisfy the following inequality:

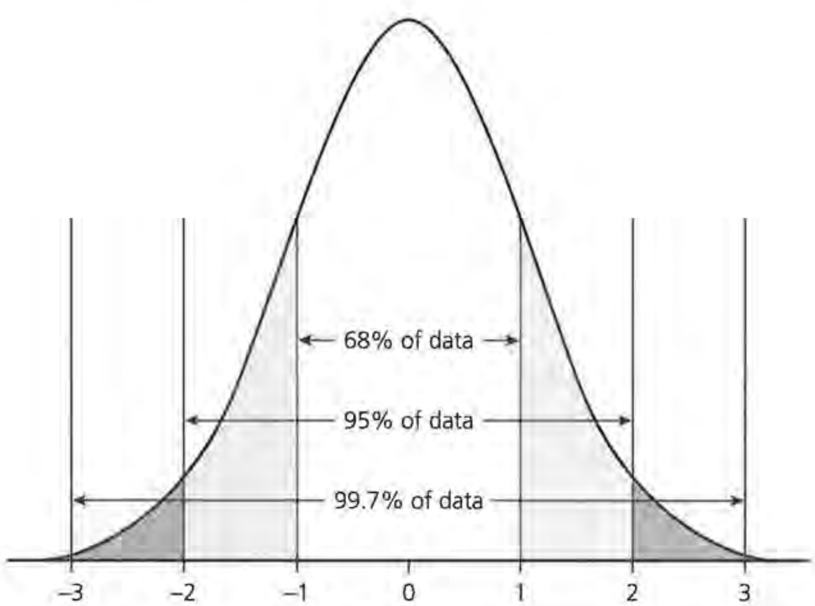
$$\overline{x} - s_x \le A \le \overline{x} + s_x$$

d. Graph the following interval on your histogram and describe in words what the interval represents:

$$\overline{x} - 2s_x \le A \le \overline{x} + 2s_x$$

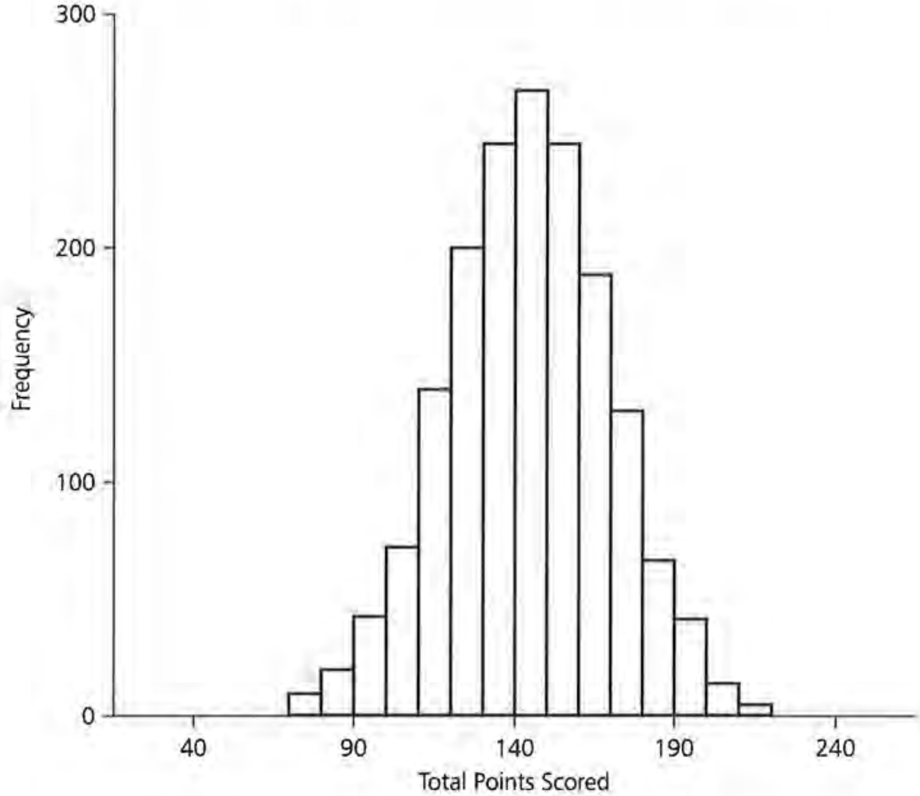
14. Would you rather describe the center and variability of your desk-measurement data in terms of the median and interquartile range or the mean and standard deviation? Explain your answer.

Unusual measurements described as outliers were found using the median and the interquartile range. There is also a technique for determining unusual measurements based on the standard deviation. If a data distribution is mound-shaped, and symmetric with a single mode, you expect to find that approximately 70% of the data lies between $\overline{x} - s_x$ and $\overline{x} + s_x$ and approximately 95% of the data lies between $\overline{x} - 2s_x$ and $\overline{x} + 2s_x$.



Under these conditions, a data point more than two standard deviations (sd) from the mean is sometimes considered an outlier.

15. The following histogram is based upon the total points scored in all the NCAA basketball playoff games from 1939 to 1994. The mean is 144, and the standard deviation is 26.



a. Look at the sketch of the distribution on *Activity Sheet 4*. Color the bars that lie within one standard deviation from the mean. Estimate the percent of scores represented by the bars that are colored.

b. Color the bars that lie within two standard deviations from the mean. Estimate the percent of scores represented by the bars that are colored.

c. Does the 70%–95% rule work well for these data?

d. Would you expect a total score to top 200 very often? What about a total score of 250?

Summary

In this lesson you learned about measures of variability. It is important to understand how things vary and whether the variability is extremely large for any of the data points. There are many ways to measure variability; range, interquartile range and standard deviation are three of the ways. A data point can be considered an outlier if it is beyond 1.5 (IQR) from either quartile and, for some distributions, if it is more than two standard deviations from the mean.

Note: If the data are from a sample of size n, the divisor used to find the average squared difference is $n - 1$ instead of n. In this case the standard deviation is labeled S_x. In this unit, use n as a divisor and s_x as the standard deviation.

Practice and Applications

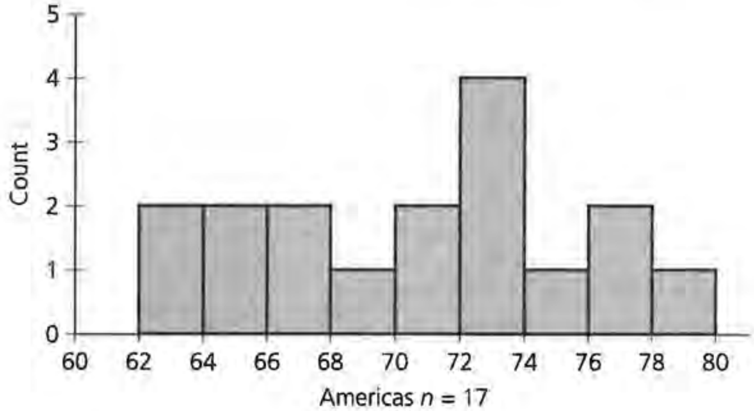
16. Recall the set of data about the life expectancies in the Americas and selected European cities.

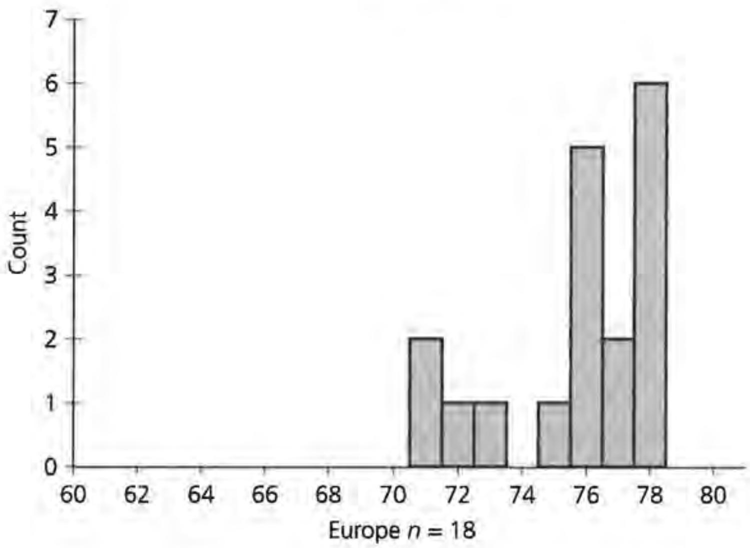
- **a.** Find the mean life expectancy for the Americas. Compare it to the median.
- **b.** Find the mean life expectancy for the Americas without Haiti. How does this mean compare to the one found in part a?
- **c.** What is the average of the deviations from the mean?
- **d.** Ignoring the sign of the deviations, what value would you choose as a typical deviation?
- **e.** About half of these deviations are negative. What does that tell you about the shape of the data distribution?

17. **a.** Calculate the standard deviations for the Americas without Haiti and for the European life expectancies. How do they compare?

- **b.** Do these two standard deviations allow you to compare Europe and the Americas in any way? Explain your answer.
- **c.** Suppose Haiti is added back into the data set for the Americas. Do you think the standard deviation will be greater? What is the new standard deviation?

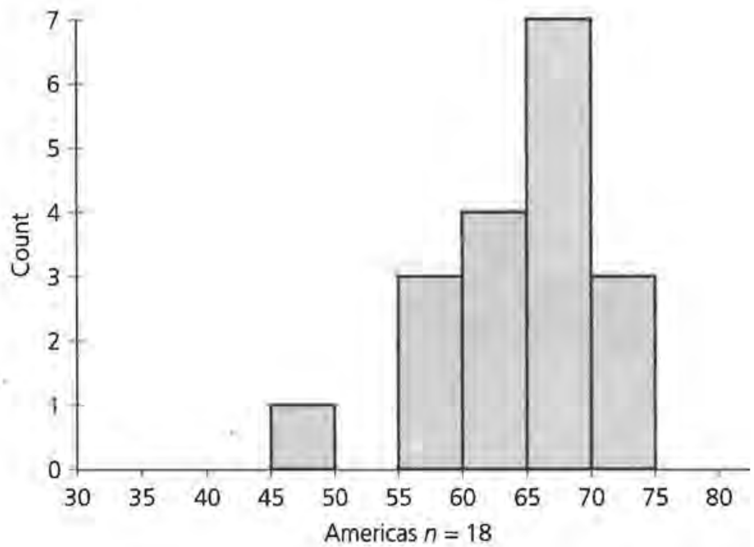
18. Histograms for the life expectancies are given below. The data for the Americas data do not contain Haiti.





On the histograms, reproduced on *Activity Sheet 4*, mark the mean for each data set. Remember that the Americas data should exclude Haiti. Then, mark a point one standard deviation above the mean and one standard deviation below the mean for each data set. For which of the two groups is the standard deviation a better measure of a typical deviation from the mean? Explain your answer.

19. A histogram of the Americas data with Haiti included is below.



Using *Activity Sheet 4,* mark the mean and the points one standard deviation above and below the mean. Remember that the data now include Haiti.

a. Does the standard deviation do a good job of describing typical deviation from the mean?

b. Based on your experience in the last few problems, do you think the standard deviation works better as measure of typical deviation from the mean for symmetric distributions or for skewed distributions? Explain your reasoning.

20. Think about all your work with standard deviation.

a. Which will have a greater standard deviation, a data set with all of its points close to each other or one with data spread out?

b. How do you think the standard deviations for each of the following two data sets compare? Make an estimate.

10, 10, 100, 100
10, 10, 10, 10, 10, 100, 100, 100, 100, 100

c. Use your calculator to check your estimate of the standard deviation for each of the data sets in part b.

d. Every data point in a set of 10 points equals either 1 or 8. Find the sets that will have the least and greatest possible standard deviations.

21. Table 7.3 shows the batting averages of the American League batting champions from 1941 to 1994. The mean is 0.344, and the standard deviation is 0.021. Are there any unusual data values here? If so, who are the players and in what year did they play?

Table 7.3
Batting Averages of American League Batting Champions

Year	Player and Club	Average
1941	Theodore Williams, Boston	.406
1942	Theodore Williams, Boston	.356
1943	Lucius Appling, Chicago	.328
1944	Louis Boudreau, Cleveland	.327
1945	George Stirnweiss, New York	.309
1946	Mickey Vernon, Washington Senators	.353
1947	Theodore Williams, Boston	.343
1948	Theodore Williams, Boston	.369
1949	George Kell, Detroit	.343
1950	William Goodman, Boston	.354
1951	Ferris Fain, Philadelphia	.344
1952	Ferris Fain, Philadelphia	.327
1953	Mickey Vernon, Washington Senators	.337
1954	Roberto Avila, Cleveland	.341
1955	Albert Kaline, Detroit	.340
1956	Mickey Mantle, New York	.353
1957	Theodore Williams, Boston	.388
1958	Theodore Williams, Boston	.328
1959	Harvey Kuenn, Detroit	.353
1960	Pete Runnels, Boston	.320
1961	Norman Cash, Detroit	.361
1962	Pete Runnels, Boston	.326
1963	Carl Yastrzemski, Boston	.321
1964	Tony Oliva, Minnesota	.323
1965	Tony Oliva, Minnesota	.321
1966	Frank Robinson, Baltimore	.316
1967	Carl Yastrzemski, Boston	.326

Table 7.3 (continued)
Batting Averages of American League Batting Champions

Year	Player and Club	Average
1968	Carl Yastrzemski, Boston	.301
1969	Rodney Carew, Minnesota	.332
1970	Alexander Johnson, California	.329
1971	Tony Oliva, Minnesota	.337
1972	Rodney Carew, Minnesota	.318
1973	Rodney Carew, Minnesota	.350
1974	Rodney Carew, Minnesota	.364
1975	Rodney Carew, Minnesota	.359
1976	George Brett, Kansas City	.333
1977	Rodney Carew, Minnesota	.388
1978	Rodney Carew, Minnesota	.333
1979	Frederic Lynn, Boston	.333
1980	George Brett, Kansas City	.390
1981	Carney Lansford, Boston	.336
1982	Willie Wilson, Kansas City	.332
1983	Wade Boggs, Boston	.361
1984	Donald Mattingly, New York	.343
1985	Wade Boggs, Boston	.368
1986	Wade Boggs, Boston	.357
1987	Wade Boggs, Boston	.363
1988	Wade Boggs, Boston	.366
1989	Kirby Puckett, Minnesota	.339
1990	George Brett, Kansas City	.329
1991	Julio Franco, Texas	.342
1992	Edgar Martinez, Seattle	.343
1993	John Olerud, Toronto	.363
1994	Paul O'Neill, New York	.359

Source: *The World Almanac and Book of Facts,* 1995

LESSON 8

Comparing Measurements

What is the purpose of standardized testing?

What standardized tests have you taken?

Headlines in the August 19, 1993 *USA Today* proclaimed: "Overall SAT scores higher . . . ." Academic life as a student is often evaluated by a test score. Two of the more widely used scores are those from the Scholastic Aptitude Test (SAT) and the American College Test (ACT). The public is interested in the results of these tests as an indication of how well our educational system is functioning and how well students are learning.

OBJECTIVES

Rescale data for making comparisons.

Use a formula to standardize scores.

INVESTIGATE

Comparing Scores from Different Tests

How do scores from the SAT and the ACT compare? Do we get the same information from the two tests? Have scores for both tests changed in the same way over the years? It is difficult to tell whether test scores on the two tests have changed similarly or even to compare scores because the results are scored on two different scales.

The data in Table 8.1 show national average SAT and ACT mathematics scores for a period of 20 years from 1970 to 1989. How can you compare these scores? How did the average test scores change over time? Consider the data.

Table 8.1
SAT/ACT Scores

Year	SAT Math	ACT Math	Year	SAT Math	ACT Math
1970	488	20.0	1980	466	17.4
1971	488	19.1	1981	466	17.3
1972	484	18.8	1982	467	17.2
1973	481	19.1	1983	468	16.9
1974	480	18.3	1984	471	17.3
1975	472	17.6	1985	475	17.2
1976	472	17.5	1986	475	17.3
1977	470	17.4	1987	476	17.2
1978	468	17.5	1988	476	17.2
1979	467	17.5	1989	476	17.2

Source: *The American Almanac*, 1995

Discussion and Practice

1. Study Table 8.1. What observations can you make by looking at the numbers?

2. Work in groups to on the following problems. Divide the tasks in parts a, b, and c among your group members. Then use your results for part d.

a. Construct number-line plots for the two data sets. Describe patterns you see. Describe key differences between the two data sets.

b. Construct parallel box plots for the two data sets. Are there any outliers? How do the box plots differ?

c. Construct plots over time for the two data sets placing the year on the horizontal axis. What patterns do you see?

d. Compare the three methods for plotting the data in parts a, b, and c. Which plots are most useful? Why? What do the line plots show that is not in the box plots?

3. a. Find the mean, the median, and the standard deviation for the SAT and ACT data. Explain what they tell you about the scores.

b. Do the numerical statistics you found in part a help you compare the ACT and the SAT scores? If not, why not?

c. Are there changes you could make that might make the comparisons easier? Explain your answer.

4. Consider the 1989 SAT value of 476 and 1989 ACT value of 17.2.

   a. Locate each value on the respective number-line plot. Approximately how many standard deviations is the 476 from the mean SAT score? The 17.2 from the mean ACT score?

   b. What does this tell you about the two scores relative to each other?

If scores are expressed in terms of the number of standard deviations from the mean, then scores from different scales can be compared. A score that is 1.5 standard deviations above its mean seems to be closer to the typical score for that data than one that is 2 standard deviations below its mean. These are called *standardized scores.*

To make comparisons of scores measured on different scales, the scores should be written in terms of the number of standard deviations from the mean.

5. Consider a set of data with mean 150 and standard deviation 10.

   a. How many standard deviations from the mean is a score of 160?

   b. How could you represent a score that is less than the mean?

   c. A score of 135 would be how many standard deviations from the mean?

   d. What score would be 2 standard deviations from the mean?

6. The number of standard deviations between a data point and its mean is called a "*z* score." The *z*-score formula is $z = \frac{(x_i - \bar{x})}{s}$, where x_i is the given score, $\bar{x}$ is the mean, and s is the standard deviation. You can use the formula with a computer spreadsheet or a calculator to make your work easier.

   a. Find the transformed or *z* scores for the SAT and ACT scores. Record your results on *Activity Sheet 5.*

b. In which year was the ACT score farthest from the average? How can you tell from the z scores?

7. Work in groups on the following problems. Divide the tasks for parts a, b and c among the members of your group, and then use your results for part d.

a. Make number-line plots of the z scores, or standardized scores, on the same number line. How do they compare?

b. Construct parallel box plots for the two sets of standardized scores. Recall the outliers from the original plot. Are the same outliers still present as z scores?

c. On the same grid, construct plots over time for the z scores for the SAT data and for the ACT data. Can you see meaningful differences in the patterns?

d. Write a paragraph comparing SAT and ACT scores from 1970 to 1989.

8. Cities in the United States are often rated in different categories according to certain criteria. The scores for the top-ten cities in the categories of jobs and recreation are listed in Table 8.2.

Table 8.2
Recreation and Jobs for Top-Ten U.S. Cities

City	Jobs	Recreation
San Diego, California	14,772	3,800
Los Angeles, California	5,547	3,857
New York, New York	1,623	2,125
Boston, Massachusetts	3,456	2,278
Washington, DC	16,288	1,857
Philadelphia, Pennsylvania	6,895	2,402
Dallas, Texas	8,964	1,338
Detroit, Michigan	5,559	2,096
Minneapolis, Minnesota	6,242	2,273
Chicago, Illinois	4,240	2,252

Source: *Places Rated Almanac,* 1994

For all 343 cities, the job-score mean was 7,509 and the standard deviation was 4,784. For the same 343 cities, the recreation-score mean was 1,586 and the standard deviation was 677.

a. Should the city of Minneapolis advertise itself as better for recreation or having a highly rated job market? Explain.

b. Is San Diego farther above the mean in jobs or recreation?

c. Were the ratings for any of the cities unusual? Explain your answer.

Summary

It is difficult to compare data that are expressed in different units. Data that are in different units or scales can be standardized by using this formula:

$z = \frac{(x_i - \bar{x})}{s}$, where $\bar{x}$ is the mean of the data, s is the standard deviation, and x_i is any one of the scores.

The standardized data are called *z* scores. Once the data have been standardized, they can be compared.

Practice and Applications

Table 8.3
SAT/ACT Scores Extended

Year	SAT Math	SAT Math	Year	SAT Math	ACT Math
1970	488	20.0	1982	467	17.2
1971	488	19.1	1983	468	16.9
1972	484	18.8	1984	471	17.3
1973	481	19.1	1985	475	17.2
1974	480	18.3	1986	475	17.3
1975	472	17.6	1987	476	17.2
1976	472	17.5	1988	476	17.2
1977	470	17.4	1989	476	17.2
1978	468	17.5	1990	476	19.9
1979	467	17.5	1991	474	20.0
1980	466	17.4	1992	476	20.0
1981	466	17.3	1993	478	20.1

Source: *The American Almanac,* 1994

9. In 1990, the ACT test was revised, and it became difficult to compare its scores with ACT scores from previous years. Notice the change from 1989 to 1990 in Table 8.3. Each year, however, there is a mean SAT and a standard deviation for that year and a mean ACT and a standard deviation for that year.

In the 1992–1993 school year, the mean SAT mathematics score was 478, and the standard deviation was 125. That same school year, the mean ACT score was 20.1 with a standard deviation of 4.5.

a. Suppose a college has set 500 as the minimum SAT mathematics score acceptable for entry into one of its programs. What ACT score is comparable to this value? How can you use a graph to help you find your answer? Try to think of another way you can answer the question.

b. What SAT average is comparable to an ACT average of 17.2?

c. Kuong took both the SAT and the ACT. On the SAT he received a 690 and on the ACT a 26. If he would like to send his better score in his application to a university, which of the scores should he send? Explain how you made your choice.

10. For a mean of 10 and standard deviation of 3, Shakia found the z score for 15 by entering the following steps on a calculator: $15 - 10 \div 3$. Did she find the correct z score? Explain why or why not.

11. Statistical information on players in the National Basketball Association Hall of Fame is contained in Table 8.4. Coaches are interested in players' skills in scoring and rebounding. How do the players in the Hall of Fame compare on these two categories?

Table 6.4
NBA Basketball Hall of Fame, 1990–1994

Year	Player	Games	Points	Field-Goal %	Rebounds
1991	Archibald, Nate	876	16,481	.467	2,046
1993	Bellamy, Walt	1,043	20,941	.516	14,241
1990	Bing, Dave	901	18,327	.441	3,420
1994	Blazejowski, Carol	101	3,199	.546	1,015
1991	Cowens, Dave	766	13,516	.460	10,444
1993	Erving, Julius (Dr. J.)	836	18,364	.507	5,601
1991	Gallatin, Harry	682	8,843	.398	6,684
1992	Hawkins, Connie	499	8,233	.467	3,971
1990	Hayes, Elvin	1,303	27,313	.452	16,279
1993	Issel, Dan	718	14,659	.506	5,707
1990	Johnston, Neil	516	10,023	.444	5,856
1992	Lanier, Bob	959	19,248	.514	9,698
1993	McGuire, Dick	738	5,921	.389	2,784
1993	Meyers, Ann	97	1,685	.500	819
1990	Monroe, Earl	926	17,454	.464	2,796
1993	Murphy, Calvin	1,002	17,949	.482	2,103
1993	Walton, Bill	468	6,215	.521	4,923

Source: *The Universal Almanac,* 1995

a. Find the rebounding rate per game for each player.

b. Find the mean and standard deviation for the rebounding rate and for field-goal-shooting percents.

c. Find the z scores for each player in rebounding and field-goal-percents. Record your results on *Activity Sheet 6*. Make a box plot and a number-line plot of the z scores.

d. Was Julius Erving more outstanding at shooting or at rebounding compared to his Hall of Fame peers? How can you tell?

e. Compare rebounding and field-goal shooting for the players. Describe any differences you see.

f. Which player was best in both shooting and rebounds? How did you decide?

12. Consider the formula for finding a z score.

a. Would it make a difference if the formula were written as $\frac{(x_i - \bar{x})}{s_x}$? Explain your answer.

b. Are $\bar{x}$ and s variables?

Cars

OBJECTIVES

Create and use formulas.

Organize and interpret data.

1. Read the attached article (pages 83 and 84) and answer the following questions.

   a. Describe at least four variables mentioned in the article that might affect the number of injuries and deaths in car accidents.

   b. Why would anyone be interested in keeping track of the number of deaths and injuries?

   c. Make a graph based on information in the article that might reflect the number of deaths since 1970.

   d. The article gives the following facts:

   - The speed limit in Virginia was raised to 65 mph in 1988.
   - Before the change, Virginia drivers averaged about 8 mph above the posted speed limits.
   - After the change, drivers in Virginia increased their average speed by about 3 mph.

   Write a description of how fast drivers in Virginia were traveling between 1985 and 1995, based on the information above.

   e. The federal speed limit was removed in 1995. In 1992, according to the U.S. Census, there were 39,235 fatalities due to motor-vehicle accidents. Based on the information in the article, a formula that could be used to project the number of deaths for the next five years is $D = 39{,}235 + (N - 1992)6400$. Explain the formula and use it to decide whether the number of deaths predicted for the year 2000 will be double those in 1992.

2. The data in Table A1.1 give the typical price and the highway gas mileage for a random sample of new cars. The cars are divided into three size classes, with 1 = small cars, 2 = compact cars, and 3 = mid-size cars.

Table A1.1
Cars: Typical Price/Highway Gas Mileage

Car	Size	Price	Highway miles per gallon (mpg)
Acura Integra	1	$15,900	31
Dodge Colt	1	$9,200	33
Honda Civic	1	$12,100	46
Hyundai Excel	1	$8,000	33
Mazda 323	1	$8,300	37
Mitsubishi Mirage	1	$10,300	33
Subaru Justy	1	$8,400	37
Suzuki Swift	1	$8,600	43
Toyota Tercel	1	$9,800	37
Volkswagen Fox	1	$9,100	33
Chevrolet Corsica	2	$11,400	34
Chrysler LeBaron	2	$15,800	28
Ford Tempo	2	$11,300	27
Mazda 626	2	$16,500	34
Subaru Legacy	2	$19,500	30
BMW 535i	3	$30,000	30
Chevrolet Lumina	3	$15,900	29
Dodge Dynasty	3	$15,600	27
Ford Taurus	3	$20,200	30
Hyundai Sonata	3	$13,900	27
Lexus SC300	3	$35,200	23
Mitsubishi Diamante	3	$26,100	24
Nissan Maxima	3	$21,500	26
Pontiac Grand Prix	3	$18,500	27
Toyota Camry	3	$18,200	29
Volvo 850	3	$26,700	28

Source: Robin Lock, "1993 New Car Data," *Journal of Statistics Education*, Vol. 1, No. 1, July 1993

Use the data in the table. In each case, explain how you found your answer.

a. What is the mean gas mileage for small cars?

b. What is the average difference in gas mileage between small and compact cars?

c. What is the average price difference between small and compact cars?

d. What is the average price difference between compact and mid-size cars?

3. A car is typically driven about 12,000 miles per year. Suppose gasoline costs $1.19 per gallon.

a. How much would it cost to buy and drive a Hyundai Sonata for one year?

b. Write a formula for the cost of buying and driving any one of the cars in Table A1.1. Use your formula to determine which car is the most economical. How did you make your choice?

c. Suppose you drove 20,000 miles per year. Would it be better to get a car that gets better gas mileage? Explain your answer.

4. The typical price of new cars in 1993 related to number of airbags is given in Table A1.2.

Table A1.2
Cars: Typical Price by Number of Airbags

Number of Airbags	Number of Cars	Mean Price ($1,000)	Median Price ($1,000)	Minimum Price ($1,000)	Maximum Price ($1,000)	Q1	Q3
0	33	13.70	11.60	7.4	23.3	9.15	16.45
1	43	21.22	19.90	9.8	47.9	15.60	26.30
2	16	28.37	25.55	15.1	61.9	17.88	35.88

Source: Robin Lock, "1993 New Car Data," *Journal of Statistics Education*, Vol. 1, No. 1, July 1993

a. Describe the differences in the mean price and in the median price depending on the number of airbags.

b. Why are there such differences?

c. Would you use the mean price or the median price in a newspaper ad? Why?

5. Use the data in Table A1.2.

a. Draw box plots for the costs of the cars in the three airbag categories. Compare the prices.

b. The standard deviations for the three categories are 4.421, 8.24, and 12.55. Match each standard deviation to an airbag category.

c. The maximum price of $47,900 for a car with one airbag is an outlier. The next higher price in that category is about $38,000. Find an approximate mean for cars with one airbag from the given information without the outlier.

Death toll from increasing speed limits yet to be told

By Don Phillips, *Washington Post,* December 10, 1996

WASHINGTON—The freedom that Congress gave states to set their own speed limits—a freedom that began on Friday—comes without an owner's manual.

There are many variables in highway safety and few serious scientific studies on the complicated relationship between speed and accidents. Experts agree that deaths and injuries probably will increase with higher speeds, but they cannot predict by how much.

The question affects not only highway safety, but also state budgets, partly because one-third of all Medicaid costs are the result of traffic accidents, and states can no longer count on automatic reimbursement of those costs by the federal government.

Transportation Secretary Federico Pena has been telling states that "some day you may have all the Medicaid costs in your state" and that one of the ways to control health costs is to reduce traffic accidents.

Traffic accidents already are the greatest killer of people age 5 to 28, a group whose members are less likely to have health insurance and, therefore, more likely, if they have disabling injuries, to become lifetime wards of the state.

That same age group is likely to be affected disproportionately by another section of the new federal legislation, the one that allows states to repeal motorcycle helmet laws without penalty.

Charles Hurley of the Insurance Institute for Highway Safety—which favored retaining the federal speed limit—said Congress acted on "a combination of philosophy and wishful thinking," but states will now have to act more responsibly.

"States don't have the luxury of passing the buck like Congress," he said.

Both accident rates and traffic fatalities have declined since 1974, when the federally mandated 55-mph speed limit was inaugurated. Traffic crashes now kill about 40,000 people a year, an average of about 110 per day. When there was no federal speed limit, annual fatalities were running as high as 55,000.

During the same period, miles driven per year have almost doubled and the number of the vehicles on the road has increased dramatically—all factors suggesting that deaths and accident rates should have increased, not decreased. Further, as all drivers know, speed limits often are violated, so the posted speed limit may have little relationship to the average speeds actually driven on a given stretch of road.

Many factors other than lower speed limits have contributed to reduced fatalities. Automobiles are safer, with everything from air bags to better suspension systems to federally mandated center-mounted rear brake lights that reduce rear-end collisions. Thousands of miles of new highways have been built, including many interstates and many more roads that meet interstate standards. An emphasis on getting drunken drivers off the road has clearly had some results.

In studying the effect of speed on deaths and injuries, researchers are confounded by all the variables. But even with all those caveats, speed clearly contributes to the severity of auto crashes. State troopers know it. Trauma physicians know it. The studies that have been done point in that direction.

During the debate over speed limits, the most-cited statistics stated that 6,400 more people would die every year and an extra $19 billion would be added in health care costs if the federal speed limit were eliminated.

The 6,400 number resulted not from a study but from a simple projection of data by the National Highway Traffic Safety Administration on the effects of the congressionally approved change in 1987 from a 55-mph to a 65-mph speed limit on some rural interstates.

Another study, published by the Transportation Research Board in 1990, compared driver behavior in Maryland with Virginia in July 1988, when Virginia raised interstate speed to 65 mph for cars but Maryland maintained a 55-mph limit.

Before the change, drivers in Virginia averaged about 8 mph above the posted speed limits: drivers in Maryland about 7 mph.

After Virginia changed the speed limit, automobile drivers in Virginia almost immediately increased their average speed by about 3 mph, while drivers in Maryland stayed at roughly the same speed as before.

But the number of drivers in Virginia who were exceeding 70 mph doubled after the speed limit was increased.

Source: *Milwaukee Journal/Sentinel*, December 10, 1996

Unit II

Functions

INTRODUCTORY ACTIVITY

Time on Task

Telephones

Most people spend some time each day talking on the phone. The time a person spends on the phone is a *function* of the day. For example, a student may spend more time on the phone on Saturday, while a business person may spend more time on the phone on a weekday. How much time do you spend on the phone?

1. Estimate how many minutes you think you will talk on the phone each day for the next seven days. Make a plot with days along the horizontal axis and minutes on the phone along the vertical axis.
2. If you spent the same amount of time each day on the phone, what would the plot look like? Explain your answer.
3. Suppose you and your family were discussing the total amount of time you think you will spend on the phone for a week.
   a. How do you think the total amount of time at the end of Tuesday would compare with the total amount of time at the end of Monday?
   b. How could you use the information and plot from problem 1 to help you answer the question in part a?
   c. What would the graph look like if you plotted the cumulative amount of time after each day in the week rather than the time for that day?
4. Describe what you think would be the typical amount of time you spend each day on the phone. How can you see this in your plot?
5. Use your plot to write a description of the time you think you will talk on the phone during a week.
6. Compare your plot and description to those of your classmates. How are they alike? How are they different?

7. Record how many minutes you actually talk on the phone each day for a period of seven days. Make a graph of your data on the grid you used for problem 1, beginning with the first day. At the end of each day, add the new information to your plot. How does the graph of your estimates compare to the graph of the actual times?

Studying

The amount of time students spend studying each day and each week will vary according to the work assigned and the amount of effort exerted by students. Are other variables important? Is the time spent studying a function of the time spent watching television?

8. Estimate how many minutes you will study each day over the next seven days. Then estimate how many minutes you think you will watch television each day for the next seven days.

    a. Make a plot (*minutes studying, minutes watching TV*) of your estimates.

    b. Describe the relationship that might exist between studying and watching TV if you use your estimates.

    c. Based on the plot and your data, decide if the following statement is accurate: "Studying is a function of the time spent watching television." Explain your decision.

LESSON 9

An Introduction to Functions

Many automobile accidents involve rear-end collisions. How fast do people react when they know they have to stop their cars?

How long do you think it takes for the car to stop after the driver steps on the brakes?

How close to other cars can someone drive safely, or does it make any difference?

Stopping a car that is traveling at a certain speed takes some time and some distance. Experts have found that the typical reaction time that a driver requires to get his or her foot on the brake after seeing an emergency is about 0.75 second. The car, of course, travels some distance during this reaction time; this distance is the *reaction distance*.

OBJECTIVES

Develop the notion that one variable depends on another.

Represent relationships with tables, graphs, and symbols.

Investigate several different functional relationships.

INVESTIGATE

How Far Before You Stop?

Table 9.1 shows the reaction distance R, for vehicles traveling at various speeds S.

Table 9.1
Reaction Distance

Speed S (miles per hour)	Reaction Distance R (feet)
20	22
30	33
40	44
50	55
60	66
70	77

Source: U.S. Bureau of Public Roads 1992

Discussion and Practice

1. Look carefully at the speeds and reaction distances in Table 9.1.

   a. Do you see a pattern? By how much does R increase for every 10-mile-per-hour increase in speed?

   b. What does $\frac{R}{S}$ represent? Add a new column to the table by calculating $\frac{R}{S}$ for each row. What pattern do you observe in the values of $\frac{R}{S}$?

   c. Given a new value of 35 mph for S, what would you predict R to be?

   d. Write a symbolic expression relating R to S.

Under ideal conditions, we can predict R perfectly from S by using the equation $R = (1.1)S$, regardless of the value of S.

R is said to be a *function* of S. This function works not only for values of S in the table, but also for any other value of S that might be needed.

2. What is the reaction distance of a car traveling at

   a. 55 mph?

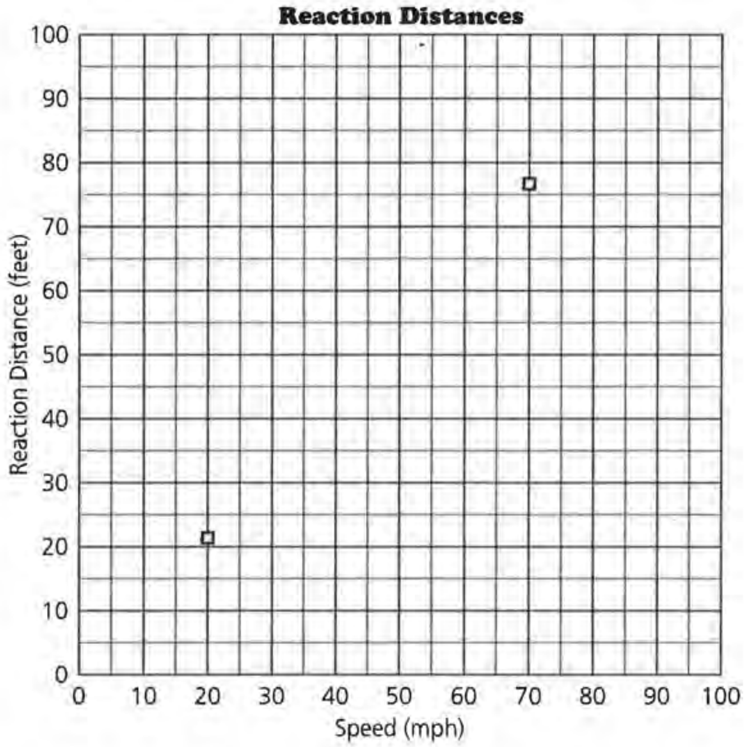
   b. 80 mph?

   c. 100 mph?

3. If it took 60 feet to get your foot on the brake, how fast had the car been traveling?

Relationships and Graphs

A plot of reaction distances as a function of speed will help you gain a visual impression of this relationship. The plot below shows such a graph with the points ($S = 20$, $R = 22$) and ($S = 70$, $R = 77$) already plotted.



4. Fill in the points for the remaining four rows of the table using *Activity Sheet 7*.

   a. What pattern emerges from the six points on the graph?

   b. Use the plot to approximate the reaction distance for a car traveling at 65 miles per hour. Then calculate R when S is 65 miles per hour. How do the two values compare?

   c. Plot the results you found in problem 2 above for $S = 55$, 80, and 100 on the same grid. Describe the graph.

   d. The graphical representation of the relationship between R and S is a straight line. Connect the points on the graph with a line, and use it to approximate the speed when reaction distance is 60 feet.

5. Does it make sense to connect the points on the graphs with lines? What can you learn by drawing line segments to connect the points?

Summary

R is a function of S when R depends in some way on S. From any value of S, not just the ones in a table, you can determine R.

6. Think about the concept of a function.

a. List three different ways to represent a function and give an example of each of the three.

b. What can you learn from each of the three ways?

Functions That Are *Not* Straight Lines

So far, only the reaction distance for stopping a car has been considered. Once the brakes are applied, the car does not stop immediately. After the driver's foot hits the brakes, the distance the car travels is called the *braking distance*. The total *stopping distance* is the sum of the reaction distance and the breaking distance. Table 9.2 is an expanded version of Table 9.1, now showing speed, reaction distance, and braking distance.

Table 9.2
Speed and All Three Distances

Speed S (mph)	Reaction R (feet)	Braking B (feet)	Stopping T (feet)
20	22	20	______
30	33	40	______
40	44	72	______
50	55	118	______
60	66	182	______
70	77	266	______

Source: U.S. Bureau of Public Roads 1992

7. Use the data in Table 9.2 to plot (*speed, braking distance*).

a. What pattern emerges?

b. Describe how braking distance depends on speed.

8. Use your graph from problem 7 to approximate each of the following.

a. The braking distance for a car traveling 55 miles per hour

b. The braking distance for a car traveling 80 miles per hour

c. The speed of a car for which the braking distance was 100 feet

9. Consider the total stopping distance *T*.

a. Write *T* as a symbolic expression involving *B* and *S*. Then use this expression to fill in the stopping distance column of the table.

b. Plot (*S*, *T*). What pattern emerges?

c. Use the plot to approximate the stopping distance for a car traveling 55 miles per hour.

d. Use the plot to approximate the stopping distance for a car traveling 80 miles per hour.

e. A police officer investigating an accident found that the car involved in the accident required 300 feet to stop. Approximate the speed at which the car was traveling.

10. Drivers must consider what is a safe distance between their car and the car ahead of them.

a. A car is about 17 feet long. About how many car lengths should you maintain between your car and the car in front of you while driving at approximately 30 miles per hour?

b. How would your response to part a change if you were traveling 70 miles per hour? 55 miles per hour?

Summary

A function is a rule that allows us to determine the value of one variable by knowing the value of another variable. Functional relationships between variables can be expressed by a rule or formula, by a graph, or by a table of values. Some functional relationships are represented by straight lines, while others are represented by curves.

Practice and Applications

11. Speeding brings fines and penalties that vary according to local and state rules, but usually a fine is a function of how much you exceed the speed limit. In one community, the fine for the first 10 miles over the speed limit is \$89.60. For each mile beyond the first 10 miles, the fine increases by \$4.80.

a. How much would you be fined for going 15 miles over the speed limit?

b. The fine F as a function of the number of miles over the speed limit can be written $F = 89.60$ for $0 < E \leq 10$ and $F = 89.60 + (4.80)(E - 10)$ for $E > 10$, where E is the number of miles over the speed limit. What is the fine for going 25 miles over the limit?

c. Graph the function and use it determine the fine for going 45 miles over the limit.

d. Use the graph to approximate by how many miles per hour you were speeding if you paid a fine of about $200.

12. According to the American Automobile Manufacturers Association, the cost of operating an automobile in 1992 was 45.77 cents a mile. In addition, insurance and license fees averaged about $926.

a. If you planned to drive about 12,000 miles per year, how much should you plan on spending to run your car?

b. Would you expect the relationship between annual cost and number of miles driven annually to be a straight line? Why or why not?

c. Show how you can use a graph, a table, and a formula to help you answer the questions in parts a and b. Which way makes the most sense to you?

13. Suppose you leased a car for $239 a month. Assume you had to pay the following costs:

Gas and oil: 6 cents per mile
Maintenance: 2.2 cents per mile
Tires: 0.91 cents per mile

a. Tiasha's group wrote two different functions to describe the cost of leasing a car for a year, using M to stand for the number of miles driven:

$C = \$239 \cdot 12 + \$0.06M + \$0.022M + \$0.0091M$

$C = \$239 \cdot 12 + \$0.0911M$

Which function should they use? Explain how you made your choice.

b. If you drive about 12,000 miles per year, would it be more or less expensive to rent a car or to drive one you owned? Explain how you made your decision.

c. Would your decision change if you increased or decreased the number of miles you drive per year? Explain your answer.

d. If a new car costs about $11,000, would you rather rent or lease a car? Explain your answer.

14. Think again about the problem posed in the investigation. It is possible to establish the formula below, relating braking distance *B* to the speed *S* at which a car is traveling.

$B = 36.2 - 2.34S + 0.08S^2$

a. What is the braking distance for a speed of 70 mph? Indicate this length by relating it to some familiar object or distance.

b. Generate at least four more sets of ordered pairs for the formula for *B*. Look carefully at the table of values you generated. Would you suspect the graph of the rule to be a straight line? Why or why not?

c. Graph the relationship (*S*, *B*) by using either the ordered pairs or the formula itself. Is the graph a straight line?

Extension

15. In an earlier lesson, you combined variables to make new variables. You can also combine formulas to make new formulas.

a. Combine the formula for *R* in terms of *S* with the formula for *B*. The new formula expresses *T* (the total stopping distance) as a function of *S* (speed).

b. Plot the new function (*S*, *T*). What happens to your total stopping distance as your speed increases?

c. Write a summary statement on the relationships among reaction distance, braking distance, and stopping distance for cars traveling at various speeds.

LESSON 10

Trends over Time

How long do animals live?

How old is a cat in "cat years" when it is five years old?

Is there a way to find an equivalent human age if you know an animal's age?

OBJECTIVES

Develop piecewise functions from information conveyed in words or data tables.

Recognize linear trends.

Use a ratio to express the slope in a linear relation.

Animals have life spans that are different from those of humans. Some animals live a very long time, while others live a very short time. How do humans' and animals' life spans compare? Read Dr. Huntington's reply to Barbara.

Dear Dr. Huntington:

How does a cat's age compare to a human's?

Barbara

Dear Barbara,

I've seen several formulas that attempt to relate animals' ages to humans' ages, but the one I'm comfortable with and feel is in the ballpark for both dogs and cats is as follows: The first year of life equals about 15 human years. The second year adds another 10 years. So a 2-year-old cat is about age 25 in human years. For each additional year of life past age 2, add about 5 years. Using this formula, a 12-year-old cat would have an age equivalent to that of a 75-year-old person....

INVESTIGATE

Animal Ages

The exchange above was reported in the November 9, 1993, *Gainesville Sun*'s column on advice for animal owners. Can the ideas presented here be generalized into a mathematical rule so that if you know the age of a cat, you can determine its equivalent human age?

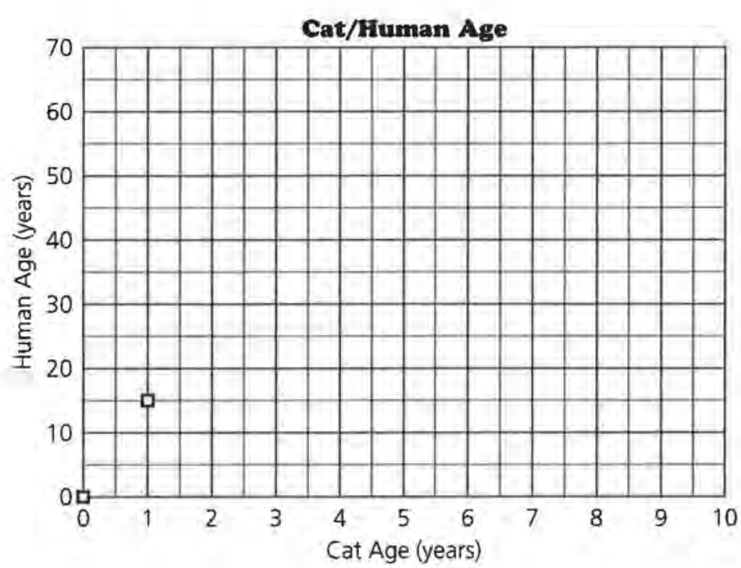
Discussion and Practice

1. Construct a table showing the actual cat age C in the first column and the equivalent human age H in the second column. To get you started, a few values have been given in the table below. Why does the first row contain (0, 0)?

Table 10.1
Cat/Human Age

C (years)	H (years)
0	0
1	15
2	______
3	______
4	______
5	______
10	______

2. After completing the table, plot the points on a rectangular grid with C, the cat's age, on the horizontal axis. Such a plot showing the first two points in Table 10.1 is illustrated below.



a. Describe the shape of the graph produced by the points.

b. What would be the equivalent human age for a cat that is 4.5 months old? How could this be determined from the graph?

3. A convenient way to answer questions like the one in problem 2b is to describe the information provided by Table 10.1 as an equation and and then graph the equation. This equation will be represented by the line segment connecting the points (0, 0) and (1, 15).

a. Suppose a cat ages from two to three months. What happens to the equivalent human age?

b. Write a ratio of H to C for the first year.

c. Use the ratio to find the equivalent human age for a cat that is 3 months old.

d. Is there any advantage to using the ratio rather than the graph to answer part c? Explain your answer.

4. The 15 indicates how fast H increases for the first-year change in C. Now, draw a segment between the point (1, 15) and the point at $C = 2$ on the plot.

a. If a cat ages from 14 to 15 months, what happens to the equivalent human age ? Is the per-unit change in H to C for the second year the same ratio as that in the first year? Why or why not?

b. What is the equivalent human age for a cat that is 1.5 years old? Explain your answer.

5. Suppose $2 < C \leq 3$.

a. What does $2 < C \leq 3$ represent on the C-axis?

b. Give the ratio of the increase in H to the increase in C in this interval.

c. If $3 < C \leq 4$, write the ratio for the change in H to the change in C.

d. What observation can you make about the ratio if $C > 2$?

6. Use the information you have found.

a. Suppose a cat is 14 years old. What is the equivalent human age?

b. Suppose a veterinarian examines a cat and concludes that the cat appears to be around 60 years old, in human terms. What is the cat's actual age?

c. What are some assumptions, not written in response to the letter, that are incorporated in the veterinarian's guidelines for determining H from C?

Personal Income

Personal income for residents of the United States has gone up considerably since 1980. But that does not mean that people are better off financially.

7. How could the general financial well-being of residents of the United States be worse today than in 1980, even though salaries and wages have increased by a large amount?

"Disposable personal income" is the amount of money people have after they have paid taxes on their earnings. Table 10.2 shows per-capita disposable personal income (DPI) for selected years since 1981. Also shown on the same per-capita basis are expenditures for food, transportation (including the purchase and operation of automobiles), housing, and medical expenses.

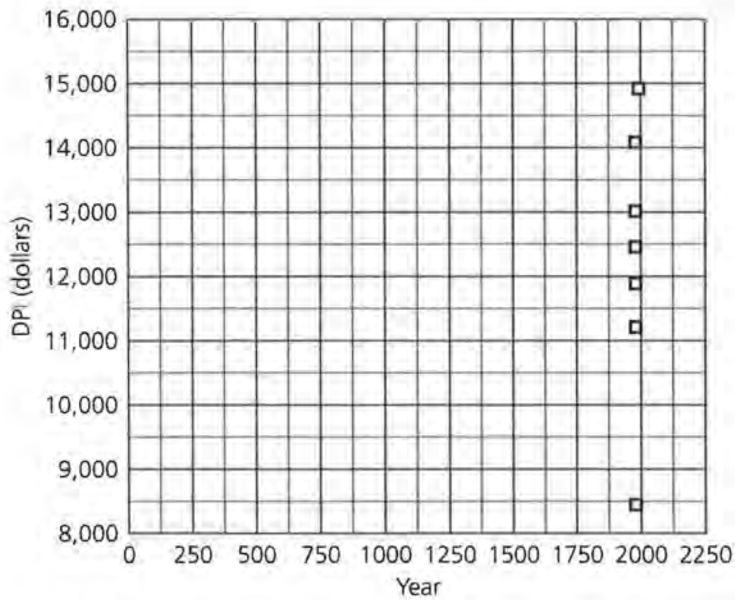
8. What is meant by the term *per capita?*

Table 10.2
Per-Capita Expenditures (in dollars)

Year	DPI	Food	Transp.	Housing	Medical
1981	9,455	1,620	1,045	1,149	823
1984	11,673	2,039	1,405	1,580	1,269
1985	12,339	2,127	1,519	1,701	1,384
1986	13,010	2,233	1,531	1,815	1,497
1987	13,545	2,345	1,575	1,944	1,654
1988	14,477	2,454	1,676	2,052	1,824
1989	15,307	2,605	1,734	2,173	1,970
1990	16,174	2,599	1,830	2,208	2,402
1991	15,658	2,651	1,746	2,288	2,615

Source: *Statistical Abstract of the United States,* 1994

9. To see how DPI changes over time, Jorge made a plot of the ordered pairs *(year, DPI)* on a rectangular coordinate system. It is shown below.



a. Is the plot useful? Why or why not?

b. Is this a good plot for seeing trends in personal income over the years? What might be done to make this plot better?

10. Begin the time axis at 1980 and the DPI axis at around 8,400 and plot *(year, DPI)*.

a. Describe the trend in DPI over the years shown.

b. Use a straightedge to draw a single straight line close to the data points that you think best represents the trend in these data. Use the line you drew to answer this question: About how much does the DPI tend to increase each year?

11. a. For 1984 onward, calculate the gain in DPI each year. (The gain from 1984 to 1985 is 12,339 – 11,673 = 666.) What is the average gain per year over the period 1984–1991?

b. How does the average gain per year compare to your conclusions in Problem 10?

Summary

A ratio can be used to describe the relationship between changes in a quantity and time. After a cat is two years old, the increase in age of the cat is related to an equivalent increase in human age by a constant ratio. Growth in disposable personal income seems to be related to time by a constant ratio. In both cases, a plot of the data seems to fall near a straight line that slopes upward as you move from earlier to later years. That is, both equivalent human age and DPI are increasing linearly.

Practice and Applications

12. What about the other per-capita personal expenditures in Table 10.2? Do they exhibit the same trends as personal income?

- **a.** Plot per-capita transportation expenditures as a function of time. Describe the trend in transportation expenditures over time. Does it look as if a single straight line would adequately explain this trend?
- **b.** Could two different straight lines be used to explain the trends in transportation expenditures over the years? If so, what years would you use for each of the lines?
- **c.** How would the ratios for the two lines compare?

13. Divide the tasks in parts a and b among your group members. Share your results to answer part c.

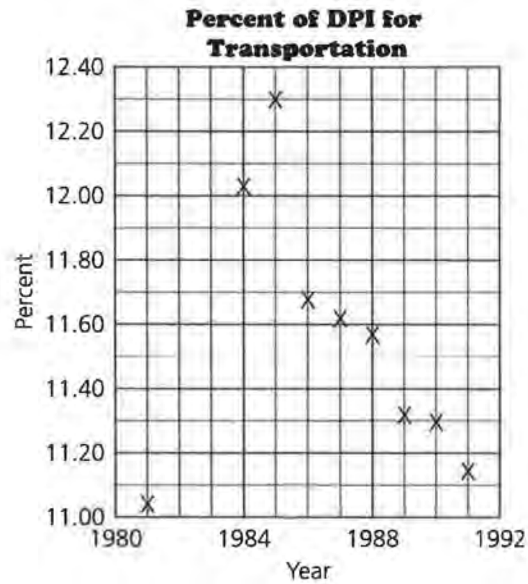
- **a.** Plot per-capita housing expenditures as a function of time. Describe the trends in the plot.
- **b.** Plot per-capita medical expenditures as a function of time. Describe the trends in the plot.
- **c.** Compare the plots you have made of transportation, housing, and medical expenditures, each as a function of time. What observation can you make?

14. So far, the investigations and discussions have led you through an analysis of trends over time for DPI and individual types of personal expenditures. This analysis does not, however, provide information on how one type of expenditure relates to income or to other types of expenditure. Are transportation costs using up more and more of DPI?

- **a.** In 1989, DPI was $15,307 and the amount spent on food was $2,605. What percent of the DPI was spent on food?

b. Calculate the transportation expenditures for 1991 as a percent of DPI.

c. The plot of percent of income spent on transportation as a function of time is shown in Figure 3. Comment on the trends you observe.



15. Divide the following topics among your group members, and fill out a table for each.

Food expenditures as a percent of DPI over time
Housing expenditures as a percent of DPI over time
Medical expenditures as a percent of DPI over time

a. On the same grid, plot each of the percents as a function of time. Do the trends seem to be the same for all expenditures?

b. Which expenditure might cause people the most concern recently? Why?

16. Write a paragraph summarizing the trends in disposable personal income and major personal expenditures during the years for which data are provided.

LESSON 11

Exponents and Growth

How many people live in your town?

Has the population of your town changed over the last ten years?

There are over 3 million people in Chicago, Illinois. How does your town's population compare to Chicago's?

How many people are on Earth?

How many people do you think will be in the United States in the year 2000?

When the birth rate is greater than the death rate, the population increases and may lead to overpopulation. Eventually living space, food, and natural resources are not sufficient to support the increased population. What are some measures governments take to control population? When would a sizeable decrease in population create problems?

OBJECTIVES

Use exponential functions as models of population growth.

Compare exponential growth and linear growth.

INVESTIGATE

Incredible Growth Rates

The article that follows appeared in newspapers across the nation in the spring of 1993. It describes the differences in growth rates of countries around the world.

World Population Growing at Record Pace

By Nick Ludington, ***The Associated Press,*** May 12, 1993

WASHINGTON—World population is growing at the fastest pace ever and virtually all growth is in the Third World, according to a survey released Tuesday by a research group.

The annual survey at the Washington, D.C.-based Population Reference Bureau said, "We are at a point where, except for the United States, population growth is essentially a Third World phenomenon."

The survey predicted world population will reach 5.5 billion by mid-1993, 40 percent of it in China and India. The bureau said population is growing each year by 90 million, roughly the population of Mexico.

Carl Haub, a demographer who worked on the study, said that world population will grow to 8.5 billion by the year 2025, "only if birth rates continue to come down as expected. If they don't, growth will be even faster."

In an interview, Haub said world population took 15 years to increase by 1 billion to its 4 billion total in 1975 and 12 years to increase to 5 billion in 1987. It is expected to take only 10 years to rise to 6 billion in 1997.

Haub said that if it were not for the relatively high U.S. increase, all growth in the world would be in poor areas.

The survey showed the United States with a growth rate of 0.8 percent a year. This rate "and the world's highest amount of immigration will now produce unexpected high growth," it said.

"With a net immigration of about 900,000 per year, the United States effectively absorbs 1 out of every 100 people added to world population each year."

Haub said that most of the immigrants to the United States are from the Third World.

Europe's population is virtually stagnant, with growth of 0.2 percent a year. "This virtually guarantees population decline by the turn of the century," Haub said. Several former Communist states, including Hungary and Bulgaria, already show negative growth.

States of the former Soviet Union have been growing at 0.6 percent. But there was a wide gap between Russia and the Ukraine, where population is declining, and the Muslim republics of central Asia, which are growing at more than 2 percent.

The world's fastest-growing area is the poorest: sub-Saharan Africa with growth of 3 percent a year, meaning population will double in 20 years. Latin America is growing at 1.9 percent.

In much of Africa and Latin America, income in the past 10 years grew more slowly than population, according to World Bank figures. So the average citizen ended the decade poorer.

Asia is growing at a rate of 1.7 percent. But without China, whose strict and controversial birth-control program is responsible for a sharp drop in growth to 1.2 percent, Asia's rate is 2.1 percent. The lowest growth rate in Asia is Japan's 0.3 percent.

Source: *The Associated Press,* Gainesville, Florida, Vol. 117, No. 306, May 12, 1993

Discussion and Practice

1. Answer the following questions based on the preceding article.

a. What is meant by *population growth rate?*

b. What is the fastest growing area in the world?

c. How does the United States's growth rate compare with that of other nations?

d. Bulgaria has a "negative growth." What do you think this means?

The estimated world population is given in Table 11.1 for 5-year intervals from 1960 to 1990.

Table 11.1
World Population

Year	Population (millions)
1960	3,049
1965	3,358
1970	3,721
1975	4,103
1980	4,473
1985	4,865
1990	5,380

Source: *Statistical Abstract of the United States,* 1994

2. Plot (year, population) using the data in Table 11.1.

a. Does the plot appear to be a straight line? Explain why or why not.

b. Find the amount of increase for each five-year period. What observations can you make about the amounts?

3. a. From 1960 to 1965, the increase in population was 309 million. What percent increase is that?

b. Add a new column to the table, and for each of the 5-year periods calculate the population increase as a percent of the year at the beginning of the period. Is there any pattern among these percent increases?

c. Based on your work, what would be a reasonable figure to quote as the typical five-year growth rate for the world's population?

d. The article says that the world's population is growing by about 90 million people per year. Do the data confirm this?

4. The article projects the world population in 2025 to be 8.5 billion. Use your work to confirm or reject this as a reasonable approximation.

5. Use the growth rate you found in problem 3c and the world population for 1990.

   a. When will the population double?

   b. Will it take the same amount of time to double again? How did you find your answer?

   c. Write a brief summary of what you have learned about the world's population and its growth.

   d. The article indicates that sub-Saharan Africa with a growth of 3% per year will double its population in 20 years. Do you think this is possible? Why or why not?

This type of population growth is called *exponential growth*.

6. A formula for the population growth function can be $P = 993.65(1.02^x)$ where x is the last two digits of the year (19)70, (19)75, . . . and P is in millions.

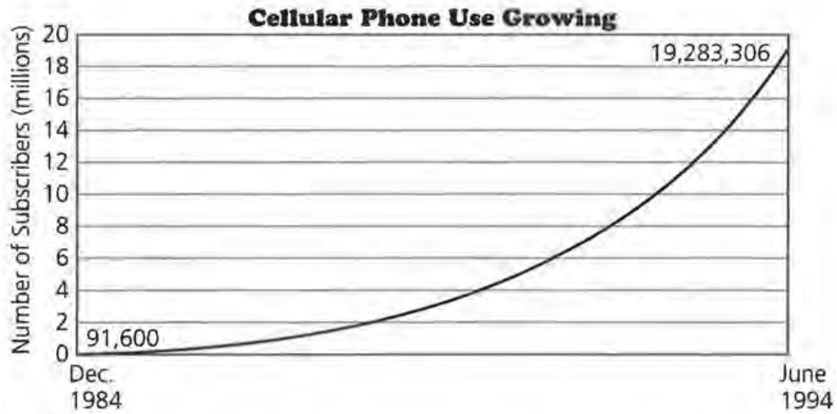
   a. Use the formula and your calculator to estimate what the population was in 1987 and in 1995.

   b. Graph the formula on the plot you made of the data. How well does the graph seem to fit the data?

   c. What will happen to the function and its graph if the formula is written $P = 993.65\ (1.02^x)$?

   d. What does the formula $P = 993.65(1.02^x)$ predict for the population in 2010? Do you think this prediction is reasonable?

7. Cellular-phone use is growing rapidly. A newspaper article contained the following information: "The number of cellular-phone subscribers has grown in the past decade from fewer than 100,000 in 1984 to more than 19 million. The number is increasing by 17,000 per day."A graph with more specific numbers follows.



a. Refer to the graph above and describe the growth of cellular-phone usage.

b. The cellular-phone data for the years 1986 to 1993 from Cellular Telecommunications Industry Association, Washington, D.C., State of the Cellular Industry, is given in Table 11.2. A formula for the function is $N = (1.96)(10^{-14})(1.56^x)$, where x is the last two digits in the year and N is the number of cellular-phone subscribers (in millions). How well does the formula work? Explain what you did to make your conclusion.

Table 11.2
Cellular-Phone Subscribers

Year	Number of Subscribers (1,000s)
1986	682
1987	1,231
1988	2,067
1989	3,509
1990	5,283
1991	7,557
1992	11,033
1993	16,009

Source: *The American Almanac*, 1995

c. Does the data point for 1994 given in the article and graph fit in the formula?

d. Draw a graph using the formula in part b. How does your graph compare to the graph above?

e. Is the statement, "The number is increasing by 17,000 per day," correct? Explain your answer.

Summary

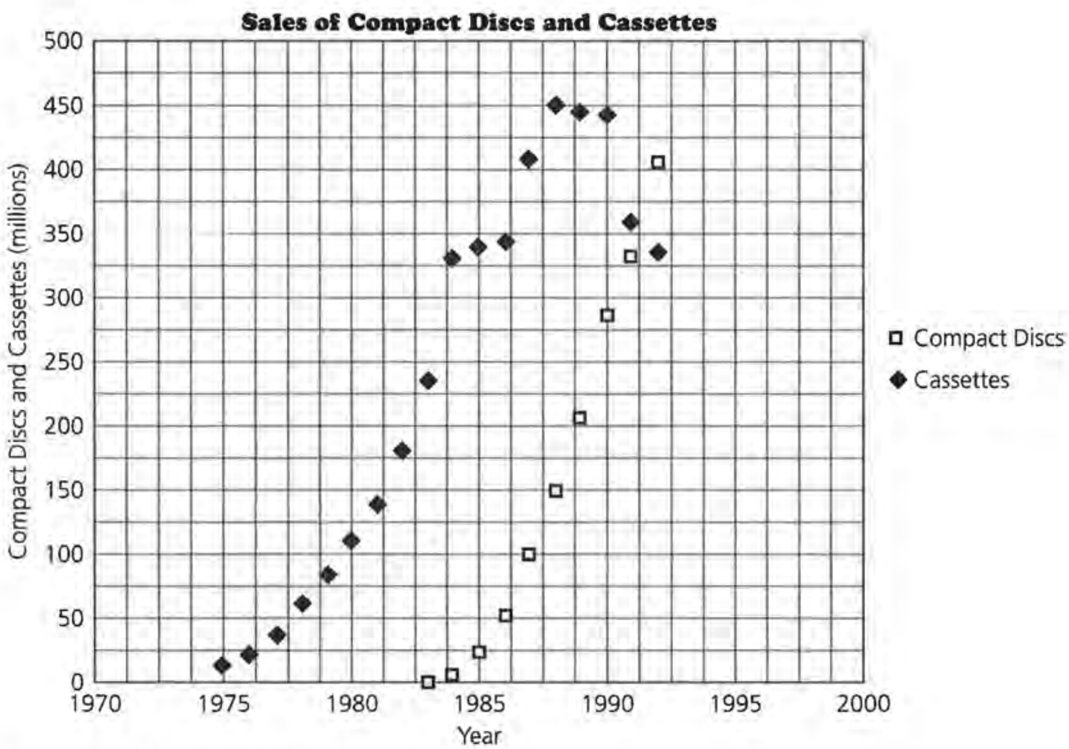
Some formulas describe exponential growth. When a quantity is growing exponentially, the rate of change becomes increasingly greater. The *percent* of increase is the same in each time period. You can study exponential growth using tables, graphs, or formulas. Population growth is one example of exponential growth.

Practice and Applications

8. Consider the two functions $y = 2x$ and $y = 2^x$.

- **a.** Graph them on the same set of axes and compare the two graphs.
- **b.** Make a table of values for the two functions and compare the values.
- **c.** The function $y = 2^x$ is an *exponential function.* Is there any set of values for x for which $y = 2x$ is greater than $y = 2^x$? Explain how you know.

9. The following graph represents the sales of compact discs and cassettes, as reported by retailers.



Source: Data from *USA Today,* July 14, 1995

- **a.** Describe both plots.

b. Does either of the plots seem to display exponential growth? Explain your answer.

10. The cost of health care is a serious issue and a matter of great debate in the United States. In 1970, spending on health care amounted to 74.4 billion dollars.

a. Suppose health-care spending had grown about 10.5 billion each year. How much had it been in 1975?

b. Make a graph and a table to show what the cost would be by the year 2000 if the change remained constant at $10.5 billion each year.

c. The approximate amounts spent on heath care for 10-year intervals and a predicted amount for the year 2000 are in Table 11.3. Plot the data on the same set of axes you used for part b. Did the spending increase at a constant rate? Explain how you know.

Table 11.3
Health-Care Spending

Year	Health-Care Spending (billions of dollars)
1970	74.4
1980	220.5
1990	650.0
2000	1,613.0

Source: *USA Today*, January 30, 1995

11. A rule for the function that describes the amount spent on health care is

$A = .0566(1.1^x)$, where x is the last two digits of the year (70, 80, 90, . . .).

a. How is this rule different from $A = 0.0566(1.1x)$?

b. Use the rule to generate a table to estimate the amount spent or predicted for health care from 1990 to 2000. How closely do the values match those given in the table?

c. Explain how you can tell that the graph of this rule is not a straight line.

12. Use an almanac to find a history of the population of your state or community. Was its growth exponential or not? Show the work and thinking you did to make your decision.

13. Explain what the following statement means: "When something increases exponentially, the percent of increase in each time period is the same."

Percents, Proportions, and Graphs

Who will support Social Security as the number of recipients increases faster than the number of workers? Whose salaries will pay the cost?

How will the country provide health care for so many people in poor health due to old age?

How will the increasing number of retired people affect society?

OBJECTIVES

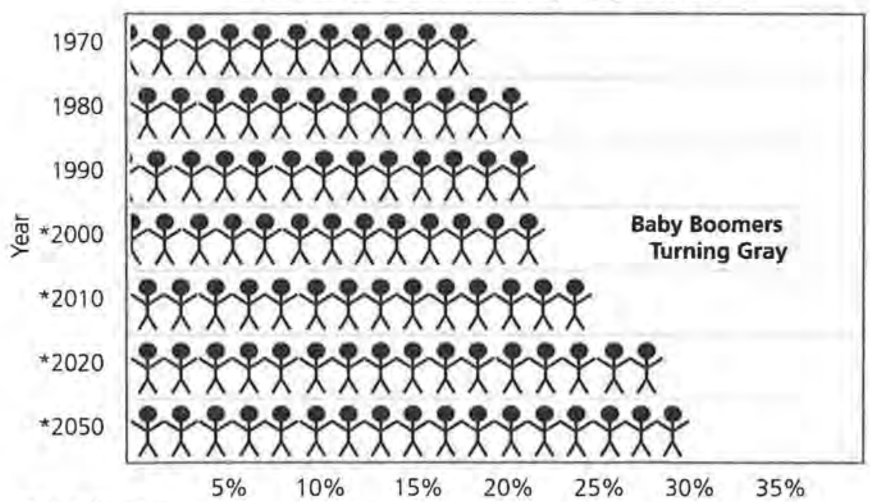
Work with inequalities.

Calculate and plot proportions.

Investigate cumulative-proportion plots.

One of the issues facing modern society is how to care for the elderly. As the "baby boomers" age, the percent of Americans aged 55 and over rises. Many news stories reflect concern over the problems posed by the aging of our society. The graph below shows how the United States population is aging.

Percent of Americans Ages 55 and Over



*Projected numbers

Source: United States Bureau of the Census, 1990

INVESTIGATE

What Is Meant by "The Graying of America"?

Is the population of the United States really getting older? Table 12.1 shows age distributions of the country's populations in 20-year intervals since 1900. The figures for the year 2000 are projections. Study these data for patterns or trends.

Table 12.1
Population Age Distribution

Age	1900	1920	1940	1960	1980	1992	2000
Under 5	0.121	0.109	0.080	0.113	0.072	0.076	0.066
5–14	0.223	0.208	0.169	0.198	0.153	0.143	0.143
15–24	0.196	0.177	0.182	0.136	0.188	0.142	0.135
25–34	0.160	0.164	0.162	0.1127	0.165	0.166	0.136
35–44	0.122	0.135	0.140	0.134	0.144	0.156	0.163
45–54	0.084	0.099	0.118	0.114	0.100	0.107	0.138
55–64	0.053	0.062	0.081	0.086	0.095	0.082	0.089
65 and over	0.041	0.046	0.068	0.092	0.113	0.127	0.130

Source: *Statistical Abstract of the United States,* 1991, 1995

Discussion and Practice

1. Refer to Table 12.1.

a. What information is given by the first entry, 0.121?

b. What was the population proportion of 15- to 24-year-olds in 1900 and how has this proportion changed over the years?

c. How has the proportion of those over the age of 65 changed over the years?

2. The data in Table 12.1 are population proportions describing age categories. Thus, some of the kinds of graphs and data summaries used earlier cannot be used here.

a. Can you make a number-line plot of the ages of people in the year 1900? Why or why not?

b. Can you accurately calculate the mean age of the population in 1980? Why or why not?

You can use the information in Table 12.1 to answer questions of the form, "What proportion of the population was under the age of 35 in 1900?" Let the variable A represent the age in years. For the data in the table, A begins at age 0 but does not have a defined *upper bound,* as the last age category is simply

65 and over. For simplicity, define the upper bound of A as 100. Certainly there were people over the age of 100 in these various populations, but their proportion was very small. Now, the phrase "under the age of 35" can be written symbolically as "$A < 35$."

3. Find the proportions of the 1900 population and the 1980 population for each interval.

a. $0 \le A < 5$

b. $15 \le A < 35$

c. $A \ge 45$

d. Use these figures to comment on the differences between the age distributions of the populations for 1900 and 1980.

Instead of looking at proportions in each age category given, the *cumulative* form of a table gives the proportions less than various ages. Based on Table 12.1, Table 12.2 is a cumulative table in which each entry is the *sum* of the entries in rows above it.

Table 12.2
Cumulative Age Distribution

A	1900	1920	1940	1960	1980	1992	2000
< 5	0.121	0.109	0.080	0.113	0.072	0.076	0.066
< 15	0.344	0.317	0.249	0.311	0.225	0.219	0.209
< 25	0.540	0.494	0.431	0.447	0.413	0.361	0.344
< 35	0.700	0.658	0.593	0.574	0.578	0.527	0.480
< 45	0.822	0.793	0.733	0.708	0.692	0.683	0.643
< 55	0.906	0.892	0.851	0.822	0.792	0.790	0.781
< 65	0.959	0.954	0.932	0.908	0.887	0.872	0.870
< 100	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Source: *Statistical Abstract of the United States,* 1991, 1995

4. Consider Table 12.2.

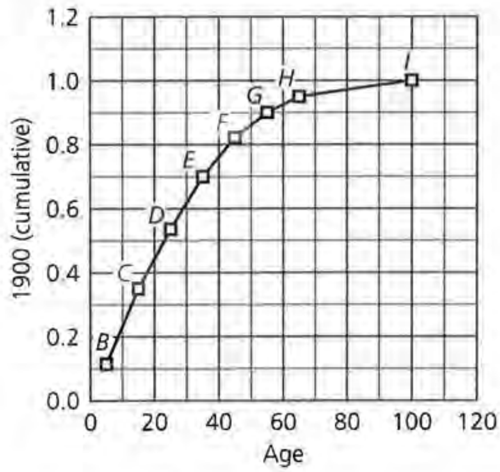
a. What information is given by the 0.344 under 1900?

b. Find the entry that corresponds to $A < 25$ for 1960.

c. Using the table, how could you find the proportions age 65 and over in each year?

d. Using the table, how could you find the proportions of the population between the ages of 25 and 44 for each year?

Differences among the age distributions over the years can be seen by plotting *cumulative proportion* as a function of the age *A*. The figure below shows a plot for 1900.



5. Study the graph above.

a. What information is given by point *E*?

b. Describe the pattern you observe.

c. Compare the change between *B* and *C* with the change between *G* and *H*.

d. Is it appropriate to connect the points with straight-line segments? Why or why not?

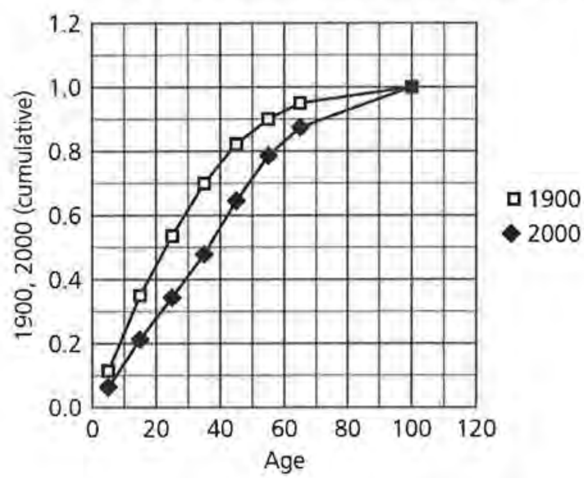
6. What age corresponds to

a. the youngest 25% of the population?

b. the youngest 50% of the population?

c. the median age? Explain your answer.

7. This graph represents plots of the cumulative proportions for the year 1900 and for the year 2000, projected.



a. Why is the first point in the plot for the year 2000 lower than the first point for the year 1900?

b. Describe the difference in the population proportion for the two plots.

8. In groups, construct cumulative-proportion plots for the other five years. Each year may be assigned to a different group of students; if done by different groups, share your results and construct all the plots on the same grid. Use whatever technology is available.

a. Study the five cumulative-proportion plots, and describe any patterns in the way the age distribution of the population of the United States is changing.

b. Can a cumulative frequency graph slope downward? Why or why not?

c. Why does the graph of cumulative proportions as a function of age have to be nondecreasing?

9. Refer to Table 12.2.

a. Approximate the median age for each year in the data table.

b. Describe how the median has changed over time.

10. Refer to either Table 12.1 or Table 12.2.

a. Compare the mean age of the population in one of the tables to the mean age projected for the year 2000. Describe your method for finding the mean.

b. How do the means for those years compare with the medians? Why is this the case?

11. Summarize your findings on the shifting age distribution of the U.S. population. Is there evidence to suggest that the problems mentioned in the opening discussion are serious?

Summary

Proportions are useful for comparing quantitative information from investigations having different numbers of elements. The proportion of something in a large population can be similar to the proportion in a small population. When categories are ordered, such as income categories or age categories, it is convenient and useful to work with cumulative proportions.

Practice and Applications

12. Is there a difference in how much education people from different countries have? Table 12.3 contains information regarding the proportions of levels of education attained by adults ages 25–64.

Table 12.3
Education Levels

Country	No High-School Diploma	High-School Diploma	College or Postsecondary Degree
Australia	.44	.25	.31
Austria	.33	.61	.07
Belgium	.57	.24	.20
Canada	.24	.36	.40
Denmark	.39	.43	.18
France	.49	.35	.15
Germany	.18	.60	.22
Ireland	.60	.24	.16
Italy	.72	.22	.06
Netherlands	.44	.37	.20
Portugal	.93	.03	.04
Spain	.78	.12	.10
Turkey	.82	.11	.06
United Kingdom	.35	.49	.16
United States	.17	.47	.36

Source: *Statistical Abstract of the United States*, 1995

a. What proportion of adults ages 25–64 in Austria have at least a high-school diploma?

b. Which country has the least percent of adults ages 25–64 with a high-school diploma?

c. Why is a proportion a useful way to record this information?

d. Which country would you say has the population with the most formal education? Justify your choice.

13. What are the age distributions for different countries? A cumulative-proportion table for several countries is given in Table 12.4.

Table 12.4
Comparative Age Distributions by Country

Age A	Argentina	Germany	South Korea	United States
Under 5 years	.09	.06	.19	.08
Under 15 years	.28	.17	.48	.22
Under 65 years	.90	.85	.97	.87
Under 100 years	1.00	1.00	1.00	1.00

Source: *The American Almanac,* 1994–95

a. What proportion of the people in Germany is in each of the two age brackets given below?

$$5 < A \leq 14; \; A \geq 65$$

b. In which country is there the greatest difference between the number of young people and the number of old people? How did you decide?

c. How does the age distribution in the United States compare to those of the other three countries?

14. Table 12.5 gives the proportion of footwear sold in 1992 to people of different ages.

Table 12.5
Footwear

Age A	Aerobic Shoes	Gym Shoes, Sneakers	Jogging/ Running Shoes	Walking Shoes
$A < 14$ years	.06	.42	.09	.03
$14 \leq A \leq 17$ years	.04	.17	.09	.02
$18 \leq A \leq 24$ years	.08	.05	.11	.03
$25 \leq A \leq 34$ years	.26	.09	.19	.10
$35 \leq A \leq 44$ years	.25	.11	.21	.18
$45 \leq A \leq 64$ years	.22	.12	.25	.36
$A \geq 65$ years	.09	.04	.06	.28

Source: *Statistical Abstract of the United States,* 1994–1995

a. What proportion of those who purchased aerobic shoes were under 25?

b. Suppose you were targeting an advertising campaign for people ages 18 to 34 because you knew they had money to spend. What kind of shoes would you feature in your campaign? Why did you select that kind?

c. Do row totals make sense? Why or why not?

d. Describe the changes in shoe preference as people grow older.

15. On the same set of axes, make a cumulative-proportion plot for each type of shoes with the age category along the horizontal axis.

a. Describe any patterns you see.

b. Approximate the median age for each shoe type. What do these numbers tell you?

Extension

The heights of males and females between the ages of 18 and 24 have the cumulative percents shown in Table 12.6.

Table 12.6
Cumulative Heights

Males		Females	
Height (inches)	Cumulative Percent	Height (inches)	Cumulative Percent
61	0.18	56	0.05
62	0.34	57	0.43
63	0.61	58	0.94
64	2.37	59	2.22
65	3.85	60	4.22
66	8.24	61	9.13
67	16.18	62	17.75
68	26.68	63	29.06
69	38.89	64	41.81
70	53.66	65	58.09
71	68.25	66	74.76
72	80.14	67	85.37
73	88.54	68	92.30
74	92.74	69	96.23
75	96.17	70	98.34
76	98.40	71	99.38

Source: *Statistical Abstract of the United States*, 1991

16. Construct plots of cumulative percent as a function of height. Draw both plots on the same grid.

a. How do the two distributions of heights differ?

b. Approximate the median height for males and the median height for females. Explain how you found your approximations.

c. Approximate the mean height for males and the mean height for females. Explain how you found your approximations.

d. Write a summary paragraph on heights of males and of females, and how the heights compare.

ASSESSMENT

Driving Records

Have you ever wondered about insurance rates for driving?

Are men really better drivers than women?
Are young people better drivers than older people?

OBJECTIVES

Use rates, information from tables, and cumulative proportions to make decisions.

Table A2.1 contains information about automobile drivers in crashes during 1994 in the state of Wisconsin.

Table A2.1
Drivers in Car Crashes

Age of Driver	Licensed Drivers	Drivers Involved in Crashes	Drivers in Fatal Crashes	Drivers in Injury Crashes	Drivers in Property-Damage Crashes
14 and under	0	224	2	108	114
15	0	375	1	152	222
16	39,505	8,290	25	2,883	5,382
17	55,144	8,627	32	3,048	5,547
18	56,965	8,487	40	3,081	5,366
19	56,802	7,378	34	2,690	4,654
20	59,116	6,698	23	2,376	4,299
21	58,474	6,173	31	2,131	4,011
22	61,883	6,435	27	2,190	4,218
23	65,079	6,541	28	2,198	4,315
24	72,337	6,221	23	2,111	4,087
25–34	770,177	55,797	239	18,418	37,140
35–44	797,171	45,114	176	14,454	30,484
45–54	565,409	27,030	106	8,465	18,459
55–64	381,391	15,702	96	4,906	10,700
65–74	319,438	11,085	58	3,610	7,417
75–84	166,417	6,133	43	2,020	4,070
85 and over	28,695	1,109	9	375	725
Unknown	0	21,830	18	3,407	18,405
Total	3,554,003	249,249	1,011	78,623	169,615

Source: *Wisconsin Traffic Crash Facts,* Wisconsin Department of Transportation 1994

1. Refer to Table A2.1.

   a. Which age group has the greatest percent of licensed drivers? Greatest percent of drivers in crashes? Greatest percent of crashes?

   b. Explain what these three percents tell you and how they are different in each case.

2. Use the data in Table A2.1.

   a. Calculate the total crash rate per 1,000 licensed drivers for the age categories 16–19, 20–24, 25–34, 35–44, 45–54, 55–64, 65–74, 75–84, and 85+. Write a formula that will explain how you made your calculations.

   b. Make a plot for (age categories, total crash rates). What observations can you make from the plot?

3. Use the data in Table A2.1.

   a. Plot the cumulative percents of drivers as a function of age for just those drivers involved in crashes. Describe the plot.

   b. Estimate the median age of those drivers involved in crashes. Explain how you made your estimate.

   c. Estimate the mean age of drivers involved in crashes. Describe the procedure you used. How does your estimate for the mean age compare to that for the median age?

4. Is there any difference in the relationship between the ages of drivers involved in fatal crashes and those involved in property-damage crashes? Show work that will justify your answer.

5. Insurance premiums are based on the number of claims made against the premiums. Table A2.2 contains crash information by age and by sex for Wisconsin drivers in 1994. Use the information in the table and the techniques you have learned to decide which age groups should pay the highest premiums for crash insurance. Be sure to explain how you used the data to make your decision. You may include graphs in your argument.

   Note: Since accurate gender data is not available for all accidents, Table 2.1 may show different totals from those of Table 2.2.

Table A2.2
Male/Female Crash Rates

	Licensed Drivers		Drivers in Fatal Crashes		Drivers in Injury Crashes		Driver/Property-Damage Crashes		Total Number of Drivers in Crashes	
Age	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**
14	0	0	0	2	35	72	38	76	73	150
15	0	0	0	1	52	100	96	126	148	227
16	19,624	19,881	11	14	1,358	1,525	2,200	3,182	3,569	4,721
17	27,070	28,074	18	14	1,381	1,667	2,195	3,352	3,594	5,033
18	27,436	29,529	16	24	1,232	1,849	1,905	3,461	3,153	5,334
19	27,465	29,337	5	29	1,071	1,619	1,749	2,905	2,825	4,553
20	28,477	30,639	5	18	951	1,425	1,561	2,738	2,517	4,181
21	28,745	29,729	8	23	844	1,287	1,527	2,484	2,379	3,794
22	30,339	31,544	4	23	901	1,289	1,662	2,556	2,567	3,868
23	32,073	33,006	8	20	947	1,251	1,601	2,714	2,556	3,985
24	35,643	36,694	9	14	882	1,229	1,585	2,502	2,476	3,745
25–34	381,312	388,865	53	186	7,584	10,834	13,835	23,305	21,472	34,325
35–44	392,865	404,306	51	125	6,071	8,383	11,741	18,742	17,863	27,250
45–54	278,117	287,292	30	76	3,499	4,966	6,698	11,760	10,227	16,802
55–64	188,056	193,335	26	70	1,765	3,141	3,514	7,186	5,305	10,397
65–74	159,812	159,626	10	48	1,346	2,264	2,591	4,826	3,947	7,138
75–84	86,339	80,078	12	31	819	1,201	1,650	2,420	2,481	3,652
85 and over	13,794	14,901	0	9	138	237	248	477	386	723
Unknown	0	0	0	5	177	779	478	2,087	655	2,871
Total	1,757,167	1,796,836	266	732	31,053	45,118	56,874	96,899	88,193	142,749

Source: *Wisconsin Traffic Crash Facts,* Wisconsin Department of Transportation 1994

6. In general, are males or females involved in more crashes? Defend your answer based on the data in Table A2.2.

Dale Seymour Publications® is a leading publisher of K–12 educational materials in mathematics, thinking skills, science, language arts, and art education.

DALE
SEYMOUR
PUBLICATIONS®

9 781572 322301
90000
ISBN 1-57232-230-6
21173