ALGEBRA

TEACHER'S EDITION

Exploring Symbols

G. BURRILL, M. CLIFFORD, R. SCHEAFFER

DATA-DRIVEN MATHEMATICS

DALE SEYMOUR PUBLICATIONS®

Exploring Symbols: An Introduction to Expressions and Functions

TEACHER'S EDITION

DATA-DRIVEN MATHEMATICS

Gail F. Burrill, Miriam Clifford, and Richard Scheaffer

Dale Seymour Publications®
White Plains, 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: Alan MacDonell

Senior Math Editor: Carol Zacny

Project Editor: Nancy Anderson

Production/Manufacturing Director: Janet Yearian

Senior Production Coordinator: Fiona Santoianni

Design Manager: Jeff Kelly

Text and Cover Design: Christy Butterfield

Composition: Joan Olson

Art: Carl Yoshihara

Cover Photo: Kirk Yarnell

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

Dale Seymour Publications
10 Bank Street
White Plains, NY 10602
Customer Service: 800–872–1100

Order number DS21175

ISBN 1-57232-234-9

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 the 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* vi

Using This Module vii

Unit I: Variables, Expressions, and Formulas

Introductory Activity: **Evaluating with Formulas** 3

Lesson 1: **Variables and Formulas** 6

Lesson 2: **Formulas That Manage Money** 18

Lesson 3: **Creating Your Own Formulas** 29

Lesson 4: **Expressions and Rates** 41

Lesson 5: **Rates, Frequencies, and Percents** 53

Lesson 6: **Formulas That Summarize Typical Values in Data** 63

Lesson 7: **Formulas That Summarize Variation in Data** 77

Lesson 8: **Comparing Measurements** 93

Assessment: **Cars** 103

Unit II: Functions

Introductory Activity: **Time on Task** 111

Lesson 9: **An Introduction to Functions** 114

Lesson 10: **Trends over Time** 124

Lesson 11: **Exponents and Growth** 132

Lesson 12: **Percents, Proportions, and Graphs** 141

Assessment: **Driving Records** 151

Teacher Resources

Quizzes 159

End-of-Module Test 167

Solutions to Quizzes and Test 173

Activity Sheets 183

Procedures for Using the TI-83 Graphing Calculator 191

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 helps teachers motivate, develop, and reinforce concepts taught in current texts. The materials incorporate the 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 the statistical concepts most important for students to know and the relationship of these concepts to the secondary mathematics curriculum.

A diagram of the modules and possible relationships to the curriculum is on the back cover of this Teacher's Edition.

Using This Module

Why the Content Is Important

Mathematical reasoning requires students to think abstractly so that a solution to one problem can be generalized to solve other similar problems. Symbolic expression enables students to engage in this abstract thinking. The reasoning process and communication of that process continues as symbols are used in expressions, equations, and functions.

In this module, lessons built around real data motivate and justify introducing students to symbols. A spreadsheet formulation of problems is used, so symbols acquire a natural meaning as column labels. Symbolic expressions then show how one column of data is related to another. Examples include rules for finding the total income of a student; common rates such as birth, death, disease, and accident rates; and a slugging percent in baseball. Students create their own formulas to explore situations and investigate formulas used for such things as calculating ratings in sports or in cities.

Mathematical Content

- 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

Statistical Content

- 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

Instructional Model

Each of the modules in *Data-Driven Mathematics* emphasizes discourse and student involvement. Each lesson is designed around a problem or mathematical situation and begins with a series of introductory questions or scenarios that can prompt discussion and raise issues about that problem. These questions can involve students in thinking about the problem and help them understand why such a problem might be of interest to someone in the world outside the classroom. The questions can be used in whole-class discussion or in student groups. In some cases, the questions are appropriate to assign as homework with some family involvement.

These questions are followed by discussion issues that clarify the initial questions and begin to shape the direction of the lesson. Once the stage has been set for the problem, students begin to investigate the situation mathematically. As students work their way through the investigations, it is important that they have the opportunity to share their thinking with others and to discuss their solutions in small groups and with the entire class. Many of the exercises are designed for groups in which each member does one part of the problem and the results are compiled for final analysis and solution. Multiple solutions and solution strategies are also possible; it is important for students to recognize these situations and to discuss the reasoning behind different approaches. This will provide each student with a wide variety of ways to build his or her own understanding of the mathematics.

In many cases, students are expected to construct their own understanding by thinking about the problem from several perspectives. They do need, however, validation of their thinking and confirmation that they are on the right track, which is why discourse among students, and between students and teacher, is critical. In addition, an important part of the teacher's role is to help students link the ideas within an investigation and to provide an overview of the "big picture" of the mathematics within the investigation. To facilitate this, a review of the mathematics appears in the Summary following each investigation.

Each investigation ends with a practice section in which students can revisit ideas presented within the lesson. These exercises may be assigned as homework, given as group work during class, or omitted altogether if students are ready to move ahead.

Periodically, student assessments occur in the student book. These can be assigned as long-range take-home tasks, as group assessment activities, or as in-class work. The assessment pages provide a summary of the lessons up to that point and can serve as a way for students to demonstrate what they know and what they can do with the mathematics. Commenting on the strategies students use to solve a problem can encourage students to apply different strategies. Students also learn to recognize those strategies that enable them to find solutions efficiently.

Teacher Resources

At the back of this Teacher's Edition are the following:

- Quizzes for selected lessons and an End-of-Unit Test
- Solution Key for quizzes and test
- Activity Sheets
- Procedures for Using the TI-83 Graphing Calculator

The *Materials* section at the beginning of each lesson in the Teacher's Edition references all Activity Sheets to be used with that lesson and any Quizzes or Test to be used following the lesson.

Where to Use the Module in the Curriculum

This module is appropriate whenever students are beginning to formalize their work with variables. It could be used as a first or second unit in an algebra class. The module should be used before the *Data-Driven Mathematics* module, *Exploring Linear Relations*, to provide students with the background necessary to study linear equations. Some of the lessons can be used independently of the others. Lessons 4 and 5 could be used to develop students' understanding of rates. Lessons 6, 7, and 8 on statistical concepts, could be used in any introductory statistics course. Lessons 9, 10, 11, and 12 could be used in teaching graphing in any course. Many of the ideas contained in this module are not usually taught explicitly in traditional mathematics classes. The concepts, however, and the real-world contexts from which they come present important and significant ways to think quantitatively about common situations.

Prerequisites

Students should have worked with ratios and percents and should be familiar with stem-and-leaf plots, box plots, and graphing points in a plane. For those who might need review, a possible resource is *Exploring Data* by Landwehr and Watkins, Dale Seymour Publications, 1995.

Pacing/Planning Guide

The table below provides a possible sequence and pacing for the lessons.

LESSON	OBJECTIVES	PACING
Unit I: Variables, Expressions, and Formulas		
Introductory Activity: Evaluating with Formulas	Introduce variables and formulas.	1/2 class period
Lesson 1: Variables and Formulas	Identify and use variables in formulas; evaluate expressions; use a spreadsheet.	2 class periods
Lesson 2: Formulas That Manage Money	Identify and use variables in formulas; evaluate expressions; use a spreadsheet.	2 class periods
Lesson 3: Creating Your Own Formulas	Use formulas written in words; use variables to develop formulas for ranking; translate sentences into symbols.	2 class periods
Lesson 4: Expressions and Rates	Understand rates in numerical and symbolic form; understand the importance of units of measurement; combine rates by using weighted averages.	1 class period
Lesson 5: Rates, Frequencies, and Percents	Apply rates in practical situations; consider time as a variable and to look at time plots; change from rates to counts and percents.	2 class periods
Lesson 6: Formulas That Summarize Typical Values in Data	Use formulas to summarize typical values; use symbolic notation as an efficient method of communication.	2 class periods
Lesson 7: Formulas That Summarize Variation in Data	Use formulas to describe the spread of variability in a set of data; use inequalities to describe unusual values; compare two formulas graphically and numerically.	3 class periods
Lesson 8: Comparing Measurements	Use scale data for making comparisons; use a formula to standardize scores.	2 class periods
Assessment: Cars	Create and use formulas; organize and interpret data.	1/2 to 1 class period
Unit II: Functions		
Introductory Activity: Time on Task	Use tables and graphs to represent functional relationships.	1/2 class period

LESSON	OBJECTIVES	PACING
Lesson 9: An Introduction to Functions	Develop the notion that one variable depends on another; represent relationships with tables, graphs, and symbols; investigate several different functional relationships.	1 class period
Lesson 10: Trends over Time	Develop piecewise formulas from information conveyed in words or data tables; recognize linear trends; use a ratio to express the slope in a linear relation.	1 class period
Lesson 11: Exponents and Growth	Use exponential functions as models of population growth; compare exponential growth and linear growth.	2 class periods
Lesson 12: Percents, Proportions, and Graphs	Work with inequalities; calculate and plot proportions; investigate cumulative proportion plots.	1 class period
Assessment: Driving Records	Use rates, information from tables, and cumulative proportions to make decisions.	1/2 class period

Technology

A scientific calculator is necessary for this module; a computer with spreadsheet software or a graphing calculator with list capabilities would be very useful. Instructions for using a TI-83 graphing calculator are included throughout the Teacher's Edition. Also, a graphing calculator section, entitled *Procedures for Using the TI-83,* is included at the end of this module. The data are available on a disk that accompanies the module. You can use the disk with spreadsheet software, or you can enter the data into a graphing calculator and have students use the link to transfer the data efficiently.

Data Disk List and Resource Materials

The table below lists the data sets available on the disks (one IBM disk and one MacIntosh disk) that accompany this Teacher's Edition.

LESSON	DATA SETS ON DISKS	RESOURCES
Unit I: Variables, Expressions, and Formulas		
Introductory Activity: Evaluating with Formulas		
Lesson 1: Variables and Formulas	Baseball Sluggers High-School Football Ranking NFL All-Time Passing Leaders After 1993–1994 1993 National Football League Passing Leaders	*Activity Sheet 1*
Lesson 2: Formulas That Manage Money	Payroll Records Wages and Commissions Gross Salary Weekly Payroll Weekly Car Rentals	*Activity Sheet 2* *Lesson 2 Quiz*
Lesson 3: Creating Your Own Formulas	Austin College Enrollment City Ratings Rating High-School Football Rating Cities Final Cities Ranking	*Lesson 3 Quiz*
Lesson 4: Expressions and Rates	Safe Driving Gas Mileage States and Safe Driving Driving Data	*Activity Sheet 3*
Lesson 5: Rates, Frequencies, and Percents	Death Rates for Leading Causes of Death for All Ages Death Rates for the Age Group 15–24 Death Rates from HIV Infections	
Lesson 6: Formulas That Summarize Typical Values in Data	Endangered Species Threatened Species Sports Drinks Quiz Scores Car Prices and MPG Family Size Track Meet Points	
Lesson 7: Formulas That Summarize Variation in Data	Life Expectancy Dimensions of Desktops Table 7.3: Batting Averages of American League Batting Champions	*Activity Sheet 4* *Lesson 7 Quiz*

LESSON	DATA SETS ON DISKS	RESOURCES
Lesson 8: Comparing Measurements	SAT/ACT Scores Recreation and Jobs for Top-Ten U.S. Cities SAT/ACT Scores Extended NBA Basketball Hall of Fame, 1990–1994	*Activity Sheets 5 and 6; Quiz to follow Lesson 1, 4, 5, 7, or 8 (Extension)*
Assessment: Cars	Cars: Typical Price/Highway Gas Mileage Cars: Typical Price by Number of Airbags	
Unit II: Functions		
Introductory Activity: Time on Task		
Lesson 9: An Introduction to Functions	Reaction Distance Speed and All Three Distances	*Activity Sheet 7*
Lesson 10: Trends over Time	Cat/Human Age Per-Capita Expenditures	
Lesson 11: Exponents and Growth	World Population Cellular-Phone Subscribers Health-Care Spending	
Lesson 12: Percents, Proportions, and Graphs	Population Age Distribution Cumulative Age Distribution Education Levels Comparative Age Distributions by Country Footwear Cumulative Heights	
Assessment: Driving Records	Drivers in Car Crashes Male/Female Crash Rates	*End-of-Module Test*

Grade Level/Course

The module is appropriate for students in grades 8–10 in the latter part of a pre-algebra course, in an introductory algebra course, or in an algebra course. The lessons might be used in conjunction with an integrated course in a variety of ways.

UNIT I

Variables, Expressions, and Formulas

INTRODUCTORY ACTIVITY

Evaluating with Formulas

Pacing: 1/2 class period

Overview

This activity is designed to have students begin to think about the factors or variables that are involved in a situation and how to quantify these factors. The process of giving grades contains many of the concepts that will be emphasized throughout the lessons. If you have already given students a formula for their grades, ask them to think of one of their own. If they think your formula is fair, students can use it; but be sure they have considered it seriously before they agree.

Teaching Notes

Students should recognize that factors such as test grades, quiz grades, class contribution, homework, and projects are variables. They should begin to recognize how variables are involved in situations, as this is a major emphasis of this entire unit. Some students will have a hard time writing a formula; they will feel comfortable using subjective measures: how hard they try, whether they were absent or sick, how fast or slow they work. Try to emphasize how a formula or procedure will enable a teacher to judge everyone on the same criteria (Lessons 1, 2, and 3). Students may suggest something such as counting tests twice as much as quizzes and homework half as much. This is really weighting the variables differently, something that will be addressed in Lesson 3. The final grade is usually made from percents and frequencies (Lesson 5) and is some measure of a typical value, the mean or median (Lesson 6). Some discussion might take place on whether the results will be fair to everyone and whether they should be. This can lead to a discussion of variation in the data that make up a student's grade, and how such variation should be handled. For instance, should a student be penalized because illness led to a low test score? To compare grades among different courses, students can think about combining grades from different teachers who use different formulas and different criteria. Does a *B* in mathematics compare to a *B* in English? This leads to the discussion in Lesson 8, Comparing Measurements.

Solutions

1. Answers will vary. They might include test scores, quiz grades, projects, homework, class participation, and extra credit.
2. Students will suggest different formulas. One possibility is $\frac{T+Q+H}{N} = G$, where T is the total number of points earned on tests, Q is the total number of points earned on quizzes, H is the total number of points earned on homework, and N is the total number of points possible. Be sure students recognize the importance of defining what the variables they use represent; if they do not, the formula will not be useful to anyone else.
3. Students should recognize that a formula will grade everyone on the same basis. They may or may not think this is a valid point.
4. Answers will vary. This is an opportunity to have students see how others feel about the same issues and to recognize that there are different ways to approach the problem.
5. Have students compare their formula to yours if you have one. You may choose not to assign this problem if you prefer.

STUDENT PAGE 3

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?

Introductory Comments on the Lessons

Lessons 1 and 2 have both been designed to introduce students to the use of variables and formulas. You should select one of the two lessons for your class, depending on the interests of your students. Lesson 1 is centered around using formulas in a variety of sports settings, while Lesson 2 is centered around formulas that occur in situations dealing with jobs and money. Lesson 1 involves ratios and fractions to a greater extent than does Lesson 2, which has more work with percents.

In these lessons, students should learn the importance of using variables—either in words or symbols—as a shorthand way to communicate quantitative information. They should recognize that using sentences or a series of words is cumbersome in expressing a relationship. They should also begin to understand that a variable represents an unknown quantity that can take on different numerical values as the situation changes.

Combining variables and operations will yield expressions–new variables that may become formulas or a procedure for performing a series of computations. These formulas can be derived from the situation itself, or can be given as a standard procedure. Substituting values for the variables allows students to apply the formula for a specific instance, such as the slugging average of a baseball player or the rating of a quarterback. It is also possible to combine variables in several ways to produce formulas that are equivalent; that is, they give the same output when evaluated for a specific situation. The Σ symbol is used to represent a sum, and the word *function* is introduced informally.

Students learn that symbols are useful in managing data and communicating with technology. Data arranged in a spreadsheet need labels in order for the arrangement to be meaningful. Using technology such as a spreadsheet or a graphing calculator with list capabilities will allow students to efficiently and effectively use a formula.

Students will find it useful to be able to talk with each other as they work their way through the problems, particularly in the beginning sections. In several cases, they are expected to compare their formulas and outcomes, decide how they are different or alike, and select the one that seems most reasonable. Many of the problems have no right answer, and student discussion is very necessary to make different responses meaningful and useful for understanding the mathematics involved. Students will have had experience with formulas such as those for area, volume, and distance from their earlier work in mathematics. It is important for them to recognize that formulas are useful in other areas, as well.

Technology

While it is important to do some of the calculations by hand, if possible students should be allowed to use technology such as a spreadsheet or graphing calculator, or the work may become tedious and the mathematics lost in the routine. One of the objectives of the lessons is to introduce students to working with data in spreadsheet format. The data are available on the data disk that accompanies this module and can be downloaded into a spreadsheet or into a graphing calculator. To reduce frustration, be sure the data are accurate before students perform an extensive set of calculations. In many cases, students will probably have to redefine their variables in language compatible with the technology they are using. For example, batting averages may become List 1, or hourly wages the entry in cell B2. To avoid problems in this respect, have students write out a clearly defined list of the variables and what they represent in each case.

Follow-Up

You might have students think about other places in and out of school where variables are important, such as calculation of average daily attendance and formulas for computing state aid to schools, for paying taxes, and for ranking tennis players. They might search newspapers and magazines for data presented in tables, and write a report about the variables represented in those tables. They might also interview teachers from other disciplines about variables or formulas that are used in that discipline. You might also ask them to invent their own formulas, carefully defining the variables.

LESSON 1

Variables and Formulas

Materials: *Activity Sheet 1*
Technology: spreadsheet program or graphing calculators
Pacing: 2 class periods

Overview

Students are introduced to variables and formulas in the context of a variety of sports situations. They find the slugging average for baseball players and rank and rate quarterbacks. In each case, they consider the variables that might make a difference and use symbols to represent the variables. For the slugging average, students develop a formula that seems quite natural. To rate quarterbacks, they are given a formula that is very complicated and asked to consider what the variables are and how the coefficients might have been obtained. They are then asked to apply the formula to several quarterbacks. Students are informally introduced to the concept of function in language such as, "The slugging average is a function of the number of triples a player hits."

Students learn to use a spreadsheet or the list function of a graphing calculator to apply their formulas.

Teaching Notes

Students do not need to know much about baseball or football in order to do the lesson. Some students, however, will not be comfortable using the sports contexts and, if that is the case, use Lesson 2 as an option. Emphasize the notion of variables, what they represent, and how they can be described using symbols. Some of the quantities in the formulas are constant for that formula, such as the coefficient for the interceptions in the quarterback rating formula. Point out the necessity of using symbols of some form to enter formulas into either a spreadsheet or a calculator.

Some formulas have several forms, yet produce the same results. This introduces students to equivalent expressions and to some of the algebraic manipulations that produce equivalent expressions. In some cases, such as the quarterback-rating formula, it makes more sense to use the formula as given (where each variable is a function of the number of attempts) than to use an alternate form (where the final divisor is the number of attempts or where the coefficients look different only because of the difference in form). Students see that these situations exist and begin to understand something about how symbols operate.

Technology

Students can do the lesson with only a scientific calculator. One of the intentions of the lesson, however, is to introduce them to either a spreadsheet or the list function of a graphing calculator. The ability to do a repeated series of calculations by giving a single formula or command emphasizes the need for symbols and formulas and demonstrates how efficiently the process can be carried out. Be sure students are comfortable with using the technology and are successful at producing results. To help facilitate the work, you may want to pair students who have some technological skills with those who don't. You may also want to have some of the students demonstrate for the rest of the class how to use the spreadsheet or calculator.

Be sure that students record what they are saving in each list or column when they use a spreadsheet or calculator. An example would be:

$L_1 = WC \quad L_2 = TC \quad L_3 = WH \quad L_4 = TH \quad L_5 = OW$

This will help the students keep in mind what they put into each list.

If the students are unfamiliar with using a spreadsheet or the graphing calculator, some time may be needed here to familiarize them with the procedures involved.

If the students are using a TI-83 calculator, they must remember to put a parentheses around the entire numerator of a fraction.

If the students want to review their formula on the TI-83, it is necessary to go to the home screen and write the formula and then store the formula under the list name, for example, L_6.

If you are using the TI-83, you may want students to download the data using the link. An efficient way to do this is to appoint one student in each group as the calculator representative who will get the data from your calculator before class, and then transfer the data to each member of the group.

Follow-Up

You might have students think about other sports situations and how formulas are used in them. How do tennis players get rated? How are tennis matches scored? What kind of scores are given in gymnastics? Are there formulas for scoring dogs at dog shows? In each case, have students identify the variables involved, indicate how important the variables are, and then describe how the formulas work.

STUDENT PAGE 4

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

Solution Key

Discussion and Practice

1. **a.** Possible answers: Letters are quicker to write and are just as representative; letters take up less space in tables and formulas; letters help organize the information; letters can be a shortcut to represent many numbers.

b. The number of singles can be found by subtracting the total number of doubles, triples, and home runs from the total number of hits.

c. Bonds: 93; Ruth: 73; Mantle: 109

d. $S = H - D - T - HR$ or
$S = H - (D + T + HR)$

Most students will likely use the letters because it makes the formula quicker to read and write. Some students may use words so that the formula is clear to those who do not know what the letters stand for.

STUDENT PAGE 5

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.

STUDENT PAGE 6

2. **a.** A triple earns 3 bases, and a home run earns 4 bases.

b. $TB = S + 2D + 3T + 4HR$, where $S = H - D - T - HR$

c. $TB = 2 + 1(3) = 2 + 3 = 5$

3. **a.** Possible answers: Find the ratio of the total number of bases to the total number of hits; find the ratio of the number of extra-base hits to the total number of hits; find the ratio of the total number of bases for extra-base hits to the total number of bases.

Encourage students to share their methods.

b. Answers are given for each method suggested in part a.

Bonds: .486; .163; 2.017

Ruth: .576; .216; 2.256

Mantle: .420; .148; 2.000

c. Answers will vary.

d. Answers will depend on students' feelings about the information given by the average. Some may think that it is a reasonable method; others may feel that too much information is lost in the computation of the average.

4. **a.** Answers will vary. If students divide by the total number of hits, their method takes the number of singles indirectly into account. Students' formulas may have taken the number of strikeouts into account.

b. Answer depends on the method students used.

5. **a.** $\frac{TB}{AB} = \frac{388}{458} = .847$

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. .705

c. .677; Bonds's average was lower than Babe Ruth's.

d. Ruth, Mantle, Bonds; answers will vary.

6. **a.** $AB = 539 + 40 = 579$; $\frac{365}{579} = .630$

b. $AB = 665$; $\text{TB} = 365 + 126$; $\frac{491}{665} = .738$

7. Possible answers. In 1993, baseball tickets cost more than they did in 1920 so there was more money to pay the players. There has been inflation between 1920 and 1993, so people earn more now than they did then. The system of players and teams has changed since 1920, which resulted in a dramatic increase in what players make. TV revenues bring in more money.

8. **a.** Minimum: $\frac{175}{500} = .350$; maximum: $\frac{700}{500} = 1.4$

b. Possible answer: all home runs

c. $175 - 60 = 115$; 60×4 (bases from a home run) + 115×3 (bases from a triple) = 585; $\frac{585}{500} = 1.17$

STUDENT PAGE 7

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

9. a.

	Season Points	Standing
Berlin	78.89	4
Darien	18.89	7
East Lyme	86.67	3
New London	56.67	5
Plainfield	56.11	6
Redding	92.22	2
Wolcott	97.22	1

STUDENT PAGE 8

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

b. Winning most of your games is better than playing a tough schedule. Winning gives either 100 or 110 points, while playing a "tough" team gives only 10 points.

10. a. Possible answers: number of jumps, difficulty of moves, falls, grace, and choreography

b. Possible answers: successful hits, misses, possible spikes, and spikes that are not hit

c. Possible answers: number of blocks and number of balls that get through

11. There is a variety of forms for this formula. One of them is discussed in the Extension. The 50 is probably a quantity introduced to produce ratings that are in a certain range of numbers (maybe around 100). Students might be encouraged to apply the formula to players in their own football conference or in local or state university conferences.

STUDENT PAGE 9

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.

(11) a. 81.86

Player	Rating	Rank
Ken Anderson	81.86	11
Len Dawson	82.56	8
Brett Favre	82.18	9
Sonny Jurgensen	82.63	7
Jim Kelly	85.85	3
Bernie Kosar	81.82	12
David Kreig	82.98	5
Neil Lomax	82.68	6
Dan Marino	82.18	9
Joe Montana	92.26	2
Danny White	81.72	13
Roger Staubach	83.42	4
Steve Young	96.75	1

Source: *The Sporting News 1995 Football Register*

b. Steve Young, Joe Montana, Jim Kelly, Roger Staubach, David Krieg

c. Possible answers: reducing interceptions because they count so much against a rating; increasing the number of touchdowns because that has the greatest coefficient

If students are interested, have them experiment with different combinations.

12. a. The ratio for interceptions is negative because they are undesirable.

STUDENT PAGE 10

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. 4 times as important; $8000 \div 2000 = 4$

c. No, a touchdown has a coefficient of 8000 and an interception has a coefficient of $-10{,}000$.

d. Possible answer: The formula uses yards per attempt rather than just total yards passing so that players who are new to the game are not disadvantaged. The number of years played is not a factor this way.

13. a.

Player	Rating	Rank
Troy Aikman	99.022	2
Steve Beuerlein	82.466	7
Buddy Brister	84.945	5
Brett Favre	85.302	4
Jim Harbaugh	72.109	9
Bobby Hebert	84.041	6
Jim McMahon	76.202	8
Phil Simms	88.313	3
Wade Wilson	70.135	10
Steve Young	101.500	1

Source: *The Universal Almanac,* 1995

For 1993, Troy Aikman and Steve Young have a higher rating than the lifetime statistics of the all-time leading passers.

STUDENT PAGE 11

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

(13) **b.** Answers will vary.

14. a. Yes

b. (See table below.)

STUDENT PAGE 12

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.

Player	NFL Rating	Rank	NCAA Rating	Rank
Troy Aikman	99.022	2	145.13	2
Steve Beuerlein	82.466	7	131.38	4
Buddy Brister	84.945	5	122.08	7
Brett Favre	85.302	4	128.76	6
Jim Harbaugh	72.109	9	113.62	9
Bobby Hebert	84.041	6	129.85	5
Jim McMahon	76.202	8	114.51	8
Phil Simms	88.313	3	133.42	3
Wade Wilson	70.135	10	112.63	10
Steve Young	101.500	1	154.9	1

The top three remained in the same order. The middle 4 to 7 positions changed slightly. In particular, because Beuerlein and Hebert had more interceptions than the others their NFL ratings were lower.

15. a. A = area; l = length; w = width; area of a rectangle = length times width

b. A = area; b = base; h = height; area of a triangle = $\frac{1}{2}$ times the base times the height.

c. V = volume; r = radius; h = height; volume of a cylinder = pi (3.14) times the radius squared times the height

d. a = leg; b = leg; c = hypotenuse; in a right triangle, the length of the hypotenuse squared equals the sum of the squares of the lengths of the legs of the triangle (the Pythagorean Theorem).

16. Possible answer: $R = \dfrac{50A + 2000C + 8000T - 10000I + 100Y}{24A}$

17. Answers will vary.

18. In most competitions, five judges are involved in scoring high-school diving competition. Each scores a dive from a low of 0 to a high of 10, scoring in quarter points, for instance, 6.25, 6.5, 6.75, 7.0. The highest and lowest scores are eliminated. The other scores are summed and then multiplied by a particular dive's difficulty factor. A typical list follows:

front two-and-a-half tuck: difficulty factor 2.2

front one-and-a-half pick: difficulty factor 1.8

front somersault, 2 twists: difficulty factor 2.3

back one-and-a-half tuck: difficulty factor 2.2

back one-and-a-half twister free: difficulty factor 2.3

inward one-and-a-half tuck: difficulty factor 2.2

reverse one-and-a-half tuck: difficulty factor 2.3

Possible formulas, where $J1$, $J2$, and so on, represent judges' scores:

$2.3\,(J1 + J2 + J3)$ = Somersault; $2.2(J1 + J2 + J3)$ = Front two-and-a-half tuck

LESSON 2

Formulas That Manage Money

Materials: *Activity Sheet 2*
Technology: spreadsheet program, or graphing or scientific calculators
Pacing: 2 class periods

Overview

In this lesson, students use the context of managing money at a school store to think about variables and to generate a formula for total salaries for the store employees. They list variables such as hours worked, hours of overtime, and pay rate; and decide how to incorporate these variables into a spreadsheet to find individual employee salaries. They use symbols to represent variables and write a formula using their variables to find the total pay. Students evaluate their formulas for different situations and compare different ways to pay their employees. Students are informally introduced to the concept of function by using language such as, "The total pay is a function of the number of hours worked per week." It is assumed that students will use graphing calculators or computer spreadsheet programs.

Teaching Notes

You may choose to have students use the data provided, or you may have them create their own data. Some students will not be familiar with terminology such as commission, Social Security, gross pay, base pay, and so on. You might need to spend some time discussing what a typical paycheck could look like and how those variables would affect the amount actually received.

Emphasize the notion of variables, what they represent, and how they can be described using symbols. Some of the quantities in a given formula are constant for that formula, such as the tax rate or the amount deducted for Social Security. Point out the necessity for symbols of some form to enter formulas into either a spreadsheet or a calculator.

Technology

It is possible for students to do the lesson with only a scientific calculator. One of the intentions of the lesson, however, is to introduce them to either a spreadsheet or the list functions of a graphing calculator. The ability to do a repeated series of calculations by giving a single formula or command emphasizes the need for symbols and formulas and demonstrates how efficiently the process can be carried out. Be sure students are comfortable with using the technology and are successful at producing results. You may want to pair students who have some technological skills with those who don't to help facilitate the work. You may also want to have some of the students demonstrate how they use their computers or calculators for the rest of the class.

Be sure that students record what they are saving in each list or column when they use a spreadsheet or calculator. An example would be:

$L_1 = R \quad L_2 = H \quad L_3 = P.$

This will help the students keep in mind what they put into each list.

- If the students are unfamiliar with using a spreadsheet or the graphing calculator, some time may be needed here to familiarize them with the procedures involved.

- If the students are using a TI-83 calculator, they must remember to put parentheses around the entire numerator of a fraction.
- If the students want to review their formula on the TI-83, it is necessary to go to the home screen and write the formula. Then store the formula under the list name, for example, L_6.

If you are using the TI-83, you may want students to download the data using the link. An efficient way to do this is to appoint one student in each group as the calculator representative who will get the data from your calculator before class, then transfer the data to each member of the group.

Follow-Up

Students could think about other places in and out of school where variables are important, such as average daily attendance, formulas for computing state aid to schools, for paying taxes, or for ranking tennis players. They might search newspapers and magazines for data presented in tables and write a report about the variables represented in those tables. They might interview teachers from other disciplines about variables or formulas that are used in that discipline or be asked to invent their own formulas, carefully defining the variables.

Students might create a project using variables and formulas around topics that involve food, entertainment, clothes, savings, or cars. The topics will vary depending on the age and interest of the students.

STUDENT PAGE 13

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

Solution Key

Discussion and Practice

1. Possible answers: inventory, wages, hours worked, products sold, money taken in, number of employees, cost of supplies

2. Possible answers are given. You may want students to form groups and have each group develop its own criteria.

   a. 45 hours; possible to be open all day, open during lunch, open before and after school

   b. 3 employees

   c. 3–6 hours

   d. \$5 per hour

   It is important for students to be reasonable in their estimates of hourly wages.

   e. Yes

3. **a.** Possible answer:

Name	Rate of Pay R	Hours Worked H	Total Pay P
Mary	\$5	3	
Lee	\$5	4	
Ami	\$6	5	
James	\$5	3	

   b. R: \$4.50–\$10, depending on the minimum wage; H: less than 15 because students need time for school

STUDENT PAGE 14

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

STUDENT PAGE 15

(3) c. $P = R \cdot H$

Name	Rate of Pay R	Hours Worked H	Total Pay $P = R \cdot H$
Mary	\$5	3	\$15
Lee	\$5	4	\$20
Ami	\$6	5	\$30
James	\$5	3	\$15

4. Answers are given for the data in problem 3. To get the summation on a TI-83 calculator, use the STAT DATA menu. Arrow to CALC and then 1-Var Stats is highlighted. Press ENTER, then enter the list name you want summarized. You will see the summation symbol. This will give you the total of the list you entered. You can also use second List. Arrow to Math. Go to Sum and ENTER. Enter the list number that you want summed.

   a. \$80

   b. $P = \$20 + (R - \$0.50)H$

   c. (See table below.)

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: L1 * 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 Σx.

   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.

Name	Rate of Pay R	Hours Worked H	Total Pay $P = R \cdot H$	New Pay = $20 + (R - \$0.50)H$
Mary	\$5	3	\$15	\$33.50
Lee	\$5	4	\$20	\$38
Amy	\$6	5	\$30	\$47.50
James	\$5	3	\$15	\$33.50
Total			\$80	\$152.50

You would pay out more under the new plan.

d. Possible answers: Advantage: Employees will earn at least \$20, even if they work only a small number of hours; disadvantage: If an employee works more hours, he or she would would receive a lower average salary than those who worked a few hours.

The original plan is cheaper, and the employees are paid for only the number of hours they work. Some people don't make very much money.

5. **a.** Total sales for each employee; rate of commission

b. Answers will vary, depending on commission; for a 10% rate of commission and sales S, $C = (0.10) \cdot S$.

6. **a.** $G = P + C$

b. $G = R \cdot H + (0.10) \cdot S$

c. This includes all relevant information except taxes and deductions. Students may want additional information included. For example, this format does not allow for different hourly wages for different times.

STUDENT PAGE 16

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.

STUDENT PAGE 17

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.

7. **a.** $P = G - 0.15 \cdot G - 0.02 \cdot G - 0.0765 \cdot G$, or $P = G - FI - SI - SS$

b. See table below. Answers will vary due to rounding. Encourage students to keep all decimals until they report the final results.

If using a TI-83 calculator, students might let

L1 = hourly wage;

L2 = hours worked;

L3 = federal income tax $= .15 \cdot L1 \cdot L2$;

L4 = state income tax $= .02 \cdot L1 \cdot L2$;

L5 = FICA (SS) $= .0765 \cdot L1 \cdot L2$;

L6 = pay $= L1 \cdot L2 - L3 - L4 - L5$

Name	R	H	$G = R \cdot H$	$FI = .15 \cdot G$	$SI = .02 \cdot G$	$SS = .0765 \cdot G$	$P = G - FI - SI - SS$
Emma	\$5.65	40 hr	\$226	\$33.9	\$4.52	\$17.289	\$170.290
Bond	4.70	30	141	21.15	2.82	10.7865	106.24
Latisha	5.80	20	116	17.40	2.32	8.8740	87.41
Matt	6.70	30	201	30.15	4.02	15.3765	151.45
Terine	9.80	35	343	51.45	6.86	26.2395	258.45
Sal	8.35	40	334	50.10	6.68	25.5510	251.67
Myrna	7.10	20	142	21.30	2.84	10.863	107.00

c. Some students may use the variables FI, SI, and SS, while others may use G and the multiplying factor each time. Some students may realize that the deductions can be combined so that just $.2465G$ is subtracted.

8. **a.** A column could be added for total overtime earnings, \$60, for each employee. Alternatively, a column could be added for the overtime rate OR and another for the overtime hours OH.

b. Possible formula:
$P = R \cdot H + OR \cdot OT - .15(R \cdot H + OR \cdot OT) - .02(R \cdot H + OR \cdot OT) - .0765(R \cdot H + OR \cdot OT)$

c.

Emma	$215.50
Bond	$151.45
Latisha	$132.62
Matt	$196.66
Terine	$303.66
Sal	$296.88
Myrna	$152.21

Practice and Applications

9. **a.** (See table below.)

STUDENT PAGE 18

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.

Name	R	H	$P = R \cdot H$	S	$C = .05 \cdot S$	$G = P + C$
Lucinda	$4.25	4 hr	$17.00	$75.99	$3.80	$20.80
Frank	$4.25	6	$25.50	$88.54	$4.43	$29.93
Kordell	$5.15	8	$41.20	$72.00	$3.60	$44.80
Brittany	$5.15	8	$41.20	$74.95	$3.75	$44.95
Mihail	$4.25	3.5	$14.88	$68.35	$3.42	$18.29
Erin	$4.25	10	$42.50	$91.50	$4.58	$47.08

b. $205.84

c. Possible answers: Sum for *H:* 39.5 hours, the total number of hours worked; sum for *P:* $182.28, the total of the base pay; sum for *S:* $471.33, the total sales; sum for *C:* $23.57, the total paid in commission; and sum for *G:* $205.84, the total amount paid out in salary

d. Possible answers: Column *R,* the mean hourly pay rate; column *H,* the mean of hours worked; column *S,* the mean sales; and column *G,* the mean weekly pay.

10. Sample answers: Look at the percent of commission in relation to the amount of sales and the salaries.

11. a. Possible formula: $AG = G + R$

b. (See table below.)

STUDENT PAGE 19

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?

Name	R	H	$P = R \cdot H$	S	$S = .05 \cdot S$	$G = P + C$	$AG = G + R$
Lucinda	\$4.25	4	\$17.00	\$75.99	\$3.80	\$20.80	\$25.05
Frank	4.25	6	25.50	88.54	4.43	29.93	34.18
Kordell	5.15	8	41.20	72.00	3.60	44.80	49.95
Brittany	5.15	8	41.20	74.95	3.75	44.95	50.10
Mihail	4.25	3.5	14.875	68.35	3.42	18.29	22.54
Erin	4.25	10	42.50	91.50	4.575	47.08	51.33
Total						\$205.84	\$233.14

12. \$27.30

13. a. Possible answers: mileage, days rented, gas needed, size of car, miles driven, and insurance

b. $C = (3D + 3I + 3S) + (3D + 3I + 3S)T$

c.

Agency	*W*	*D*	*I*	*S*	*T* (%)	Cost for 3 days
Avis	155.54	25.19	12.99	2.5	13.8	$138.88
Dollar	157.00	39.95	11.95	2.5	13.8	185.72
Hertz	182.99	34.19	12.99	2.5	19.0	177.36
Budget	158.65	27.95	11.95	2.5	13.7	144.63
Alamo	143.09	27.99	12.99	2.5	13.8	148.44
Thrifty	153.95	27.95	9.95	2.5	12.0	135.74
National	149.40	26.00	12.00	2.5	13.8	138.27

Thrifty is the least expensive.

With a TI-83 calculator, one possible formula for the rate is L6 = 3(L2 + L3 + L4) · (L5 + 1).

d.

Agency	*W*	*D*	*I*	*S*	*T*(%)	Cost for 7 days
Avis	155.54	25.19	12.99	2.5	13.8	$300.40
Dollar	157.00	39.95	11.95	2.5	13.8	293.77
Hertz	182.99	34.19	12.99	2.5	19.0	346.79
Budget	158.65	27.95	11.95	2.5	13.7	295.39
Alamo	143.09	27.99	12.99	2.5	13.8	286.23
Thrifty	153.95	27.95	9.95	2.5	12.0	270.03
National	149.40	26.00	12.00	2.5	13.8	285.52

Thrifty is still the least expensive.

With a TI-83 calculator, one possible formula for the weekly cost is: L6 = (L1 + 7(L3 + L4)) · (L5 + 1).

14. Data from Problem 9 were used for this problem. Student answers will vary based on their data sets. Answers are rounded to the nearest penny.

a.

Name	Rate	Hrs	Base Pay $P = R \cdot H$	Amount of Sales *S*	Commission $C = .1 \cdot S$	Base Pay or Commission *P* or *C*
Lucinda	$5	4	$20.00	$75.99	$7.599	$20.00
Frank	5	6	30.00	88.54	8.854	30.00
Kordell	5	8	40.00	72.00	7.200	40.00
Brittany	5	8	40.00	74.95	7.495	40.00
Mihail	5	3.5	17.50	68.35	6.835	17.50
Erin	5	10	50.00	91.50	9.150	50.00
Total						197.50

The total payroll decreased.

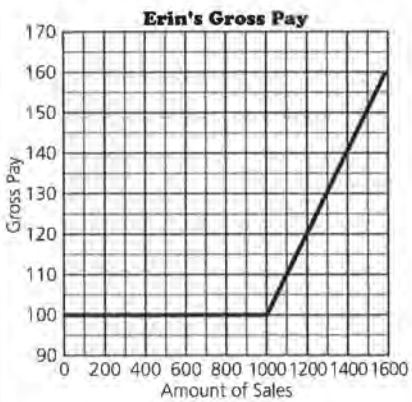
(14) **b.** Possible answers:

If $0.10 \cdot S > \$100$, then $G = 0.10G$; otherwise $G = \$100$.

$G = .01 \cdot A$, unless $A < \$1000$; then $G = \$100$.

c. $800

d.



With 10% commission, she needs to sell $1000 a week before she receives commission.

e. In this example, it appears that the store is better off under the new rules. Most people don't appear to sell enough to make any commission, and the slight increase in base pay does not make up the savings for not paying commission.

15. Answers will vary.

STUDENT PAGE 20

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

Materials: *Places Rated Almanac* (David Savageau and Richard Boyer, Prentice Hall Travel, 1994) or another book that rates cities, schools, or cars (*Consumers Report's* yearly car ratings) on hand in the classroom or in the library
Technology: scientific or graphing calculators
Pacing: 2 class periods

Overview

Many formulas are created in ways that are subjective rather than based on a mathematical characteristic that can be quantified. Often consumers are asked to create their own formulas. In this lesson, when using collected data to rank items such as cities or football teams, students both use formulas given to them and create their own. They learn the difference between a rating, a quantitative value assigned according to some criteria, and a rank, an ordering of elements. They use and analyze formulas, many given in words, to determine ratings. Students translate the words into symbols, and then evaluate the formulas for given situations. They create their own formulas to determine rank orders from the ratings, and in the process, the notion of weighted values is introduced. Throughout the lesson, students investigate a variety of quantitative relationships, "mathematize" the relationships, and apply the results in different situations.

Teaching Notes

The lesson has two purposes. One is to give students experience in reading and using formulas other than those typically found in mathematics texts (such as those for area, volume, and distance) to show how formulas can be used in the world. To build their number sense, students analyze the coefficients of the variables to determine their impact on the results. The numbers for the actual situations vary, provide students with practice in decimals and percents, and help develop their number sense. In order to combine categories that involve variables with different units or in vastly different magnitudes, students rank order each category and combine the ranks in some way. They should recognize that this transformation loses some of the information that can be observed from the actual numbers. If both are in rank orders, the relative difference is no longer a factor; the difference between 100 and 1 is the same as the difference between 2 and 1. With nearly every mathematical transformation, you can perceive a characteristic of the data in one form that you cannot see in another form. Think of the difference in what you see in $y = (x + 2)(x + 2)$ and in $y = x^2 + 4x + 4$ or what you can see in a stem-and-leaf plot and a histogram of the same data.

The second purpose of the lesson is to promote student awareness of the concept of rating and ranking that is prevalent in society. Nearly every week, the media report the results of new rankings for colleges, football coaches, hospitals, restaurants, and books. The criteria used for these rankings are often obscure, and the procedures subjective. Before accepting the results of the studies as they are published, students should understand the importance of knowing what was done. They should recognize the need for an objective list of criteria and for an algorithm that can be examined to see whether the criteria are appropriate. They should be alert to what variables were considered and how the data were gathered and by whom. Essentially, students should take away from the lesson a notion of what questions should be asked when they read about ratings or rankings.

You might foreshadow the lesson by finding articles about rating and ranking in current newspaper and magazines. Bring these to class and use them for discussion to begin the lesson.

Technology

While it is important to do some of the calculations by hand, if possible students should be allowed to use technology such as a spreadsheet or a graphing calculator, or the work may become tedious and the mathematics lost in the routine. The data are available on the data disk that accompanies the module and can be downloaded into a spreadsheet or into a graphing calculator. Before the students do an extensive set of calculations, be sure that the data are accurate.

Follow-Up

Students can prepare a report about a rating or ranking situation they found in a newspaper or other media. Working in pairs would give students the opportunity to discuss and investigate the article and its implications while they produce the report. Give them a specific list of questions: Who did the rating or ranking, what variables were selected, were there other important variables, how were the data collected, how were the data combined to produce a rating, and how was the ranking produced? You might also have students evaluate the significance of the report: Who would be interested in the results, how reliable do you think the results are, what would you do differently if you were responsible? Students might share the results with a brief (5-minute maximum) oral report to the class; you might use the reports for a bulletin board. If the report is particularly good, you might send it to the sponsors of the original rating situation.

STUDENT PAGE 21

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.

STUDENT PAGE 22

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

Solution Key

Discussion and Practice

1. **a.** PG means Parental Guidance. There is some content or language that is not appropriate for children.

Four stars is usually an indication that the reviewer thought very highly of the movie.

Best American movie indicates there are no movies that are any better, in the reviewer's opinion.

b. There are many movies that have the rating of PG and four stars. There should only be one movie that is the best American Movie; "best" is a rank.

2. Students need to think about whether the highest rating means the car has the most or the least number of complaints.

a. If you were buying a car, you would be interested in one that has a high ranking, because you wouldn't want to have many repairs. If you were opening a car repair business, you might want to specialize in cars that had a low ranking.

b. Possible answers: number of repairs needed, cost of repairs, and ease of getting repairs done

3. **a.** A rating is a judgment about quality. A rating is given based on a set of criteria, and many items can have the same rating. A ranking is an ordering of several items. Only one item can have a particular ranking unless there are ties. In the case of a tie, the next rank order is usually omitted: 1, 2, 2, 4, . . .

b. This is a rating because it gives information on popularity, how many people watched, but does not give a comparison with other television shows. The 15% market share does not put the show in any order unless it is used with other market shares.

c. This is a form of a rank because it compares *Frasier* to other television shows.

d. Possible answers:

i. Ranked: athletes—professional tennis standings are used to determine pairings; television shows—the most watched show can charge the most for commercials; people—the best candidate gets the job.

ii. Rated: books—like movies, reviewers give ratings to books to help readers make choices; hotels—some people will only stay in hotels that have a particular rating from AAA or another organization; restaurants—people go to those that have the best rated food.

Consumer Reports magazine rates many items to help people make choices.

STUDENT PAGE 23

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?

4. This may be difficult for some students. You might read the formula for one of the variables aloud in class and discuss how you would use it: what numbers would you have to know before you could use the formula and what would you do with the numbers to find the result? Then have students do the problem on their own or in pairs.

a. The variables are the estimate of *percent* increase in new jobs (1993–1998) and prediction of *number* of new jobs (1993–1998).

b. $(.0275)(36{,}053 - 1763) + 2000 = (.0275)(34{,}290) + 2000 = 942.975 + 2000 = 2942.975$

c. Predictions are made by looking at trends, both local and national, and by studying data from local surveys, the Census, and the Department of Labor. The actual process that takes many variables, such as the economy, technology development, and population shifts, into account is highly mathematical.

STUDENT PAGE 24

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.

---

5. Students might question why the divisors of 100, 75, and 50 are used and whether the formula seems to be an adequate one for providing an education rating. They might suggest other variables they would find useful in making a rating.

a. Two-year college enrollment, public-college enrollment, and private-college enrollment

b. $\frac{23067}{100} + \frac{5501}{75} + \frac{71045}{50}$

$= 230.67 + 73.35 + 1420.90$

$= 1724.92$

c. $\frac{\text{two-year}}{100} + \frac{\text{private}}{75} + \frac{\text{public}}{50} = R$

d. An increase of 1000 students in public colleges would have a greater effect on the rating, an increase of 20. An increase of 1000 students in a two-year school would increase the rating by only 10. Some students might need to work an example to see this. Others will recognize that dividing by 100 reduces the effect of the numerator more than dividing by 50.

6. **a.** In jobs, education, and recreation, a high rating is good. If students read the description of the formulas, they should see that you would like a large number of jobs, a large number of students enrolled in educational institutions and a large number of recreation opportunities. In housing costs and crime rates, a low rating is good.

b. Possible answer: Because there tends to be more educational institutions in large cities, education would help produce high ratings for large cities.

7. **a.** $\frac{1}{10}$ reduces the importance of property crimes in relation to violent crimes. The use of 10 seems arbitrary; why not $\frac{1}{5}$? More information would have to be obtained to see whether the $\frac{1}{10}$ were based on fact.

b. Answers will vary.

8. **a.** Possible answers: San Diego (ocean, weather, zoo), Denver (mountains, skiing, hiking), Orlando (Disneyworld, Epcot, Seaworld), and Minneapolis (lakes, fishing)

b. Possible answers: Grand Forks, ND; Racine, WI; Terra Haute, IN. Small, inland cities would probably not have professional sports or coastline, for example.

9. It appears that education ratings are higher in larger cities.

10. a. Housing ratings, from 4,116 to 20,322; this, however, depends on the potential minimum and maximum from the scoring procedure. Students should recognize the importance of thinking carefully about the variability. A 100-point difference in recreation might indicate a greater variation than a 1,000-point difference in housing.

b. Crime ratings, from 869 to 1474; this is reasonable because the values for crime are reported as rates in terms of crime per 100,000 residents. This means that city size has been factored out and has no direct effect.

11. Note: This ranking assumes that low housing costs are desirable.

Boston	7
Washington	6
Atlanta	5
San Diego	8
Terra Haute	1
Lincoln	3
Greenville	4
Salem	2

The size of the differences is lost.

12. It is important to realize that higher crime rate means that a community is less desirable. Thus, when ranking from good to bad, the lowest number in this category is good. You might have students write "high-good" for education, recreation, jobs, and "high-low" for crime and housing on the columns.

STUDENT PAGE 25

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?

a.
Rankings

City	Housing	Jobs	Education	Crime	Recreation
Boston	7	5	1	6	2
Washington	6	2	2	5	3
Atlanta	5	1	4	8	4
San Diego	8	3	3	7	1
Terra Haute	1	8	7	1	7
Lincoln	3	7	5	4	6
Greenville	4	4	6	3	8
Salem	2	6	8	2	5

Possible answers: No one city is rated high in all categories. Terra Haute is either the best or the worst. San Diego is nearly the opposite of Terra Haute over the categories.

(12) b. Possible answers: Ranks make it much easier to see where cities are in relation to one another. They can be combined more easily because there is no problem with relative size of numbers. One category may have ratings from 10 to 20; another from 10,000 to 200,000. In one category high numbers may be good; in another, low numbers are good. A rank allows these categories to be considered at the same time.

13. a. $\frac{20{,}322 - 15{,}466}{15{,}466} = \frac{4{,}856}{15{,}466} \approx$ 31.3% decrease

b. $\frac{15{,}466 - 6{,}911}{15{,}466} = \frac{8{,}555}{15{,}466} \approx$ 55.3% decrease

c. $\frac{HA - HB}{HA}$

14. a. Possible answers: Add all the rankings, and the city with the highest total would be the best. Find the average ranking. (This will produce the same result as the lowest total. Ask students why this is so.) Students may have problems combining crime with the others by adding. Remind them that a 1 for crime means a good rating even though it came from a low value. Some weighting could be used in the combination of ranks.

b. Possible answer: No; recreation, housing, and crime may be more important to grandparents; all of the categories may be important to families; education and recreation may be more important to students.

STUDENT PAGE 26

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?

15. Possible answer: Housing 1, Jobs 3, Education 2, Crime 3, Recreation 1. Be sure students weigh the ranks, not the ratings. Answers are given for this weighting.

a. $1(7) + 3(5) + 2(1) + 3(6) + 1(2) = 44$; $44 \div 10 = 4.4$

b. $\frac{1H + 3J + 2E + 3C + 1R}{10}$

c.

City	Housing	Jobs	Educ.	Crime	Rec.	Weighted Ranking
Boston	7	5	1	6	2	4.4
Washington	6	2	2	5	3	3.4
Atlanta	5	1	4	8	4	4.4
San Diego	8	3	3	7	1	4.5
Terra Haute	1	8	7	1	7	4.9
Lincoln	3	7	5	4	6	5.2
Greenville	4	4	6	3	8	4.5
Salem	2	6	8	2	5	4.7

16. a.

City	Rank from Weighted Ranking
Boston	2
Washington	1
Atlanta	2
San Diego	4
Terra Haute	7
Lincoln	8
Greenville	4
Salem	6

b. Answers will vary. Here, jobs and recreation appear to have rankings similar to the weighted rankings. Boston, Washington, Atlanta, and San Diego are in the top half in both.

STUDENT PAGE 27

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?

Practice and Applications

17. a. Possible answers: number of players, number of players recruited by colleges, number of players going to particular college programs, number of players placed on national all-star teams, number of coaches with professional experience, and number of players that make it to the NFL

b. Answers will vary depending on procedures used to combine the data. In the example below, the numbers were added.

State	NFL	College	Rating	Ranking
AL	38	70	108	6
CA	169	272	441	1
FL	106	316	422	2
GA	71	46	117	4
IL	38	61	99	7
MA	23	27	50	9
MI	35	55	90	8
TX	125	218	343	3
VA	43	69	112	5

18. Possible answers: Divide the ranks into categories; for example 1–3 is top, 4–6 is middle, and 7–8 is bottom. Then the city with the most 1–3 ratings is first, and the middles and bottoms are used to complete the ranking. An advantage is that the procedure uses small numbers and is easy to do. A disadvantage is that it is hard to account for ties (there are three 3s and no 4 or 5) or extremes. A city the best in two categories can be the worst in others. Some students may suggest a

"+/–" system, giving a "+" if a rank is above average and a "–" if it is below average. Then the pluses and minuses are summed. Again, the method is simple but the relative differences are made even less important. You might have students use several of the methods on the data and discuss the difference in the results.

19. Possible questions: What were the variables, how were the variables measured, where did the data come from, how were ratings for each determined, and how were the variables weighted to determine the ranking?

20. **a.** Answers will vary. Students may use formulas similar to the ones they used for the other city rankings.

b. *Money* uses different weightings and scales. It appears that a scale of 0–100 is used for the individual ratings.

STUDENT PAGE 28

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

STUDENT PAGE 29

21. a. *Places Rated* finds the sum of the ranks for each of the ten individual categories, then orders these sums for a composite rank. The Cincinnati, Ohio, metropolitan area is ranked number 1. Yuba City, California, was 343 out of 343 cities according to the 1993 *Places Rated Almanac.*

b. Answers will vary.

22. Students should be encouraged to look carefully to determine the criteria for the rankings.

As a class activity, students could brainstorm a list of the qualities of a good teacher. They should recognize these qualities as variables. They can weight the categories, select five or six teachers, and rank them. A typical group's data are below, with 1 being a low rating and 10 a high rating.

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.

Qualities	Teacher 1	Teacher 2	Teacher 3	Teacher 4	Teacher 5
Homework *H*	10	9	10	10	8
Discipline *D*	8	4	4	4	2
Sense of Humor *SH*	9	3	10	2	5
Likes Kids *K*	8	5	8	1	10
Knows Subject *S*	10	8	8	10	7
Caring *C*	6	1	6	1	10
Amount Learned *L*	9	6	7	3	5
Interesting *I*	10	2	5	1	5

$$\text{Rating} = \frac{H + 2D + 2SH + 3K + 2S + C + 2L + 2I}{8}$$

The rankings came out as follows:

First	Teacher 1	16.5
Third	Teacher 5	12
Fourth	Teacher 2	8.9
Second	Teacher 3	13.5
Fifth	Teacher 4	6.8

LESSON 4

Expressions and Rates

Materials: *Activity Sheet 3*
Technology: scientific or graphing calculators or spreadsheets
Pacing: 1 class period

Overview

The focus of this lesson is on using variables in expressions that describe rates. Rate is a ratio in which one quantity, amount, or measure is related to another used as a base. Students analyze data to decide in which state it is safest to drive. As they work with rates, students discover the need for variables to keep track of their work and to communicate their process and results. By manipulating existing variables, they create new variables. The data are in units of different magnitude, which must be dealt with carefully in order to make sense of the information. Ignoring the "thousands" or "in millions" will produce misleading results for any set of calculations. Students also explore weighted and unweighted averages for the data and think about which is more appropriate for the given situation. This builds on the concept of weighting from Lesson 3.

Teaching Notes

Some of the concepts in this lesson are difficult for students. Just working with numbers and trying to understand what a rate actually is can provoke much discussion. Rates allow us to compare numbers from different populations or with very different bases. They are relative measures and are commonly used when describing values such as characteristics of states where population size (Delaware as very small, California as very large) could be a factor. Note that a rate of 1 out of 10 does not indicate how large the population is, but does provide a means of comparison to 3 out of 10. If the original numbers were 100 out of 1,000 and 30 out of 100, without rates, you would be comparing 100 to 30. Be sure to spend time exploring these ideas before answering the question, "In which state is it safest to drive?"

The use of symbols here is primarily for convenience in organizing and manipulating the data. Students investigate a small set of data (five states) and decide how they want to use the numbers to answer the question. They then can apply their process to a much larger data set, all 50 states. Note that M_A is a shorthand form for "miles driven in the state of Alaska." Allow students to use the words or to write word phrases if that will help them keep the values straight.

Finding the weighted average as a way to bring in the size is an important concept. If the size is not included as a factor, a single data point that represents only a small set of the data could have a disproportionate influence on the results. You might need to demonstrate this with an example, particularly if you think students did not understand the point when it was covered in Lesson 3.

If you use a graphing calculator or a spreadsheet, be sure students keep a record of what each variable represents. They can do this on the chart or can keep a separate listing for the table. This helps eliminate confusion when students begin to do operations with the variables.

Technology

Students should have access to a spreadsheet or a graphing calculator in order to do the calculations.

The data are available on the data disk that accompanies this module and can be downloaded for student use. It is strongly suggested that you use the data set containing all 50 states.

Follow-Up

You might have students investigate the driving habits of their own family, analyzing the number of miles driven per year, estimating the amount of fuel used and the cost of fuel. They might also check out the safe-driving data for your state. They can use current information from an almanac to answer questions such as, "Has the death rate decreased or increased?" Students might also check in papers and magazines for the use of rates; write a report on what they found, including why they think rates were used to convey the information; and share the results with the class.

Solution Key

Discussion and Practice

1. **a.** Possible answers: number of accidents, type of roads on which accidents typically occur, seriousness of accident, speed, seatbelt use, involvement of alcohol or drugs, road conditions, and whether the accident was caused when a driver broke a law

   b. Possible answers: In many communities, police patrol streets; in some places, videos are monitored by police; public-service announcements can be used to promote safety awareness.

STUDENT PAGE 30

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. a. Answers will vary.

   b. Possible answers: highway-patrol records, police records, and insurance-company reports

3. This problem can be extremely difficult for students. Registration is given in thousands and miles driven is given in billions. This makes the calculations confusing if students do not use care in dealing with the units and the zeros. It may be necessary to actually write the problems, eliminate the zeros, and then divide if students are to understand what is happening with the numbers in the problem.

   a. 10,232,000 registered vehicles

   b. 262,500 million miles

   c. There were fewer motor-vehicle deaths in Rhode Island than in Alaska. In a direct comparison, there were 27 fewer deaths. In a relative comparison, there were .000218 deaths for every registered car in Alaska and only .000127 deaths per registered car in Rhode Island.

4. Possible answers: R, the number of registrations, found by checking state car-registration records; M, the number of miles driven, gauged from gasoline taxes; and D, the number of deaths, collected from police records

5. Possible answer: Probably not. D could be used, but this disadvantages states that have many cars that travel many miles. Some might argue that a great number of cars makes the traffic worse and so the number of cars should not be factored out.

STUDENT PAGE 31

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.

STUDENT PAGE 32

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?

6. a. Possible answers: 2.79 deaths for every 100,000,000 miles driven; .0279 deaths per million miles driven; .0000000279 deaths per mile driven

b. It measures the number of billions of miles driven for every death. In Alaska, there were .036 billion miles driven for every death.

c. In this problem, it is important that students recognize whether they are looking for a large value as good ($\frac{M}{D}$ for example) or a small value ($\frac{D}{M}$). Depending on which combination of variables students choose to use, they may make different conclusions. A table of some possible procedures is below. Students may think of others that are not listed.

	$\frac{M}{D}$	$\frac{R}{D}$	$\frac{D}{MR} \div 10^{12}$
Alaska	35,849,056	4,585	0.0574
California	68,789,308	5,818	0.0007
Florida	46,088,709	4,126	0.0020
New York	61,055,555	5,433	0.0017
Rhode Island	97,468,354	7,873	0.0165

Rhode Island has the greatest number of miles driven per death and the most registrations per death. Note that Alaska is the least safe based on $\frac{M}{D}$, while Florida is least safe based on $\frac{R}{D}$. If you factor in both registration and miles driven, $\frac{D}{M} \div R$, the least number of deaths with both factors occurs in California. If $\frac{D}{M}$ is used, Rhode Island has only 10.2597 deaths per billion miles driven while California has 14.5371 deaths per billion miles driven.

Some students may use $\frac{D}{M} \div R$. You might point out that while this is a logical way to think about the rate in terms of the problem, its equivalent $\frac{D}{M \cdot R}$ would not make sense as a way to start thinking about the problem. While one form of an expression may be easier to use, it may not be the form created by thinking about the problem initially.

7. Possible answers: the number of serious accidents and the number of nonserious accidents

8. a. Possible answer: Add M for all the states and then divide by the sum of the Rs.

b. With the method in part a, about 11,500 miles

9. a. 7818.93 miles;

$$\frac{M}{R} = \frac{3{,}800{,}000{,}000 \text{ miles}}{486{,}000 \text{ registrations}}$$

STUDENT PAGE 33

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

(9)b. Estimates for the states other than Alaska are between 11 and 13 thousand.

Alaska	7,818.9 miles
California	11,823.3 miles
Florida	11,170.8 miles
New York	11,237.2 miles
Rhode Island	12,379.4 miles

c. Answers will vary. Some students may think the procedure doesn't account for cars that aren't driven anywhere near that many miles.

10. The symbolization in this problem may need to be explained to students. The concept is not difficult if students understand the meaning of the variables.

a. Miles per registered car in Alaska

b. Average of the miles-per-car registration for the five states $(7{,}818.9 + 11{,}823.3 + 11{,}170.8 + 11{,}237.2 + 12{,}379.4) \div 5 = \frac{54{,}429.6}{5} = 10{,}885.9$ miles per registration

c. Phil's method weights each of the five states equally to find the average for the five states. He divides by 5 to get a value that is the average for a single state.

11. a. The total number of miles driven in the five states, 498.2 billion miles

b. Jerene is computing the typical number by dividing the total number of miles by the total number of registrations.

c. Possible answer: Jerene's method is better because Phil's method gives each of the states equal weight regardless of the number of cars or the miles driven in the state. Jerene's method avoids this problem by using only the totals.

d. Phil's method:

$$\frac{\frac{3.8}{486}+\frac{262.5}{22{,}202}+\frac{114.3}{10{,}232}+\frac{109.9}{9{,}780}+\frac{7.7}{622}}{5}$$

$$= \frac{0078 + .0118 + .0112 + .0112 + .0124}{5}$$

= 0.01089 billion miles per thousand cars

= 10,890 miles per car

Jerene's method:

$$\frac{3.8 + 262.5 + 114.3 + 109.9 + 7.7}{486 + 22{,}202 + 10{,}232 + 9{,}780 + 622}$$

$$= \frac{498.2}{43{,}322}$$

= 0.0115 billion miles per thousand cars

= 11,500 miles per car

12. Possible answer: Probably not. Even using Jerene's method, which gives the greater result, a 3-year-old car would have only 34,500. Since the car has been driven 35,000 miles, it was driven a typical distance and should not be advertised as having low mileage.

Practice and Applications

13. a.

Date	Miles per Gallon
6/19	22.6
6/22	27.8
6/28	24.0
7/1	25.5

b. 25.33 miles per gallon for all four trips:

$$\frac{126.48 + 230.74 + 115.20 + 188.70}{5.6 + 8.3 + 4.8 + 7.4} =$$

25.33 miles per gallon

Be sure students do not use 24.98 by averaging the results from part a. This would give each trip equal weight, but he drove much farther on two of the days. You may need to use an extreme case to make the point clear to students: 10,000 miles and 285.7 gallons for one trip; 100 miles and 4.5 gallons for the second. The individual averages are about 35 mpg and 22 mpg, which average to about 28.5 mpg. The total distance, however, was 10,100 miles with 290.2 gallons of gas for an average of 34.8 mpg. The small number of miles driven at 22 mpg does little to the overall average.

STUDENT PAGE 34

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

14. a. $\frac{5 \cdot 3 + 3 \cdot 4 + 4 \cdot 2}{5 + 3 + 4} = \frac{35}{12} = 2.92$

b. With one credit-hour of A work: $\frac{39}{13}$, or 3.0 Some students may set up and solve an equation to find the number of hours: $\frac{35 + 4n}{12 + n} = 3$, $36 + 3n = 35 + 4n$, $n = 1$. Others may just guess.

c. No; since his average is currently less then 3.0, he would need a grade that would give more than 3 points to raise the average. He needs at least one hour of an *A* as well as all the rest *B*s to get a 3.0 average. Three hours of *A* does not make up for five hours of *C*.

STUDENT PAGE 35

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.

15. a. Answers will vary. The table shown here contains the results of spreadsheet calculations for three different procedures. Note that the decimals are from the spreadsheet calculations and are accurate to only four significant digits because of the units given in original data.

Students could use the value of deaths per mile as a measure of safety. If they do this, Massachusetts is the safest state with only 10.25 deaths per billion vehicle miles. If deaths per total number of registered vehicles is used, Connecticut is the safest. Students should explain why they chose their measure of safe driving.

State	$\frac{M}{D}$	$\frac{R}{D}$	$\frac{D}{MR} \times 10^{12}$
Alabama	45,754,245.754	3,300.699	0.0066
Alaska	35,849,056.604	4,584.906	0.0574
Arizona	43,209,876.543	3,458.025	0.0083
Arkansas	39,182,282.794	2,558.773	0.0170
California	68,789,308.176	5,818.134	0.0007
Colorado	55,684,007.707	5,616.570	0.0062
Connecticut	89,527,027.027	8,206.081	0.0046
Delaware	49,285,714.286	3,892.857	0.0372
Washington, D.C.	Inc data	Inc data	
Florida	46,088,709.677	4,415.323	0.0020
Georgia	58,881,330.310	4,458.806	0.0029
Hawaii	62,500,000.000	6,046.875	0.0207
Idaho	44,444,444.444	4,255.144	0.0218
Illinois	63,709,090.909	5,805.091	0.0020
Indiana	63,192,904.656	5,006.652	0.0035
Iowa	54,691,075.515	6,192.220	0.0068
Kansas	62,532,299.742	4,963.824	0.0083
Kentucky	46,520,146.520	3,642.247	0.0072
Louisiana	38,920,780.712	3,552.239	0.0083
Maine	57,276,995.305	4,591.549	0.0179
Maryland	63,102,409.639	5,555.723	0.0043
Massachusetts	97,525,773.196	7,552.577	0.0028
Michigan	65,019,305.019	5,645.56	0.0021
Minnesota	70,912,220.310	5,996.558	0.0040
Mississippi	43,377,483.444	3,235.099	0.0118
Missouri	54,111,675.127	4,064.975	0.0046
Montana	44,736,842.105	4,773.684	0.0246
Nebraska	54,074,074.074	5,018.519	0.0136
Nevada	43,426,294.821	3,669.323	0.0250
New Hampshire	82,113,821.138	7,268.293	0.0136
New Jersey	77,545,691.906	7,298.956	0.0023
New Mexico	40,130,151.844	2,932.755	0.0184
New York	61,055,555.556	5,433.333	0.0017
North Carolina	53,486,529.319	4,205.230	0.0035
North Dakota	69,318,181.818	7,443.182	0.0220
Ohio	66,111,111.111	6,270.833	0.0017
Oklahoma	56,704,361.874	4,421.648	0.0064
Oregon	60,129,310.345	5,566.810	0.0064
Pennsylvania	57,734,627.832	5,293.851	0.0021
Rhode Island	97,468,354.430	7,873.418	0.0165
South Carolina	43,370,508.055	3,223.048	0.0089
South Dakota	44,720,496.894	4,360.248	0.0319
Tennessee	43,290,043.290	4,021.645	0.0050
Texas	53,418,384.037	4,153.418	0.0015
Utah	60,594,795.539	4,654.275	0.0132
Vermont	62,500,000.000	4,843.750	0.0344
Virginia	75,566,150.179	6,244.338	0.0025
Washington	75,883,256.528	6,860.215	0.0030
West Virginia	39285714.286	3030.952	0.0200
Wisconsin	73,913,043.478	5,799.689	0.0036
Wyoming	52,542,372.881	4,093.220	0.0394

STUDENT PAGE 36

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

STUDENT PAGE 37

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?

(15) b. Using number of miles per death or deaths per mile, the five safest states are Massachusetts, Rhode Island, Connecticut, New Hampshire, and New Jersey.

Using number of vehicles per death or deaths per registered vehicle, the five safest states are Connecticut, Rhode Island, Massachusetts, North Dakota, and New Jersey.

Using all three variables, Texas and California are the safest states.

Using either of the two common ways to determine averages, a vehicle is typically driven about 11,750 miles per year (11,731 vs. 11,760).

16. a. Possible answer: If you use a rate, you can make more valid conclusions than if you compare the numbers themselves because the rates are in the same scale. The size of the population or set is not a factor.

b. Possible answer: In a weighted average, the data values are multiplied by numbers, or weights, before they are used to find the average. In the grade point average in Problem 14, the weights were the point values assigned to each letter grade. For a simple average, each value counts the same.

17. a. $\frac{1,219,000,000,000}{73,000,000} = 16,690$ miles per household in 1983; and $\frac{1,510,000,000,000}{81,000,000} = 18,642$ miles per household in 1988; Some students may reason that billions divided by millions equals thousands and use thousands as a unit for their answers (16.698 thousand and 18.642 thousand miles per household).

You might ask whether the change is constant and also whether you could predict the number of miles traveled per household at five-year intervals, 1993, 1998, and so on.

b. $\frac{\$95,000,000,000}{73,000,000} = \$1,301$ in 1983; $\frac{\$81,000,000,000}{81,000,000} = \$1,000$ in 1988

c. Based on part a, no

d. $\frac{1219}{81} = 15.05$ miles per gallon in 1983; $\frac{1510}{82} = 18.41$ miles per gallon in 1988; so fuel efficiency is improving if you assume the same driving pattern and use of motor fuel.

18. a. Vehicle miles traveled, motor fuel consumed, and motor-fuel expenditures

b. All of the variables

STUDENT PAGE 38

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

Materials: graph paper, straightedges
Technology: scientific or graphing calculators or spreadsheets
Pacing: 2 class periods

Overview

In this lesson, students apply their knowledge of rates from Lesson 4 to real data, statistics on the deaths due to a variety of causes and over a span of years. They look at trends in deaths over time, plot the data, and make comparisons among the plots. As rates are applied to situations, it becomes necessary to use percents and frequencies in order to answer certain questions. Students investigate the relationship between a population and a rate or percent and use variables to understand how these may be combined and sometimes simplified.

Teaching Notes

In this lesson, students interpret and use data contained in a table. Much information is presented in this way, but rarely are students given real experiences in making sense of such data. Often, students study tables for patterns from some given functional operation; but with real data in actual tables, they have to confront the data and decide what to do with the numbers. It is sometimes not easy to decide where to place information. Let students experiment with their tables in order to see how important some of these ideas are.

This is a good lesson in which to use a spreadsheet in order to understand why the labels on the columns are important and to recognize how carefully you must think about your data in order to do the calculations that will answer the questions. The data on death rates are given in table form, with each column containing the information from a given year. Students manipulate the data and sometimes find it useful to work with the columns as they are. In other cases, it is more useful to work with the rows. Such an approach may not be apparent to many students who will have trouble recognizing what they should do when. You might have to give them a few examples and ask them to interpret the results of their work for several cases before they understand the task. For example, to find the number of deaths for a given population, you would add a row to the table (population per year) and find the product of a death rate due to a certain cause and the population in 100,000s.

The correct table for translating death rates into number of deaths is shown below. Since population P, is in millions and death rates are per 100,000, the correct formula for transforming from rates to numbers is:

$$D = \frac{(rate)\boldsymbol{P}(1{,}000{,}000)}{100{,}000} = (rate)\boldsymbol{P}(10)$$

	1979	1989	1990	1991
Population (millions)	225	247	249	252
Rate AC	852.2	871.3	2150.9	2168.0
Deaths; all causes	1,917,450	2,152,111	2,150,862	2,163,420
Rate A	46.9	38.5	37.0	36.4
Deaths; accidents	105,525	95,095	92,130	91,728

The calculations can be done easily on a spreadsheet or on a graphing calculator using the List function. In both cases, students have to think carefully about what they wish to do with the numbers and use some form of symbolic notation to organize their procedures. With some software, you can transpose rows and columns of data just as you do in a matrix. Have students work together on these problems, and then take time to explore with the class the different methods they used to do the calculations. Be sure to have students share their strategies for using the technology, as some students will have very innovative ways.

The plots can be made by hand or with a graphing utility. Be sure the scales are accurate if students do them by hand or else the trends will be difficult to see. The plots shown in the text and in the solutions do not have a complete vertical scale. This can be misleading, particularly when comparing trends or when looking at an increase or decrease in a trend. Be sure students are aware of this and understand that their interpretation should be guided by thinking about the plot as a whole.

Technology

Students should have access to a spreadsheet or a graphing calculator in order to do the calculations. The data are available on the data disk that accompanies the module and can be downloaded for student use.

Follow-Up

Have students find tables of data in the media and share them with the class. They might write a report about the data and the effectiveness of the tables. Have them tell if another arrangement would be possible. Have them make a list of the questions they could answer by inspecting the data in the table and think about why the data are important.

STUDENT PAGE 39

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?

OBJECTIVES

Apply rates in practical situations.

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

Change from rates to counts and percents.

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.

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

Solution Key

Discussion and Practice

1. **a.** To make it easier to compare places with different populations

b. If P is the population and D is the number of deaths in a year per 100,000, the number of deaths will be $\frac{P \cdot D}{100{,}000}$.

2. **a.** The total number of deaths per 100,000 in 1979

b. Yes; 17.9 per 100,000 from motor-vehicle accidents and 10.5 per 100,000 from homicides

c. Possible answers: No records were kept about HIV deaths in 1979. If records had been kept but there were no HIV deaths, then a value of 0.0 would have been listed.

STUDENT PAGE 40

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?

STUDENT PAGE 41

3. **a.** No; there are other causes of death besides those listed. Consider deaths by "natural causes," or deaths due to nonmalignant diseases. These would contribute to "all causes" but not to any of the listed categories.

b. Possible answers: Accident rates, including those for motor vehicles, have decreased over the years, while suicide and homicide rates have remained reasonably constant. Malignancy rates have been about the same since 1989, but are higher than in 1979. There has been an increase in HIV deaths, while deaths due to heart disease have decreased.

4. If students have difficulty with this, display the complete computation. You may also "reduce" the zeros to help them see exactly how the numbers are interacting.

a. $(225{,}000{,}000 \div 100{,}000) \cdot 852.2 = 1{,}917{,}450$ people

b. $(252{,}000{,}000 \div 100{,}000) \cdot 36.4 = 91{,}728$ people

c. $(249{,}000{,}000 \div 100{,}000) \cdot 18.8 = 46{,}812$ people

5. **a.** In a row, with the population figures corresponding to the years.

b. Possible answer: $D = (P \cdot R) \div 100{,}000$, where D is the total number of deaths in a given year for a particular cause, P is the population in millions of that year, and R is the death rate per 100,000 for that cause. If students are using a calculator, they might enter all of the rates R for a given year in a list, then use the formula above to generate a new list with the number of deaths for each cause for that year.

---

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?

Number of Deaths for Leading Causes of Death for All Ages

	1979	1989	1990	1991
Population (P) (millions)	225	247	249	252
All Causes (A)	1,917,450	2,152,111	2,150,862	2,163,420
Accidents	105,525	95,095	92,130	91,728
Motor-Vehicle Accidents*	53,550	47,671	46,812	45,108
Suicide	27,225	30,134	30,876	30,240
Homicide	22,500	22,971	24,900	27,468
Heart Disease	734,625	734,331	720,855	717,696
Malignancies	404,100	496,470	505,968	514,080
HIV	——	21,983	25,149	29,484

*Motor-vehicle accidents are a subset of all accidents.

STUDENT PAGE 42

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.

6. a. The HIV death rate increased by 1.6 units, so the percent of increase is $(1.6 \div 10.1) \cdot 100 = 15.8\%$. The homicide rate increased by 0.9 units, so the percent of increase is $(0.9 \div 10.1) \cdot 100 = 9.0\%$.

b. Yes; no other death rate increased by more than 0.9 per 100,000.

7. a. Answers will vary. In general, if the death rate for a particular cause increases, then that cause has become a bigger threat to society. However, all that you actually know is that a greater fraction of the population is dying of that cause than before. It is possible that the increase is nothing to worry about. For example, if a new safety feature drastically reduces the motor-vehicle accident death rate, then you might expect a very slight increase in some other death rates because a fraction of all those people not dying in auto accidents can be expected to die in some other way. As people continue to live longer, there could be an increase in deaths from illness associated with old age, such as Alzheimer's. Also, sometimes changes in the way data are collected can affect death rates. For example, as HIV becomes less stigmatized, you might expect more HIV deaths to be accurately reported and fewer to be disguised as if from some other cause. This could explain a rise in the reported HIV death rate.

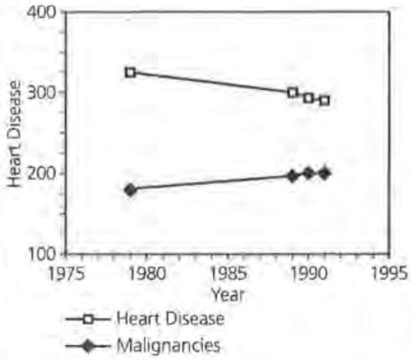
b. Answers will vary. In general, if the death rate for a particular cause decreases, then that cause has become less of a threat to society.

c. No; if the death rate decreases but the population increases, then the total number of deaths may increase, may decrease, or may stay the same.

Students may need an example before this makes sense. An extreme case can help make the point. For a death rate of 15 per 1000 people, there would be 15 deaths for 1000 people. If the rate decreases to 1 per 1000 people, but the population increases to a million, there would be 1,000 deaths.

8. a. No; since every motor-vehicle accident death is also counted as an accident death, the accident death rate must always be at least as large as the motor-vehicle accident death rate.

b. They decrease over time, and they decrease more rapidly starting in 1989 (accident death rate) or 1990 (motor-vehicle accident death rate).



c. The heart disease death rate decreases over time, and it decreases more rapidly starting in 1989; the malignancy death rate increases over time, but it increases less rapidly starting in 1990.

9. **a.** In 1979, 114.8 of every 100,000 people aged 15–24 died.

b. In 1990, 34.1 of every 100,000 people aged 15–24 died in a motor-vehicle accident.

c. In 1991, 1.7 of every 100,000 people aged 15–24 died of an HIV infection.

10. **a.** The number of people aged 15–24 in 1979, 1989, 1990, and 1991

b. The death rates for people aged 15–24 are significantly greater than (roughly twice) the death rates for the population as a whole.

11. **a.**

STUDENT PAGE 43

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?

b. The motor-vehicle accident death rate decreased faster between 1990 and 1991, but the accident death rate has been decreasing faster overall.

12. The accident death rate for people aged 15–24 decreases fairly smoothly over time. The motor-vehicle accident death rate for people aged 15–24 decreases over time, but less rapidly than for all accidents.

13. The suicide death rate has remained fairly constant, while the homicide death rate has risen sharply. Comments will vary.

14 a. $\frac{23.8}{46.9} \approx \frac{1}{2}$, or about 0.50746

b. It has remained fairly constant, although it dipped below 0.49 in 1991.

15. No; there cannot be more motor-vehicle accidents than all accidents.

16. 50.7% is the percent of all accidental deaths that are deaths due to motor-vehicle accidents.

17. a. $\frac{A}{100{,}000} \cdot P$

b. If A is the accident death rate per 100,000 then $\frac{A \cdot P}{100{,}000}$ is the number of deaths due to accidents for a given population P. If MV is the motor-vehicle accident death rate per 100,000, $\frac{MV \cdot P}{100{,}000}$ is the number of deaths due to motor-vehicle accidents.

c. If A and MV are death rates per 100,000, then $\frac{A \cdot P}{100{,}000}$ is the number of deaths due to all accidents and $\frac{MV \cdot P}{100{,}000}$ is the number of deaths due to motor-vehicle accidents. Therefore, the ratio of these two terms gives the fraction of accidental deaths that are due to motor-vehicle accidents. The ratio of the total number of deaths is the same as the ratio of the death rates. $\frac{P}{100{,}000}$ is a common factor and can be reduced from the numerator and denominator.

d. No; the population size is not a factor in the ratio of deaths due to the two causes, so the value of the population P is unimportant for this calculation. It is meaningful to compare two death rates even if the population size is not known.

18. a. 72.8% in 1979, 77.2% in 1989, 77.7% in 1990, 76.2% in 1991

STUDENT PAGE 44

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.

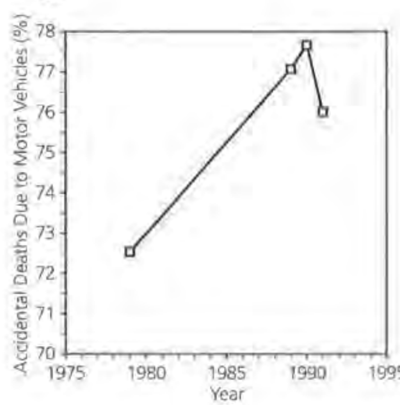
d. Is it necessary to know population size to compare rates as in part c?

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.

b.



STUDENT PAGE 45

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. In the first plot, we see that the motor-vehicle accident death rate is decreasing less rapidly than the death rate due to all accidents. In the second plot, we see that the motor-vehicle accident death rate is increasing as a fraction of the death rate due to all accidents. The plots show two ways of viewing the same phenomenon.

19. a. Possible answers: Is the rate for all causes increasing again? Is there a difference in the rates for ages 15 to 19 and 20 to 24? Does the increase in categories such as homicide, suicide, and HIV infections affect the decrease in the other categories (motor-vehicle accidents, accidents, malignancies)?

b. Possible answers: The percent of accidents from any cause for 15–24-year-olds is higher than than that for all age groups. The percent of motor-vehicle accidents for 15–24-year-olds is much higher than that for all age groups. There has been a steady decline in the number of accidents for this age group.

Practice and Applications

20. The size of the population

21. Yes;
$HD = (\text{number of } HD) \cdot \frac{100{,}000}{P}$

and $M = (\text{number of } M) \cdot \frac{100{,}000}{P}$.

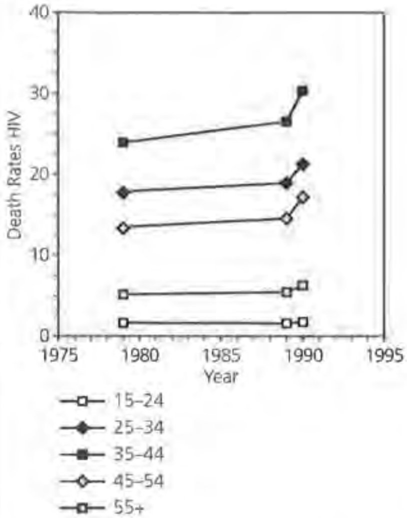
So the rate for the combined diseases where P is the same is

$(\text{no. of } HD + \text{no. of } M) \cdot \frac{100{,}000}{P}$.

22. Extension a. It lists death rates from the *same* cause for different age groups.

b. Answers will vary. A possible plot is shown below. The death rates for the 35–44 age category are much higher than those for the other categories, followed by those for the 25–34 category. The death rates for those in the 45–54 and 35–44 age categories are increasing at a faster rate than the others. This information is important because it is needed to define treatment plans and to decide which age group should be targeted.

What has caused the death rate to increase in the last two years? Scientists use this information to help in their research as they search for a cause and for a cure. They can tell from the trends when something significant has occurred and try to discover what variables might be involved.



(22) c. No; the two populations are of different sizes so the rates cannot be added. Find the rates for the first two age classes by computing the following, where P_i denotes the population size of the age class:

$$(\text{No. of deaths in 15–24}) \cdot \frac{100{,}000}{P_1}$$

$$(\text{No. of deaths in 25–34}) \cdot \frac{100{,}000}{P_2}$$

STUDENT PAGE 46

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

Materials: none
Technology: graphing calculators
Pacing: 2 class periods

Overview

In this lesson, students study formulas that reflect certain characteristics of the data. They review the mean and the median as measures of center, and use the total, the maximum, and the minimum as other measures to describe data. The processes are generalized using symbols, subscripts, and summation notation, Σ.

Teaching Notes

The lesson assumes that students have learned how to calculate the mean and the median in earlier grades. The focus in this lesson is on the formula and writing the formula using symbolic notation. You might have to review the difference between the mean and the median and also help students recall how to find the median for an even-numbered set of objects (the mean of the two middle values) and for an odd-numbered set of objects (the middle value). Some students may need to be reminded that the data must be ordered to find the median. The mode is not really a measure of center because it indicates only the most common data point. The mode does not have to be anywhere near the center of the distribution of the data.

Subscripts are commonly used to identify different items within a given category. For example, in later work in mathematics students will use (x_1, y_1) and (x_2, y_2) to describe points on a line. In this lesson, subscripts are used to describe different members of the junior and senior classes and different-sized cars.

The summation notation is commonly found on calculators and is used extensively in the study of sequences and series. Students have little trouble understanding the notation as a shortened way to think about a sum. The notation depends on subscripts to indicate the number of terms to be summed and on a rule or variable denoted by a subscript (x_i) that refers to the index on the Σ symbol. These indexes on the summation symbol indicate the beginning and the end of the sequence for that particular sum.

$\sum_{i=2}^{10} x_i$ indicates a sum from the second to the tenth elements of the sequence.

Be sure that students distinguish between an exponent and a subscript.

Most of the manipulation rules involving Σ are easily understood if students remember what the notation represents. Some students might need to write out the sum.

($\Sigma cx_i = c\Sigma x_i$ because $\Sigma cx_i = cx_1 + cx_2 + \ldots + cx_n = c(x_1 + x_2 + \ldots + x_n) = c\Sigma x_i$ by the Distributive Property. Have students note that the variable is x_i and the constant is c.)

Students may have trouble recognizing that the summation of a constant is just the product of the number of times the constant is used and the constant. You might ask them to think about $\Sigma 1$ for a given number of terms.

The distinction between $\Sigma(x_i)^2$ and $(\Sigma x_i)^2$ is an important one. Be sure that students understand that in the first expression the elements are squared and then summed, and in the second expression the elements are summed and that result is squared. This is a fundamental process that students will see many times. $(x + y)^2 \neq x^2 + y^2$ and $\sqrt{(x^2 + y^2)} \neq x + y$ are two examples.

Technology

Much of the lesson can be done either by simple hand calculations or by using a scientific calculator. Point out the key on the scientific calculator with the Σ symbol and help students understand how to use it. Using a graphing calculator throughout the entire unit, however, will be helpful to students as they process the data from the lessons. In some cases, the size of the data sets is quite large. The graphing calculator with lists also encourages students to explore and to try different things, such as omitting a value to see the effect on a statistic.

Follow-Up

You might ask students to explore finding a formula for the sum of the whole numbers from 1 to a certain number n. They can do this a variety of ways. Some might look at the number patterns: 1 sums to 1, $1 + 2 = 3$; $1 + 2 + 3 = 6$; $1 + 2 + 3 + 4 = 10$; and notice that they are adding $n + 1$ to the sum of the first n terms each time. This way of reasoning builds on the previous results. Some might look at a graph and discover a rule to describe it. Others may guess and test. The formula is $\frac{n\,(n + 1)}{2}$ for $n \geq 1$. (Some students may find $\frac{n^2 + n}{2}$ as a rule that will also work, but in this case n begins with 2, not 1.) Other equivalent expressions might also be suggested.

You might also ask students to look for other kinds of data where the total would be important (basketball points, amount of earnings in one week), where the maximum or minimum would be significant (temperature high or low, Olympic events where you get the best distance out of three tries). This might be a long-range assignment that could be collected after three or four days, so students have a chance to investigate other sources of data. Have them write a report describing their results, illustrating them with a graph, and including articles from a newspaper or some other source.

You may want to have students represent the data in the table about endangered species using Venn diagrams. This might enable some of them to see the relationship among the different sets more clearly and can be a very useful tool to organize and to represent information.

STUDENT PAGE 47

LESSON 6

Formulas That Summmarize 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

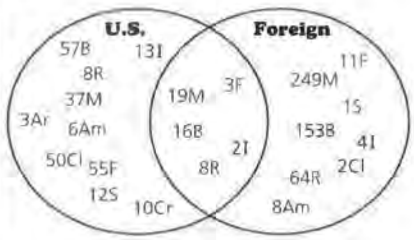
Summary Numbers

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

Solution Key

Discussion and Practice

1.



a. 226 birds

b. Answers will vary. Many students will think that mammals are the most endangered because they have the most species on the endangered list. Some students with other knowledge may suggest that another category is more endangered because a higher percent of the animals in that category are endangered.

2. **a.** There are 299 endangered species in the U.S. This is the sum of all of the categories in the U.S. Only column and in the U.S. and Foreign column. This sum gives all of the endangered species that can be found in the U.S. Students may find the Venn diagram helpful.

b. There are 791 endangered species worldwide. The is the sum of all of the values in the table. This represents all of the endangered species in the U.S. and in foreign countries.

c. No; the average would be the average number of endangered species per category, and this does not have any real meaning since the categories are arbitrary and the underlying populations are not the same size.

d. Answers will vary. Paragraphs should include information from the previous problems. You could encourage them to take and defend a position in the paragraph. Some might point out that 38% of the endangered species are in the U.S.

STUDENT PAGE 48

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.

STUDENT PAGE 49

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.** 86

b. The number of threatened species in foreign countries; 36

c. The number of threatened species from the first six categories in the U.S. and foreign countries; 23

4. Students may recognize the difference in the values for $\Sigma(x_i)^2$ and $(\Sigma x_i)^2$ from the numbers on the calculator screen. In one case, $\Sigma(x_i) = 583$ and $\Sigma(x_i)^2 = 38{,}345$. Students might see that squaring 583 will result in something less than 360,000, but not as small as 38,345. You can point out the cross products in the expression $(\Sigma x_i)^2$ will change the answer considerably:

$(60 + 70 + 80 + 70 + \ldots + 50)(60 + 70 + 80 + 70 + \ldots + 50) = (\underline{60 \cdot 60} + 60 \cdot 70 + 60 \cdot 80 + \ldots + 60 \cdot 50) + (70 \cdot 60 + \underline{70 \cdot 70} + \ldots + 70 \cdot 50) + \ldots + (50 \cdot 60 + 50 \cdot 70 + \ldots + \underline{50 \cdot 50})$,

while only the squared (underlined terms) are present in $\Sigma(x_i)^2$. An example with only three terms may make more sense at first to students.

a. Σx^2 represents the sum of the squares of the values. This is not very useful in this case. It is useful when some of the values are negative and others are positive and a sum of some kind without reference to the negative sign is desired.

b. $(\Sigma x)^2$ is the square of the sum of the values. Σx^2 is the sum of the squares of the values. These values are not the same in all nontrivial cases.

c. 109; the total number of threatened species that live in the U.S. and in foreign countries; the total of the sums of List 1 and List 2

d. $\Sigma USF = 23$; $\Sigma F = 36$

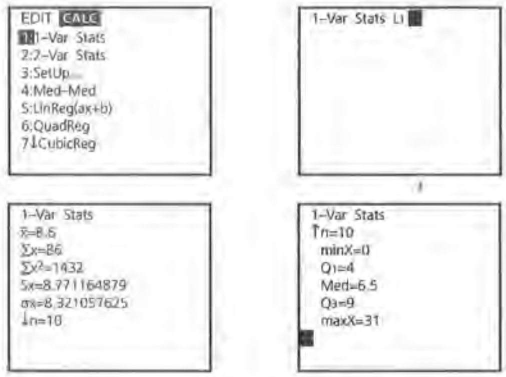
STUDENT PAGE 50

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.

STUDENT PAGE 51

5. **a.** Possible answers: for the number of calories, the mean; for cost per serving, the median, since there is one high-priced drink

b. No; no; the totals would represent the number of calories altogether and the cost for all nine drinks.

c. Eliminating Hydra Fuel changes the mean number of calories from 64.78 to 64.63. It changes the mean cost from \$0.29 (29.44 cents) to \$0.27 (26.63 cents).

d. Possible answers: Gatorade and Snapple Snap-Up, because they are at the extremes where the effect is greatest; removing Gatorade changes the mean number of calories from 64.78 to 66.625, an increase of 1.845; removing Snapple Snap-Up changes the mean number of calories from 64.78 to 62.875, a decrease of 1.905.

6. Answers will vary. It is fair to include the calorie values in the means, because they are comparable to the other drinks; including the price is more problematic. Exceed Powder at \$0.29 is comparable to the other products; however, Gatorade Powder at \$0.11 is much lower and will lower the mean significantly. If the number is to give an accurate value for all of the products available, then the powder should not be included or the median used as a summary measure.

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

7. a. Including the new varieties has less than a $0.01 effect on the mean price. However, it has a 9.33 calorie effect on the mean calories (from 64.78 to 55.45). This is a 14% change.

b.	mean	64.78	0.29
	median	66	0.26
c.	mean	55.45	0.28
	median	60	0.26

For the number of calories, in both cases the mean is less than the median. The difference is greater when the low-calorie drinks are included. For the prices, the median is lower than the mean. The mean is more affected than the median by adding the light drinks in the calorie category.

8. a. $D_T = \sum_{i=1}^{9} D_i$

mean cost $= \frac{1}{9} \sum_{i=1}^{9} \text{Cost}_i$

b. the sum of the costs for the first eight drinks

c. $\frac{1}{11} \sum_{i=1}^{11} D_i$

d. Hydra Fuel has the greatest effect on pulling the mean upward; Gatorade Powder has nearly the same effect on pulling the mean downward.

9. Answers will vary. Students may include information about the mean's being affected by unusually small or large values, or outliers. They may state that the totals would have little meaning because the number of values combined is not known if just the total is given. Students may feel that the median is the best summary in this case. Check students' answers to see if they understand the advantages and disadvantages of the various summary numbers.

STUDENT PAGE 52

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.

10. a. It is always true. Putting the constant before the summation is the same as factoring out the value from the sum. This is true for any constant.

$$\sum_{i=1}^{n} 3x_i = (3x_1 + 3x_2 + \ldots + 3x_n) =$$

$$3(x_1 + x_2 + \ldots + x_n) = \sum_{i=1}^{n} 3x_i$$

b. There is no difference.

11. a. 25 students; four students earned a 2.

b. 64% of the students got a 3 on the test.

c. 2.32; sample formula:

$$\frac{\sum (\text{Score} \cdot \text{Frequency})}{\sum \text{Frequency}} = \frac{48 + 8 + 2 + 0}{25}$$

d. Instead of using the frequencies, the proportions can be used and then no division is needed. $\sum(\text{Score} \cdot \text{Proportion}) = 3(0.64) + 2(0.16) + 1(0.08) = 2.32$

STUDENT PAGE 53

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 was not correct. L3 contains products, and when Tong takes the mean of the column, the calculator divides by 4, the number of entries. The total should be divided by 25 because there were 25 students.

STUDENT PAGE 54

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.

13. a. Possible answers, where D represents distance and G the amount of gas:

$$\frac{D_{city} + D_{country} + D_{express}}{G_{city} + G_{country} + G_{express}};$$

$$\frac{\text{City}_{mpg} + \text{Country}_{mpg} + \text{Express}_{mpg}}{3};$$

the latter formula will be misleading. Students may need to revisit the idea of weighted averages from Lesson 4.

This example might also help:
120 miles at 6 gallons, 20 mpg;
120 miles at 4 gallons, 30 mpg;
120 miles at 3 gallons, 40 mpg

360 miles used 13 gallons for 27.8 mpg; averaging the three different miles per gallon would give $\frac{90}{3} = 30$ mpg.

b. 296; the total miles per gallon for the first 10 cars on the list

c. Small cars: 29.6; mid-sized cars: 21.43

d. If you bought a small car instead of a mid-sized one, you could expect an increase of about 8 miles per gallon. This is because the mean gas mileage for a small car is 29.6 and for a mid-sized car is 21.43.

14. Answers will vary. Students could use the low and high ends of the ranges to get a mean for each interval, or they could find the midpoint of each range, then use either of these as a representative value to find a mean. They might also find a mean for the lower bounds for each range and a mean for the set of upper bounds to obtain a mean range or price interval.

STUDENT PAGE 55

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

15. a. 42% of the families in America have 2 persons.

b. 21% of 2,000 = 420 households

c. Student results may have more than expected in the higher categories because they are members of younger families. As the adults in a family get older, the mean tends to get smaller since the statistic measures only the people living in one house.

16. a. 3.14 persons

b. The calculation assumes that all of the 2% listed in the final category have exactly 7 people in their family.

c. The value will be a little low but will be very close because the percent of people with families having more than 7 persons is very small.

STUDENT PAGE 56

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

17. a. 62

b. $\sum_{i=1}^{4} P_i = 279$

18. a. 279; total number of points earned by the first four students, the juniors

b. 550; total number of points earned by the students not counted in part a, the seniors

c. 829; total number of points earned by all of the students

d. 69.08; mean number of points earned by a student

STUDENT PAGE 57

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 1997 team has the better record. Some students may realize that no calculation is needed. Since the score of the student who was disqualified was lower than the mean, removing that student will raise the mean. Since the mean of the 1997 team was already higher than that of the 1996 team, the record of the 1997 team will remain better. Some students may calculate: multiply 77 by 12 for the total points for the school, 924; subtract 52 for the new total, 872; then divide by 11 to get a new mean of 79.27.

20. **a.** $\sum_{i=5}^{12} 2P_i$; $2\sum_{i=5}^{12} P_i$

b. The first formula

21. **a.** If all of the team members have the same number of points, the total is the number of points times the number of team members, or $70n$.

b. If c is constant and n is the number of team members, then $\sum_{i=1}^{n} c_i = cn$.

22. Possible rules: If there is a constant times the variable, it can be moved outside the summation sign:

$$\sum_{i=1}^{n} cx_i = c\sum_{i=1}^{n} x_i.$$

If the value inside the summation sign is constant, the sum is the constant times the number of iterations indicated by the summation:

$$\sum_{i=1}^{n} c = cn.$$

If a value is added to the variable, it cannot be moved outside the summation sign:

$$\sum_{i=1}^{n} x_i + c \neq \sum_{i=1}^{n} x_i + \Sigma c$$

STUDENT PAGE 58

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

Materials: *Activity Sheet 4,* metric rulers
Technology: graphing calculators or spreadsheets
Pacing: 3 class periods

Overview

Based on their work from Lesson 6 with mean and median as measures of center, students now use formulas to describe the spread or variability in a set of data. They consider the advantages and disadvantages of using just the range as a measure of spread. Students learn to use quartiles and the interquartile range to measure variability around the median and the standard deviation as a measure of variability that corresponds to the mean. In both cases, they address the issue of an *outlier,* or data point, that seems to be an extreme value in relation to the other points. Mathematical rules are used to quantify these outliers, although students are encouraged to think about the appropriateness of the rules and whether they might devise better ones. Students also investigate different forms of symbolic representations to see if they are equivalent. The lesson content centers initially around life expectancies in various countries. The development of the standard deviation is done with data to enable students to understand what the statistic represents.

Teaching Notes

Each measure of center has a corresponding measure of variability. The median is associated with the interquartile range; the mean with the standard deviation.

Center	Variability	Plot
Median	Interquartile range	Box plot
Mean	Standard deviation	Histogram or stem-and-leaf plot

The range is not really adequate as a measure of variability because it does not give any indication of how the data are distributed throughout the interval. They might be clustered at one end or the other, equally distributed throughout the data set, or clustered around the middle. Some other measure is necessary to better understand the distribution, but the range does give the difference between the extremes, which can be helpful in situations comparing temperatures.

Students are expected to have some familiarity with a box plot, but you may need to remind them how to construct one. The box is drawn parallel to a number line around the upper and lower quartiles for the data set with a line through the box indicating the median. Line segments connect the box to the lower and upper extremes. A box plot shows five summary numbers for the data set and can help give a picture of both the center of the data and of the spread. Outliers are sometimes represented by an asterisk with the line segment reaching only to the last data point that is not an outlier. For more details see *Exploring Data* by Landwehr and Watkins, 1996, published by Dale Seymour, Incorporated.

Standard deviation will be a new concept for many students. It is not a difficult idea, but one that is not often introduced to students at this level. It is an example of a formula created according to some reasonable criteria that helps you understand something about the data. The important variables are the mean and difference for each of the data points from the mean. The essential element is to think about finding a measure that tells how data are clustered around the mean. An intuitive method is to find the difference for each data point from the mean. When students begin to write formulas to perform the calculations, however, they discover the need for some rules. Subtracting on a calculator or computer defines a process: The difference between a data point and the mean $(x_i - \bar{x})$ gives some positive and some negative values. Students should recognize that negative values can be assigned to points that are below the mean. If you find an average difference, the sum of the differences equals zero, a surprise to many students. The mean compensates by taking from one large value and giving to another smaller value. Thus, when you sum the differences and some are positive and some are negative, you add and subtract all of the parts used to do the compensation.

This situation motivates the need for eliminating the negatives. A logical step is to take the absolute value, but for mathematical reasons, the procedure followed is to square the differences. (For more information, see *Exploring Least Squares Linear Regression*, Landwehr, Hopfensperger, Burrill, *Data-Driven Mathematics,* Dale Seymour, Incorporated, 1996). These differences are then summed and averaged. The result is still in squared units, so to return to the original units, you take the square root. For some classes, you may want to write the steps out and save the formula until students are very comfortable or need to generalize in some way. Scientific calculators and graphing calculators report the standard deviation as one of the summary statistics. The notation may vary, so be sure students recognize which notation is used for their machine. On the TI-83, *Sx* is the standard deviation for a sample, *x* is the standard deviation for a population. The difference is the divisor: for the sample standard deviation, the divisor is $n - 1$ where n is the number of data points. This helps to adjust for the impact of a small sample size on the standard deviation. For the population standard deviation, the divisor is n. For most situations at this level, it would be reasonable to use the population standard deviation.

Outliers can also be determined based on the mean for a *normal distribution,* one that is symmetric and mound-shaped, where the mean and median coincide and most of the data are clustered in the middle, sloping off to each end. There is a function that explicitly defines a normal distribution in terms of its mean and standard deviation, but that is not necessary to use at this point. Students should just recognize that for a usual distribution that is "normal," approximately 99% of the data should lie within two standard deviations of the mean; about 66% of the data within one standard deviation of the mean.

Follow-Up

Students might search the media or at school for examples such as grades, sports, chemistry experiments, and so on, of situations with outliers; analyze whether the extreme is really an outlier or only unusual; and decide whether an outlier is unusually "good" or "bad." They might also look for situations in which the standard deviation is used.

STUDENT PAGE 59

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.

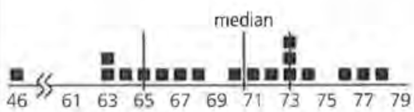
Compare two formulas graphically and numerically.

Solution Key

Discussion and Practice

1. You may have to review how to draw a number-line plot with the class or have one of the students who recalls the procedure go over it with those who do not remember.

a.



Possible answers: There are 18 countries so there are nine on each side of the median. The median of each half is a quartile, the fifth country from each end. Be sure that students understand the use of Q1 and Q3 as representations for the first and third quartiles.

STUDENT PAGE 60

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. IQR = 8; Canada, Cuba, Chile, and the United States; $\frac{2}{9}$ of the countries

c. Haiti seems to have an unusually low life expectancy.

d. 70.5, the median

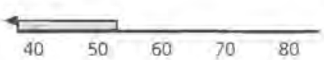
2. a. Possible answers: A country should not be farther away from Q1 or Q3 than twice the distance from Q1 to Q2 or Q2 to Q3. A country should not be farther away from the previous point than twice the distance between any other pair of points. Haiti is an outlier.

b. Possible answer: Outliers are worthy of investigation, and it is important to have a consistent way to identify them. Sometimes they indicate problems that should be investigated, such as health and traffic. Sometimes outliers identify unusually good outcomes, such as athletes' accomplishments and sales.

c. Possible answers: I'd want to be a high outlier for a test grade or income. I would not want to be an outlier for weight.

3. No; $65 - (1.5)(73 - 65) = 53$; $64 \nleq 53$

4. a.



$x \leq 25 - 75 = -50$; $Q1 - 1.5(IQR) = -50$

b. They are the same, because $Q3 - Q1 = IQR$.

c. $46 \leq 65 - 1.5(8)$; $46 \leq 53$

d. No

STUDENT PAGE 61

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.

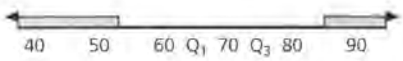
b. How does the rule in problem 3 compare to the rule $x \leq Q_1 - (1.5)$ IQR?

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. **a.** Possible answer: If the point is more than one and a half times the IQR above the third quartile or below the first quartile, then it is an outlier.

b. The "or" could not be replaced by "and," because then there would never be outliers. No number is both greater than Q3 and less than Q1 when Q1 < Q3.



c. $x > Q1 - (1.5)(Q3 - Q1)$ and $x < Q3 + (1.5)(Q3 - Q1)$

d. Not necessarily; it depends on where the least and greatest points lie; there is no guarantee that they will be symmetric around the median.

6. **a.** It is the same; $x \le Q1 - 1.5Q3 + 1.5Q1$, so $x \le 2.5Q1 - 1.5Q3$.

b. $x \ge 2.5Q3 - 1.5Q1$

7. In order to compare data using box plots, the scales must be the same or there is no basis for comparison. It sometimes helps to have two students each draw one of the plots and then try to compare. They will usually use a different scale and often the one with the least IQR looks greater because it is on a different scale.

STUDENT PAGE 62

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*.

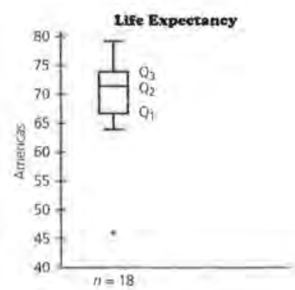
d. Is the distance between the least data point and the lower boundary described by the rule equal to the distance between the greatest data point and the upper boundary described by the rule? Explain your answer.

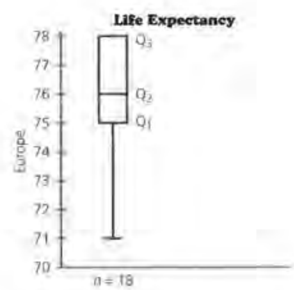
6. Jessica devised this rule for an outlier for small values: x is an outlier if $x \le 2.5\,(Q_1) - 1.5\,(Q_3)$.

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. The scales are different, one from 40 to 80 and the other from 70 to 78; it appears that there is a greater spread for Europe than for the Americas. There is an outlier in the Americas, but none in Europe.

b. The maximum value must be the same as Q3.

c. It is clear that there is a greater range of life expectancies in the Americas than in Europe.

8. Distributions are skewed in the direction of the tail. If the tail is to the right, the distribution is skewed right. Do not make this an important point; it is more important that students recognize the difference in a mound-shaped distribution and one that has the data piled at either end and how the measures of center relate to each type of distribution. You might ask students to think of situations where the data are skewed. (Contributions in a church collection—mostly dollar bills; wages in a company—most of the employees earn a small amount, with a few earning more and more as the level of management rises. Students will suggest many other instances.)

a. It becomes much more symmetric.

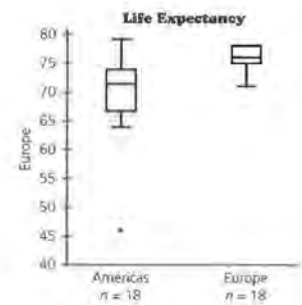
b. It is slightly skewed with the lower numbers forming the tail.

STUDENT PAGE 63

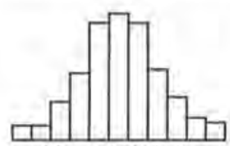
a. Describe the differences between the two box plots.

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

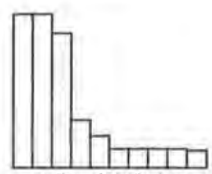
c. The two sets of data are plotted on the same grid below. Does your impression of the differences change? Explain your answer.



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.



A Symmetric Distribution



A Skewed Distribution

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. Students' work should include differences in the spread, comments about outliers, and possibly comments on the values of Q1 – Q3.

10. The median-based measure of variability (IQR) is based on counting ordered data, just as the median is. The mean-based measure is based on the weight or amount of each data point just as the mean is. In the next series of problems, students will generate their own data by measuring their desks. You might wish to assign this as group work, with whole-group discussion and sharing of results. Have each students do his or her own measuring to the nearest millimeter. Students should not talk or discuss the procedure with anyone else in the group, as it is important to have results that vary. If students agree beforehand on a given procedure, the variability will be reduced. You might even use different measuring devices, such as rulers, tape measures, and paper rulers, to increase the variability in the results. As students measure, some will read the ruler incorrectly and others will count the blunt end. Once students have made their measurements, do not allow them to remeasure or correct their errors. Have students imagine they are taking part in an extremely expensive experiment and that it cannot be repeated.

a. Possible data are given. Notice that the difference in the area is much greater than the difference in any one of the measurements, because the product used to find the area increases the error. Suppose the error in length is 0.2 and the error in width is 0.1. The area would be (length + 0.2) · (width + 0.1) = length · width + 0.2 · width + 0.1 · length + 0.02. The two middle terms increase, or compound, the error much more than expected. You might show this to students using one of their lengths and widths as an example.

STUDENT PAGE 64

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?

Dimensions of Desktops

Student	Length mm	Width mm	Area sq mm
Sam	60.5	44.5	2692.25
Tony	60.0	45.2	2712
Lou	60.5	45.0	2722.5
Kelsy	60.4	44.8	2705.92

b. Answers will vary. Some students may use the mean length and mean width of their group, possibly dropping outliers. Others may use the mode of 60.5 for the length, or the median 44.9 for the width to calculate the area. Others may use the mean area. You might question whether using the mean length and the mean width will produce the same mean area. An example should show this is not the case. In the example above, the mean width is 44.875 and the

STUDENT PAGE 65

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.

---

mean length is 60.35 for an area of 2708.20625. The mean area is 2708.1675.

11. a. Center is important because it gives a typical value. Spread is important because it tells how far many of the values are from the center and how the values are distributed around the center.

b. The actual data, any extreme values, and information about how the data cluster

c. Using two numbers is compact and describes two features (center and spread) that are important to understand. Two numbers make it easier to compare two sets of data.

d. It depends on the spread. In some cases the minimum and maximum may be sufficient. However, this doesn't give any kind of density information; the data may be clustered around the mean with an outlier. Using the minimum or maximum would be a misleading description of this distribution. Students may use the life-expectancy data, with the outlier to support their statements. All amounts but one could be at the minimum or the maximum number.

12. a. Find the mean for the values, subtract the mean from each of the data points, then square each result, add the squared differences, divide the sum by the number of data points, and finally take the square root of this quotient.

Student (S_i)	**Area A_i**	$A_i - A$	$(A_i - A)^2$
Sam	2692.25	–5.9175	253.3668
Tony	2712	3.8325	14.6881
Lou	2722.5	14.3325	205.4206
Kelsey	2705.92	–2.2475	5.0513

The sum of the squares is 478.5267; the mean is 119.6317, and the square root of the mean is 10.9376, or about 11 mm^2 as a standard deviation.

When students do their calculations, they may be off slightly because of rounding error. It is usually the best course to keep all possible digits in the computations, and to then round to those digits that make sense in the context of the data for your final answer. Here, the data were measured to the nearest millimeter, so it makes sense to have the answers accurate to square millimeters.

(12) b. Possible answer: The number represents the square root of the average of the squared distances of all of the values from the mean.

c. Groups with little variability will have small standard deviations; those with a large amount of variability will have greater standard deviations. You might have students write their areas and standard deviations on the board and compare them. Ask the class to find an object that has the same area as one of their standard deviations. For example, if the standard deviation is 11 mm^2, an object that is about 5.5 mm by 2 mm will work. The top end of a TI-83 with the screen is about this size. Students have trouble with this reversal; many will not understand that they have to find numbers whose product is 11.

13. Be sure to point out that you cannot take the average of the standard deviations of each group and use it to calculate the average standard deviation for the class. The problem is similar to problems in earlier lessons on weighted average. You might have to work out an example with two or three of the groups to make the point. Also be sure that all students know how to find the standard deviation on their calculators.

a. Answers will vary.

b. If the students have used graphing calculators, they can also make the histogram. The more variability, the more the spread on the graph.

c. Answers will vary. If the distributions are mound-shaped and relatively normal, there will be about 70%.

STUDENT PAGE 66

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

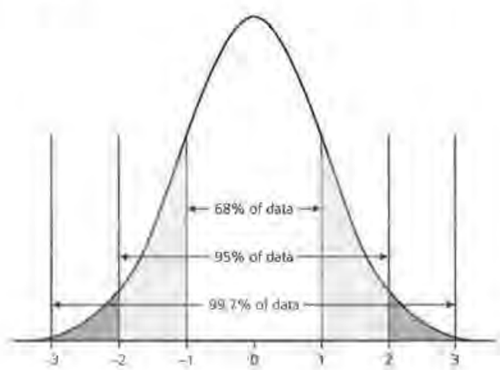
$$\bar{x} - s_x \le A \le \bar{x} + s_x$$

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

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

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 $\bar{x} - s_x$ and $\bar{x} + s_x$ and approximately 95% of the data lies between $\bar{x} - 2s_x$ and $\bar{x} + 2s_x$.



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

d. This can be done by using the vertical line command from the DRAW menu of the graphing calculator. The area represents everything within two standard deviations of the mean.

14. Answers will vary. Encourage students to measure center with the median for skewed data.

15. a. The bars 118 to 170 should be shaded, 144 – 26 and 144 + 26. This appears to be about $\frac{2}{3}$ of the area in the distribution.

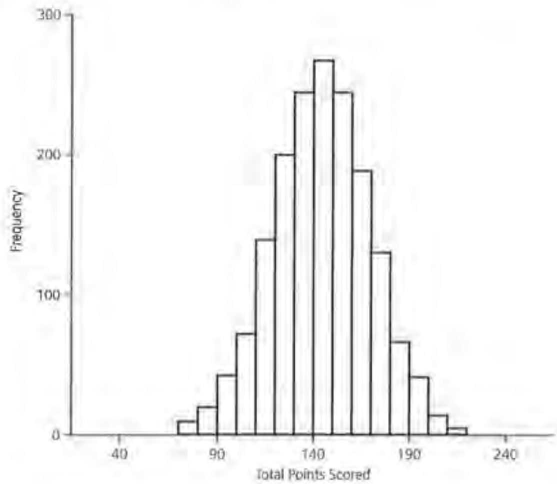
b. The bars 92 to 196 should be shaded. This appears to be about 90% of the graph.

c. The estimated percents from the plot are relatively close.

d. If you look at the plot, you would expect the score to be over 200 less than about 5% of the time. A score over 250 would occur even less frequently, maybe as little as 1% of the time.

STUDENT PAGE 67

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?

Practice and Applications

16. a. Mean: 68.83; median: 70

b. 70.18; it is greater than the mean with Haiti. Haiti's very low life expectancy lowers the mean.

c. With the positive and negative signs, the mean is zero, except for the rounding error. With absolute values of the differences, the average deviation is 5.41. These calculations are for Americas with Haiti.

d. Possible answer: Between 5 and 6

e. It indicates that half of the life expectancies are above the mean and half below the mean; the mean and the median are relatively close together.

17. a. The standard deviation for the Americas without Haiti is 4.76; for Europe it is 2.37. The standard deviation for the Americas is greater than that for Europe.

b. Yes; we can say that there is more variation in the life expectancies in the Americas than in Europe.

c. The standard deviation increases from 4.76 to 7.21.

STUDENT PAGE 68

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?

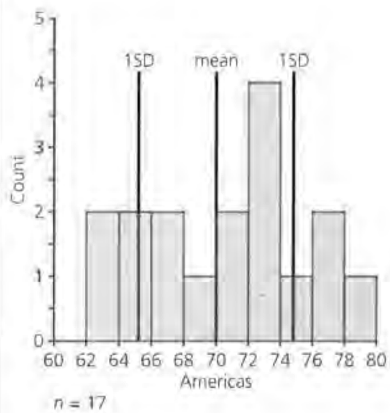
e. About half of these deviations are negative. What does that tell you about the shape of the data distribution?

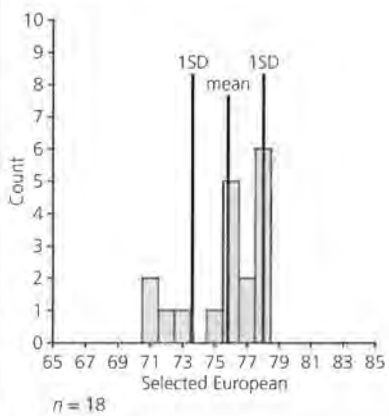
17. a. Calculate the standard deviations for the Americas without Haiti and for the European life expectancies. How do they compare?

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.

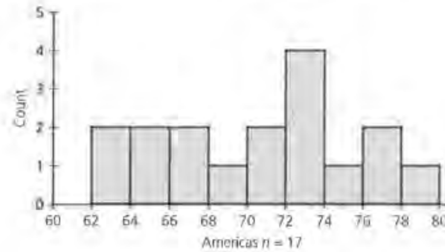


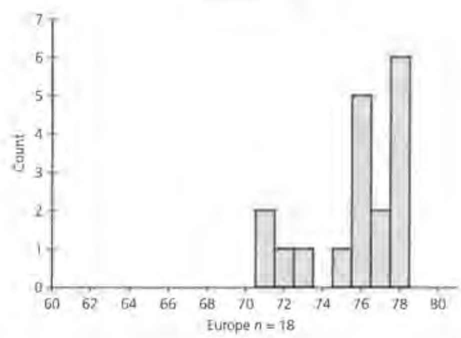


Answers will vary. It appears that the standard deviation is a better measure for the Americas than for Europe because the distribution seems more balanced around the mean for the Americas.

STUDENT PAGE 69

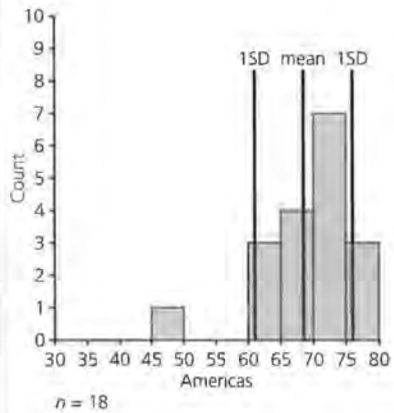
18. Histograms for the life expectancies are given below. The data for the Americas data do not contain Haiti.





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. The standard deviation appears too big, because most of the points fall within the range of one standard deviation.

b. Possible answer: It works better when the distributions are symmetric. If a distribution is skewed, the extreme values make the standard deviation very large and not a typical distance from the mean.

20. a. One with data spread out

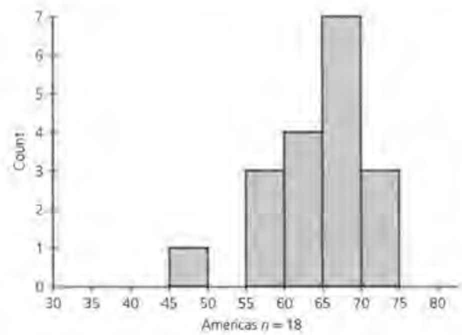
b. Answers will vary. The two standard deviations are the same, but students might not see this initially.

c. They are both 45.

d. Least standard deviation: all 1s or all 8s; greatest standard deviation: five 1s and five 8s

STUDENT PAGE 70

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. The following players are outliers: Ted Williams (1941, 1957), Carl Yastrzemski (1968), Rodney Carew (1977), and George Brett (1980).

STUDENT PAGE 71

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

STUDENT PAGE 72

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

Materials: *Activity Sheets 5* and *6*
Technology: graphing calculators or spreadsheets
Pacing: 2 class periods

Overview

In this lesson, students are asked to compare measurements that are in different units, using both numerical and graphical techniques and building on their understanding of mean and standard deviation. The difference in units makes comparison difficult. Students learn to standardize the data sets by using the mean and standard deviation to derive an equivalent z-score, or distance from the mean in terms of standard deviations, for each data point. This provides an opportunity for students to derive a formula in a different context, one that will facilitate their understanding of data. Using a z-score to compare provides a relative measure and relates each individual score to its position in the distribution of the data; those data that are comparable are the same distance from their respective means measured by their respective standard deviations. z-scores provide another motivation for negative numbers; positive z-scores indicate data above the mean, while negative z-scores indicate data below the mean.

Teaching Notes

Students investigate the ACT and SAT scores from 1970 to 1989. The data do not continue past 1989 because the SAT changed its scoring procedures, and any comparison would be unfair. You might check with SAT to see how they themselves standardized the new scoring procedures and how they make comparisons between scores from prior to 1989 with the current scores. The test has also been substantially redone and and the scoring renormed. These changes add another point of interest in any attempt to get a clear picture of the SAT scores over the years. Similarly, the ACT test was revised in 1990.

Initially, students will have trouble comparing the data sets because the SAT scores and ACT scores are in different units, SAT in 100s and ACT in 10s. A graphical comparison should have the same scale, impossible in this case. To see any changes and whether they are comparable is difficult. Students have encountered this situation in several earlier lessons; for instance, the difference in crime ratings was not the same as the difference in recreation ratings. How does the 0.3-point drop in ACT scores in 1971 compare to the 4-point drop in SAT scores? Students may offer different methods to make their comparison; one example is to divide every SAT score by 20. Consider their options carefully. If you divide each score by 20, you make the numbers nearly comparable. But there is no real basis for choosing 20 rather than 30, and you have not accounted for the difference in ranges: the ACT scores range from 20 to 16.9 while the SAT scores range from 488 to 466. In order to make any reasonable comparison, you need to have some common point from which to begin thinking. The means of the data sets provide this point. Some students will want to use the median; this is reasonable although the IQR as a measure of variability does not involve the weight of the data point, but rather a distance between quartiles.

The z-score is a typical statistic used in inferential statistics to investigate normal distributions. The standard normal distribution has a mean of 0 and a standard deviation of 1; other normal distributions are isomorphic to the standard normal distribution, and each data point can be compared by using z-scores. You do not have to indicate this to students, but they should understand the use of the z-score as a means of stripping the units from a given set of numbers and, consequently, as a way to discuss data sets that have very different measurement units. You might ask students to think about the transformed data and what was gained or lost in terms of understanding the data. The exact values are lost but you have gained some reference measures.

Technology

Students will need to transform each data point using the formula for a z-score, then work with the transformed data. The list capability of a graphing calculator will make this easy and also allow students to see how the data are changed.

Follow-Up

Students could think of other situations where they want to compare results that are in different units and how the comparison might be carried out. Suggestions might include sports statistics or grades given by different teachers (The units might appear to be the same but in actuality reflect different procedures.)

STUDENT PAGE 73

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.

STUDENT PAGE 74

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.

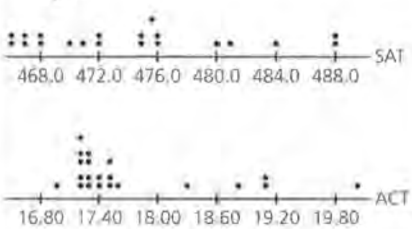
Solution Key

Discussion and Practice

1. Answers will vary. It looks as if the SAT math scores decreased from 1970 to 1981 and then they began increasing again. Some students may notice there was a drop between 1974 and 1975. The ACT math scores follow a similar trend, decreasing from 1970 to 1982 when they became about constant, with the exception of 1983 when the scores were low.

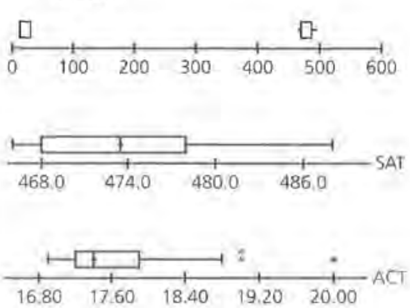
2. Notice that the number-line plots do not show the change over time but rather the total distribution of all the scores. Different plots show different characteristics of the data.

a.



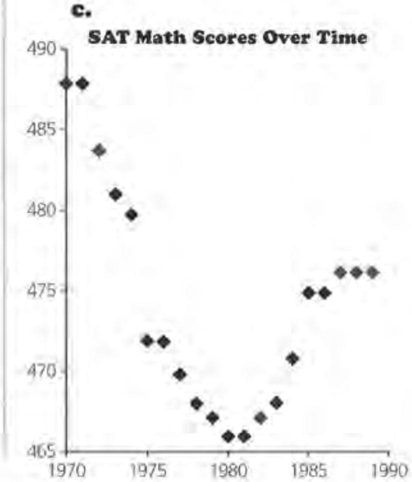
The SAT scores appear to be more evenly distributed than those for the ACT. The ACT appears to be skewed to the right. There is more of a clump in the ACT than in the SAT.

b. It is difficult to make parallel box plots because of the difference in units.



Possible answers: There are 3 outliers in the ACT data, two at the same point just above the cutoff, and one far above. In the SAT box plot, the median is close to the middle of the IQR. In the ACT plot, the median is to the left of the center of the box. Both of the box plots have small left tails.

c.



Possible answers: Both of the plots show an overall downward trend from 1970 until the early 1980s. Then the SAT increases and appears to level out, while the ACT levels out at the minimum score.

d. Answers will vary. Some students may think that the number-line plots are best because they show clusters of scores, the box plots are best because they give a summary, or the plots over time are best because they show trends. The number-line plots show the actual data points that are lost in the box plots.

3. **a.** SAT: mean, 474.3; median, 473.5; standard deviation, 6.959

Possible answers: Because the mean and median are close together, the distribution might be fairly symmetric. The standard deviation seems small compared to the size of the numbers, indicating there is not too much spread.

ACT: mean, 17.75; median, 17.4; standard deviation, 0.840

Possible answers: Because the mean and median are close together, the distribution might be slightly skewed to the right. The standard deviation seems to indicate some spread.

b. They do not allow you to compare because the scales are so different.

STUDENT PAGE 75

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?

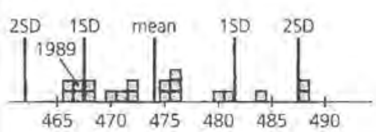
d. What score would be 2 standard deviations from the mean?

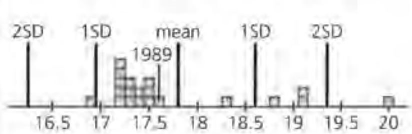
6. The number of standard deviations between a data point and its mean is called a "z score." The z-score formula is $z = \frac{(x_i - \bar{x})}{s}$, where x_i is the given score, $\bar{x}$ is the mean, and s is the standard deviation. You can use the formula with a computer spreadsheet or a calculator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

c. Change both plots to a common scale.

4. **a.**





SAT score: a little more than one standard deviation from the mean; ACT score: about $\frac{1}{4}$ a standard deviation from the mean

b. The ACT score that year was closer to the average score than the SAT. We could say that the SAT scores were lower than the ACT scores in 1989 because they were farther from the mean.

5. **a.** 1 standard deviation

STUDENT PAGE 76

(5) b. You could use a negative number to indicate a score below the mean; –0.5 would indicate a score half a standard deviation below the mean.

c. –1.5 standard deviations

d. 170 and 130 are 2 standard deviations from the mean.

6. a. (See table below.)

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.

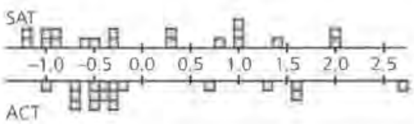
a. Should the city of Minneapolis advertise itself as better for recreation or having a highly rated job market? Explain.

SAT/ACT Scores with *z*-Scores

Year	SAT Math	SAT *z*-score	ACT Math	ACT *z*-score
1970	488	2.0198	20.0	2.7479
1971	488	2.0198	19.1	1.6488
1972	484	1.43	18.8	1.2824
1973	481	.98776	19.1	1.6488
1974	480	.84034	18.3	.67171
1975	472	–.3391	17.6	–.1832
1976	472	–.3391	17.5	–.3053
1977	470	–.6339	17.4	–.4275
1978	468	–.9288	17.5	–.3053
1979	467	–1.076	17.5	–.3053
1980	466	–1.224	17.4	–.4275
1981	466	–1.224	17.3	–.5496
1982	467	–1.076	17.2	–.6717
1983	468	–.9288	16.9	–1.038
1984	471	–.4865	17.3	–.5496
1985	475	.1032	17.2	–.6717
1986	475	.1032	17.3	–.5496
1987	476	.25063	17.2	–.6717
1988	476	.25063	17.2	–.6717
1989	476	.25063	17.2	–.6717

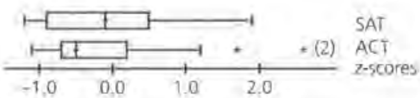
b. 1970; the 2.7479 has the greatest absolute value.

7. a.



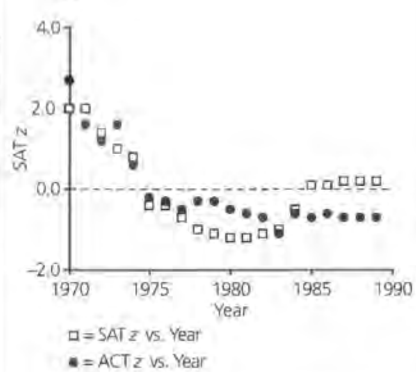
Possible answer: It is clear that there is more clustering below zero in the standardized ACT scores than in the SAT scores.

b.



There are still three outliers in the ACT plot and no outliers in the SAT plot.

c.



Possible answer: From this graph, it is clear that until the late 1970s the scores moved in a similar way. After this point, the SAT scores dipped in relation to the ACT scores and then rose. Both sets leveled off in about 1985, but the SAT was higher than the mean, and the ACT was lower.

d. Student paragraphs should include information about the variation and trends over time.

STUDENT PAGE 77

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	ACT 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.

Practice and Applications

8. a. Better for recreation; it is about 15 standard deviations below the mean in terms of jobs and about one standard deviation above the mean in terms of recreation.

b. Recreation; it is below the mean in jobs and higher than the mean in recreation.

c. Possible answer: To determine limits for outliers, add or subtract twice the standard deviation from the mean. In terms of jobs, anything over 37,077 would be a high outlier. There are none of these. Anything under 29,941 would be a low outlier. All of the cities listed would be job outliers. In terms of recreation, anything over 2,940 would be a high outlier, which would include San Diego and Los Angeles. Anything under 232 would be a low outlier, but there are none of these in the list.

STUDENT PAGE 78

9. **a.** z-scores can be used to compare the two. The z-score for a 500 SAT with the given mean and standard deviation is 0.176. Using this and working backwards will give an ACT score of 21.392.

$$0.176 = \frac{ACT - 20.6}{4.5}$$

$$(4.5)(0.176) = ACT - 20.6$$

$$21.392 = ACT$$

b. 383.56

$$z = \frac{17.2 - 20.6}{4.5}$$

$$z = -0.7555\ldots$$

$$-0.7555\ldots = \frac{SAT - 478}{125}$$

$$-94.44 = SAT - 478$$

$$383.56 = SAT$$

c. The SAT score; the z-score for a 690 on the SAT is 1.696; the z-score for a 26 on the ACT is 1.2. He should send in the SAT score because it is has a higher z-score. The SAT score is more standard deviations above the mean, and comparatively, he did better on that test.

10. Most likely she did not. Most calculators use the standard order of operations, so the calculator would perform the division before the subtraction.

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?

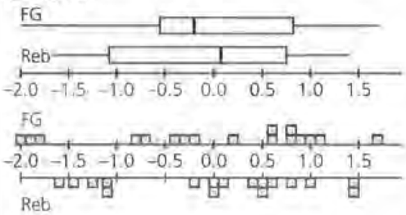
11. **a.** See table below. The rebound rate is the number of rebounds divided by the number of games $R \div G$.

b. rebound rate: mean, 8.10; SD, 3.7758
field-goal %: mean, 0.4749; SD, 0.0416

c.

Year	Player	Field-Goal %	Field-Goal z-scores	Rebound Rate $R \div G$	Rebound Rate z-scores
1991	Archibald, Nate	.467	–0.1899	2.3356	–1.4809
1993	Bellamy, Walt	.516	0.98798	13.6539	1.4272
1990	Bing, Dave	.441	–0.8149	3.7958	–1.1057
1994	Blazejowski, Carol	.546	1.7091	10.0495	0.5011
1991	Cowens, Dave	.460	–0.3582	13.6344	1.4223
1993	Erving, Julius (Dr. J.)	.507	0.77163	6.6998	–0.3595
1991	Gallatin, Harry	.398	–1.849	9.8006	0.43729
1992	Hawkins, Connie	.467	–0.1899	7.9580	–0.03634
1990	Hayes, Elvin	.452	–0.5505	12.4935	1.1291
1993	Issel, Dan	.506	0.7476	7.9485	–0.0387
1990	Johnston, Neil	.444	–0.7428	11.3488	0.8350
1992	Lanier, Bob	.514	0.9399	10.1126	0.5174
1993	McGuire, Dick	.389	–2.065	3.7724	–1.1117
1993	Meyers, Ann	.500	0.60337	8.4433	0.0884
1990	Monroe, Earl	.464	–0.262	3.0194	–1.3052
1993	Murphy, Calvin	.482	0.17067	2.0989	–1.5417
1993	Walton, Bill	.521	1.1082	10.5192	0.6218

(11) d.



He was better at shooting, because in shooting his *z*-score is positive and in rebounds his *z*-score is negative. This means that in shooting, he was above the average, but in rebound rate, he was below the average.

e. It is interesting that scoring seems to be skewed to the left while the rebound rate is skewed to the left. This indicates that there is more variation below the mean in shooting and more variation above the mean in rebounding.

f. Answers will vary. One method is to add the *z*-scores for both attributes for each player and then choose the highest sum. With this method, Walt Bellamy is the best followed by Carol Blazejowski.

12. a. Yes; it would change the sign of all of the *z*-scores; all of the values below the mean would have a positive *z*-score and all those above would have a negative *z*-score.

b. No; they are constants for a particular data set.

STUDENT PAGE 79

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?

ASSESSMENT

Cars

Materials: none
Technology: same as used in lessons
Pacing: 1 class period; you might give it as homework and spend a short time having students compare their work and discuss their results in class.

Overview

In this assessment, students demonstrate what they know about variables and formulas and the use of symbols to work with spreadsheets. They read an article about changing the speed limit and interpret the numbers discussed. They identify the variables mentioned in the article that affect the number of deaths and create a graph that could represent some of the data in the article. They also interpret a formula based on the information in the article. Data about the prices of cars and the corresponding miles per gallon provide the opportunity to use some of the statistical ideas covered in the lessons as well as create a formula to find the cost of owning one of the cars for a year.

Teaching Notes

Students can do the assessment alone or with a partner. It will probably take them more than a class period unless they have been encouraged to do some preparation at home before the day of the test. Be sure students explain what they are thinking and how they find their solutions. The last problem in which students are asked to find a new standard deviation after they have eliminated an outlier will be a challenge. They should be able to reason about the reverse of the formula but may need to do some experimenting with a set of known values before they see how to work out a solution.

STUDENT PAGE 80

ASSESSMENT

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.

Solution Key

1. **a.** Some of the variables that students might mention are: speed, miles driven in a given state, number of cars on the road, age of cars (new ones have better safety equipment), age of highways, (new ones are better designed), and number of drunk driving arrests. Be sure they mention variables that can be quantified or classified in some way. For example, students should not just list *drunk drivers* but should specify some way to quantify them: number of arrests, number involved in accidents, and so on.

b. People are interested in the number of traffic accidents because of safety in driving in that state or area and because one third of all Medicaid costs are the result of traffic accidents and states have to pay for these costs. Young people are the most likely to be in an accident and are the least likely to have health insurance and so become an expense to the state when they are in disabling accidents.

c. Possible graph:

d. Answers will vary. Be sure students mention the variability in the averages. If the speed limit was 55 miles per hour and drivers averaged about 8 mph above the posted limits, they were probably going anywhere from 55 to 71 miles per hour, with the typical driver going about 63 mph. There might be outliers, especially on the higher side which would tend to pull up the average, so it might be more representative to say around 60 mph. After the speed limit was raised, drivers increased their average speed by about 3 mph so they were now typically going 11 miles faster than 55 mph, anywhere from 55 to 77 mph. The average driver would probably be going about 6 miles over the speed limit, but there is no evidence to show that figure.

e. $D = 39{,}235 + (N - 1992)6400$. Student explanations will vary. 39,235 is the number in 1992. N is the year. $N - 1992$ gives the number of years since 1992. You multiply by 6400 to find the number of times you have to add 6400 for a 6400 per year increase. *D* is the number of deaths. In the year 2000, there will be 90,435 deaths according to the formula. This is more than double the number of deaths in 1992.

2. **a.** Small-car mean: 36.3 mpg

b. Compact-car mean: 30.6 mpg; difference: 5.7 mpg

c. Small-car mean: $9,970; compact-car mean: $14,900; difference: $4,930

d. Mid-sized-car mean: $21,980; compact-car mean: $14,900; difference: $7,080

STUDENT PAGE 81

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.

STUDENT PAGE 82

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.

3. **a.** $13,900 + 12,000 m · (1.19 gal)/27 mpg ≈ $14,429

b. Possible formula: 1,000*P* + 12,000 (1.19)/mpg = *TC*, where *P* is the initial price of the car and mpg is the number of miles per gallon listed for that car.

Car	Price *P* ($1,000)	Mpg	Total Cost *TC* ($)
Acura Integra	15.9	31	16,360.65
Dodge Colt	9.2	33	9,632.73
Honda Civic	12.1	46	12,410.44
Hyundai Excel	8.0	33	8,432.73
Mazda 323	8.3	37	8,685.95
Mitsubishi Mirage	10.3	33	10,732.73
Subaru Justy	8.4	37	8,785.95
Suzuki Swift	8.6	43	8,932.09
Toyota Tercel	9.8	37	10,185.95
Volkswagen Fox	9.1	33	9,532.73
Chevrolet Corsica	11.4	34	11,820.00
Chrysler LeBaron	15.8	28	16,310.00
Ford Tempo	11.3	27	11,828.89
Mazda 626	16.5	34	16,920.00
Subaru Legacy	19.5	30	19,976.00
BMW 535i	30.0	30	30,476.00
Chevrolet Lumina	15.9	29	16,392.41
Dodge Dynasty	15.6	27	16,128.89
Ford Taurus	20.2	30	20,676.00
Hyundai Sonata	13.9	27	14,428.89
Lexus SC300	35.2	23	35,820.87
Mitsubishi Diamante	26.1	24	26,695.00
Nissan Maxima	21.5	26	22,049.23
(cont.)			
Pontiac Grand Prix	18.5	27	19,028.89
Toyota Camry	18.2	29	18,692.41
Volvo 850	26.7	28	27,210.00

Answers will vary. Students may choose different cars to be the most economical. Many will select the least expensive, the Hyundai Excel at $8,433, as it was inexpensive to purchase and gets good gas mileage.

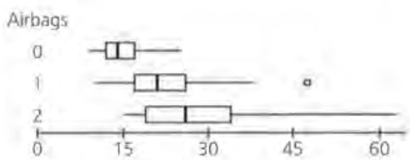
c. No; if you drove 20,000 miles per year, the Hyundai Excel is still the least expensive at $8,721, followed by the Mazda 323 at $8,943 and the Suzuki Swift at $9,153.

4. **a.** Mean: $7,520 for 0 to 1 airbag and $7,150 for 1 to 2 airbags; median: $8,300 for 0 to 1 airbag and $5,650 for 1 to 2 airbags

b. The mean prices must be affected by some very low prices for 2 airbags, which would keep the mean lower than the median.

c. Answers will vary. It would seem logical to use the mean change in price; you could say, "Add another airbag for the same cost."

5. a.



Answers will vary. The prices increase as the number of airbags in the car increases. The price for the lowest three fourths of the cars without airbags is about the same as the price for the lowest fourth of cars with 1 airbag. The median price for cars with 2 airbags is about the same as the lowest three fourths of the cars with 1 airbag. There is one car with 1 airbag that costs much more than any of the other cars with 1 airbag.

b. 4.21 went to the car without airbags; the 8.24 went to the car with 1 airbag; and the 12.55 to the car with 2 airbags. There is more variability in the prices as the cars get larger as evidenced by the ranges and the interquartile ranges.

c. There were 42 cars without the outlier and 43 with. $43 \times 21.22 = 912.46$. Subtract the 47.9 value for the outlier to get 864.56 for the 42 cars. Divide by 42 to get an average of 20.585 per car with 2 airbags.

STUDENT PAGE 83

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.

STUDENT PAGE 84

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 Jounal/Sentinel*, December 10, 1996

UNIT II

Functions

INTRODUCTORY ACTIVITY

Time on Task

Pacing: 1/2 class period

Overview

In this activity, students investigate the data they obtain from collecting information over time. They look at the amount of time they spend each day on the phone to see what patterns might result. They also investigate the relationship between the amount of time they spend studying and the time they spend watching television each day. They study a table of values, the plot, and a cumulative plot—the total for that day and for all of the preceding days. This gives students the opportunity to think about representing data in tables and graphs, and by a rule.

Teaching Notes

This activity is designed to have students think about one quantity as a function of another, that is, when an independent variable determines a dependent variable, as a precursor to Lessons 9 to 12. In the first setting, talking on the phone is presented as a function of time; and in the second, time spent studying is a function of time spent watching television. (In most actual situations, there is no relationship between the time students spend watching television and the time they spend studying, but often students and adults think that one exists. They believe that the more time watching TV, the less time studying. You might have the class collect data to check the hypothesis.)

You may choose to have students do either one of the two parts in the activity or both. They may estimate the time they spend watching television or the time they spend studying. You might want to have students begin to record data a week before they begin Lesson 9 so they have real data to use rather than estimates. You might have them make estimates and then record the actual time spent during a week and compare the two.

The notion of function is intuitive in this activity, in keeping with the lessons. Students are not meant to engage in a formal study of functions but rather to begin to get a feeling for what a function is and how one might behave graphically. The time spent studying is a function of each day. Is the time spent studying a function of the time spent watching television? Do students see any relationship? Each of the results is a function of time, although other variables may also be significant. Collecting data, thinking about a table used to record data, and how this might be reflected in a graph are the focus of Lesson 9. Students will tend to make a line graph for the plot in Problem 1. Raise the issue of continuous functions and ask them what interpretation they would give to the point where $x = 1.5$. This might be confusing, but students should see that for day 1 the total time is 25 minutes. There is no way to make sense of the values in between. This function is strictly defined by the ordered pairs.

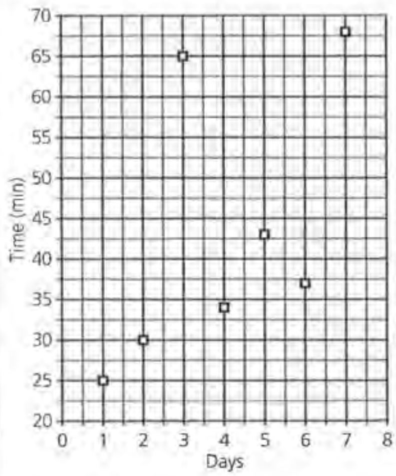
If students visualize the time they spend studying or watching television as a constant, their estimates will be a straight line, discussed in Lessons 10 and 11. When they think about the time over a week as a cumulative function, they will plot the cumulative time or the total time for the first day, then for the first two days and so on, discussed in Lesson 12. Comparing plots and thinking about how each variable behaves as a function of time is a theme of several of the lessons in this unit.

If students think about the time spent as a constant, they will begin thinking about linear relationships. When they are asked to describe an average or typical event based on the plot they have created, they may project the times onto the y-axis and estimate a mean, or they may use the cumulative plot and estimate a median. Some students might think of other ways to make an estimate.

STUDENT PAGE 87

Solution Key

1. Possible answer:



2. If the time were constant each day, the plot would be a series of points along the same horizontal line.

3. **a.** Answers are given for the sample data.

Day	Minutes	Cumulative Minutes
1	25	25
2	30	55
3	65	120
4	34	154
5	43	197
6	37	234
7	68	302

b. Possible answers: Add the values or add the increments on the plot from day to day.

c. If time is spent on the phone each day, the cumulative graph will increase throughout the week. The slope over each one-day interval will relate to the length of time spent on the phone during the corresponding day. The greater the slope, the greater the amount of time spent on the phone during that day.

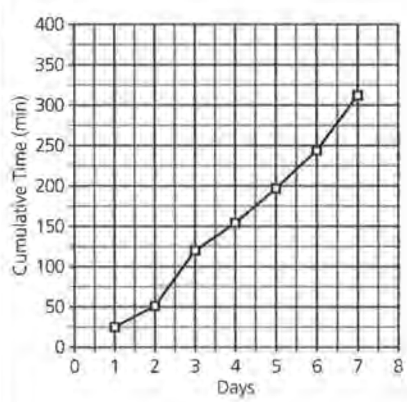
INTRODUCTORY ACTIVITY

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?

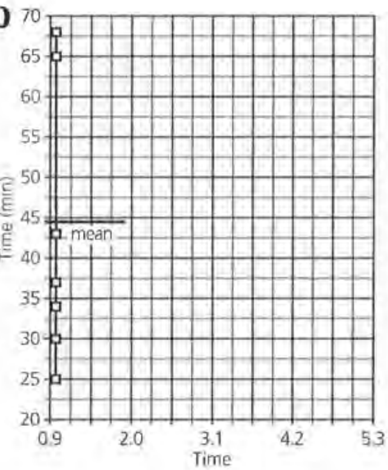
1. Estimate how many minutes you think you will talk on the phone each day for the next seven days. Make a plot with days along the horizontal axis and minutes on the phone along the vertical axis.
2. If you spent the same amount of time each day on the phone, what would the plot look like? Explain your 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?



4. A typical amount of time spent on the phone could be a mean, a median, or some other statistic students might suggest. The mean can be found by thinking about each of the data points in the first plot as projected onto the y-axis. In this case, the mean would be about 44. The median could also be used, in this case, 37 minutes per day.

STUDENT PAGE 88

(4)



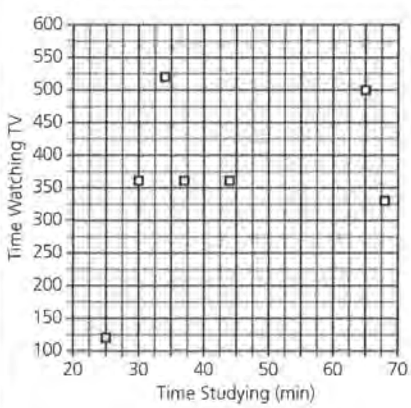
5. Answers will vary.

6. Students should realize they need to have comparable scales in order to make a comparison. Ask them to decide whose plot looks more consistent over the week and whose has the most variability. How can they see this in their plots?

7. Answers will vary.

8. Answers will vary.

a. Possible answer:



b. There is no relationship between time watching television and time spent studying in this plot.

c. If "being a function of" implies there is some relationship that can be observed, there is none in these data. Some students may have a clear relationship, possibly linear, where the more time spent studying, the less time spent watching television. Again, answers will vary. They should, however, recognize that they are looking for patterns in the plot and their data that can be described by some mathematical regularity.

---

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

Materials: *Activity Sheet 7,* graph paper
Technology: graphing calculators or spreadsheets
Pacing: 1 class period

Overview

In this lesson, students learn that in some relationships between sets of data, one variable depends on another and that this dependency can be visualized graphically, numerically, and symbolically. Although the ideas are not formalized, the problems provide an introduction to the concept of function. Students see how a rule can be represented by ordered pairs of data and how these ordered pairs can be plotted. They work with relationships that are linear and quadratic as well as reinforce some of their earlier work with inequalities.

Teaching Notes

Using speed, reaction distance, and stopping distance for driving, students investigate linear and quadratic relationships. You might begin the lesson by having students think about some of the measurements involved in the lesson: How long is a typical car? How long does it take to stop a car after you step on the brakes? If you are going fast, will it take longer to stop than if you are going slow? How do you think the police measure stopping distance and speed in an accident? The fact that some of these questions can be answered by using a regular pattern between speed and reaction distance or stopping distance will be new to many of the students. Again, they are using a spreadsheet format to do some calculations. Although an informal definition of function is given, do not place any emphasis on it. Rather, let students begin to understand what it might mean to say that stopping distance is a function of speed. Be sure students understand the difference between braking distance and stopping distance.

Reinforce the relationships between the table of values, the graphical representation, and the formulas. The patterns that appear in the graph may be visible in the table. Have students locate the ordered pairs from the table on the plot and discuss what the points indicate about the distances and speed. Follow the same procedure, beginning with the plot. Students need to assimilate the different representations and what they can learn from them. Discuss the difference in what can be learned from each form of representation. The graphical representations here are continuous; the ordered pairs on the graph are not just the discrete elements from the table. Is it always correct to move from the discrete to the continuous (connect the points)? Think of situations where this is not the case, for example, (year, number of award winners). You can often see trends more easily in a graph than a table, especially if the independent variables are not spaced at regular intervals. Students may offer other suggestions.

For the reaction distance, the differences are constant; for every 10-mile-per-hour increase in speed, the reaction distance increases 11 feet. For stopping distance and braking distance, the differences are not constant, but the pattern that emerges graphically in each case is curved. Students should realize the dependence of each of these distances on the speed, an essential element in thinking about a functional relationship. They should also realize that the independent variable (speed) is plotted on the horizontal axis and the dependent variable is plotted on the vertical axis.

Technology

Students should do some of the plots by hand, but encourage them to use a calculator or spreadsheet for the calculations. The relationship between the table of values, the graph, and the symbolic relationship is reinforced by using a graphing utility.

Follow-Up

Students should think of factors other than speed that might influence stopping distance.

Have students check with a driver-education instructor to verify the distances. Some students might be interested in talking to the highway patrol or local police force about how they measure speed in accidents and how they use skid marks and other evidence to make their judgments. Students could also be encouraged to discuss the lesson with those in their family that drive cars. Have them share the information, and then write a report indicating what the drivers in their family already knew about the stopping distances.

STUDENT PAGE 89

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.

Solution Key

Discussion and Practice

1. a. 11 feet

   b. Distance per mile per hour

S (mph)	R (feet)	$\frac{R}{S}$
20	22	1.1
30	33	1.1
40	44	1.1
50	55	1.1
60	66	1.1
70	77	1.1

   $\frac{R}{S}$ is a constant, 1.1.

   c. 38.5 feet

   d. $R = 1.1S$ or $\frac{R}{S} = 1.1$. Some students might be curious why both of these represent the same rule. Take time to discuss the point, either by showing the equivalent algebraic equations and expressions or by using examples or definitions.

2. a. 60.5 feet

   b. 88 feet

   c. 110 feet

3. About 54.55 miles per hour

STUDENT PAGE 90

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.

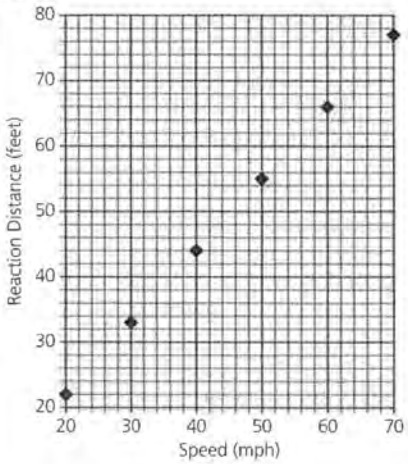
Under ideal conditions, we can predict R perfectly from S by using the equation $R = (1.1)S$, regardless of the value of S.

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?

STUDENT PAGE 91

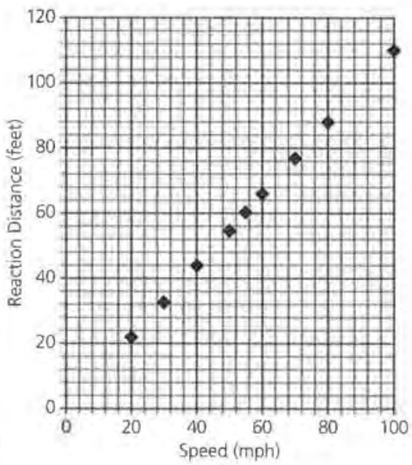
4.



a. A straight line

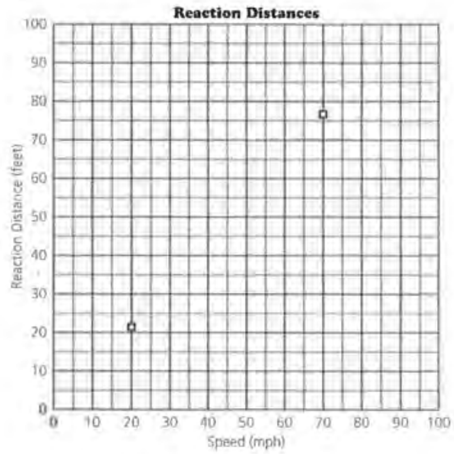
b. From the graph it is about 70; with the formula it is 71.5 feet.

c. A straight line



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?

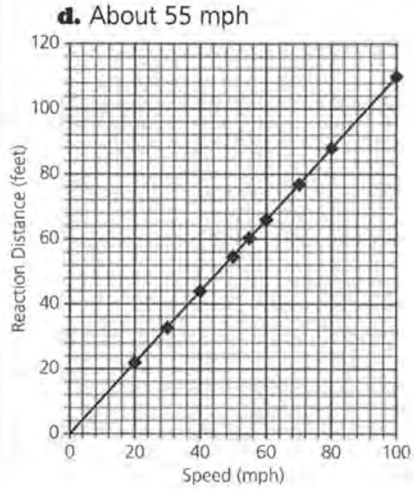
b. Use the plot to approximate the reaction distance for a car traveling at 65 miles per hour. Then calculate R when S is 65 miles per hour. How do the two values compare?

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?

d. About 55 mph

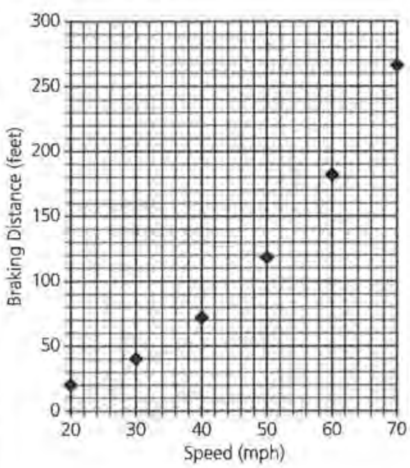


5. Yes, because you could be going any speed between the speeds given both speed and braking distance are continuous. The line allows you to approximate the stopping distance for any speed.

6. **a.** You could give a rule such as $R = 1.1S$, a table like Table 9.1 in the text, or a graph like the one for 4d.

b. Possible answer: The rule gives a description for finding the value at any point. The table gives a set of particular values and might make it easy to see patterns. The graph allows you to estimate many values and shows any trends in the data.

7. **a.** Curve



b. Answers will vary. A typical response might be that the braking distance gets longer and longer as the speed increases. (The braking distance does not increase at the same rate as the reaction distance did, but it does increase.)

8. **a.** About 150 feet

b. About 360 feet

STUDENT PAGE 92

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

STUDENT PAGE 93

(8) c. About 50 mph

9. a. $T = B + S$; See table below.

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.

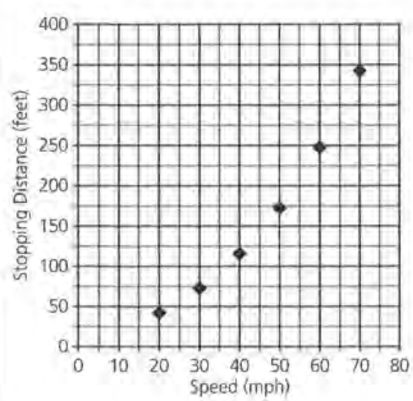
Practice and Applications

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.

Speed and All Three Distances

Speed S (mph)	Reaction R (feet)	Braking B (feet)	Stopping T (feet)
20	22	20	42
30	33	40	73
40	44	72	116
50	55	118	173
60	66	182	248
70	77	266	343

b. A curve flatter than the graph of (S, B)



c. About 220 feet

d. About 400 feet

e. About 75 mph

10. a. $\frac{73}{17} \approx 4.3$ car lengths; to be safe, allow 5 car lengths. Ask students why it makes sense to round up.

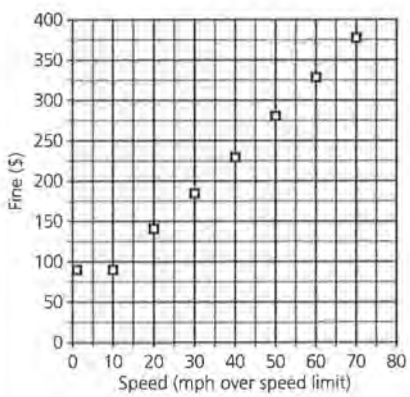
b. $\frac{343}{17} \approx 20.1$, so leave 21 car lengths; $\frac{60.5 + 150}{17} \approx 12.4$, so leave 13 car lengths. (Students will have to interpolate for 55 mph.)

Practice and Applications

11. a. $89.60 + 5 · $4.80 = $113.60

b. $89.60 + 15 · $4.80 = $161.60

c. About $250



d. About 32 miles over the speed limit

12. a. $6,418.40

b. Possible answer: Yes, because each mile increases the cost by 45.77 cents no matter how many miles driven.

c. Possible answers: You could make a graph of (miles driven, cost); the plot will be a line; find 12,000 miles and go up to the plot and then over to find the cost.

The values on a table will have a constant difference of $457.70 for every 1000 miles; if a table has values for every 1000 miles, then the cost for 12000 miles could be found directly; if not, some estimating between given values would be needed.

A formula is $C = 926 + 0.4577M$, where M is the number of miles driven and C is the cost; evaluate the formula for $M = 12,000$. Some students may recognize that this is the form of a linear equation.

STUDENT PAGE 94

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:

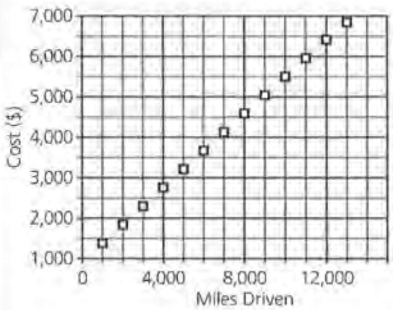
Gas and oil: 6 cents per mile
Maintenance: 2.2 cents per mile
Tires: 0.91 cents per mile

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.

Miles Driven	Cost ($)
1,000	1,383.70
2,000	1,841.40
3,000	2,299.10
4,000	2,756.80
5,000	3,214.50
6,000	3,672.20
7,000	4,129.90
8,000	4,587.60
9,000	5,045.30
10,000	5,503.00
11,000	5,960.70
12,000	6,418.40



13. a. Answers will vary. Students could use either because they are equivalent. The second might be easier because there aren't as many chances of pushing the wrong calculator buttons or doing the wrong calculation. Some students may have to experiment with the graph or a table of values to see that the expressions are equivalent. Others may recognize that you can combine the coefficients for the number of miles *M*.

b. From problem 11, if you owned the car it would cost $6,418.40 to drive 12,000 miles. Leasing would cost $3,961.20, so for one year it would be cheaper to lease.

c. If you drive less than about 5,000 miles per year, then it is cheaper to own a car using the cost information given. Otherwise, it is cheaper to lease. Students can use either a table or a graph of the two functions to answer the question. If they use a table, students will see that they have to use intervals less than 1,000 miles in order to estimate their answer. This is a good place for them to make use of the table function on a graphing calculator; have them set the increment at 25 miles beginning from 4,000, the last value where the cost to own is less than the cost to lease.

STUDENT PAGE 95

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

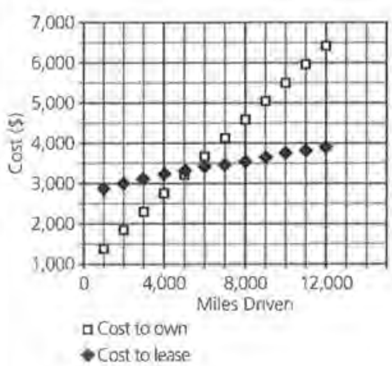
15. In an earlier lesson, you combined variables to make new variables. You can also combine formulas to make new formulas.

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.

Miles Driven	Cost to Own ($)	Cost to Lease ($)
1,000	1,383.70	2,959.10
2,000	1,841.40	3,050.20
3,000	2,299.10	3,141.30
4,000	2,756.80	3,232.40
5,000	3,214.50	3,323.50
6,000	3,672.20	3,414.60
7,000	4,129.90	3,505.70
8,000	4,587.60	3,596.80
9,000	5,045.30	3,687.90
10,000	5,503.00	3,779.00
11,000	5,960.70	3,870.10
12,000	6,418.40	3,961.20



d. Answers will vary. This depends on how long the student thinks he or she will own the car and how many miles per year he or she expects to drive. It appears that with these values it is better to lease. If you expect to keep the car for 10 years and not recoup any cost when you get rid of it, it is more cost-effective to buy only if you drive less than 2,500 miles per year. This assumes you can lease the car using the given costs.

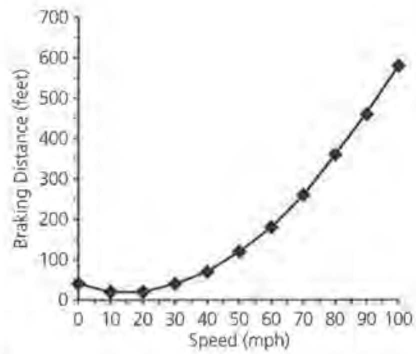
14. a. 264.4 feet, about 15.5 car lengths, or close to the length of one football field

b. No; the intervals in the *B* column increase in value, while those in the *S* column are constant.

Sample values are given.

S	*B*
20	21.4
30	38
40	70.6
50	119.2
60	183.8

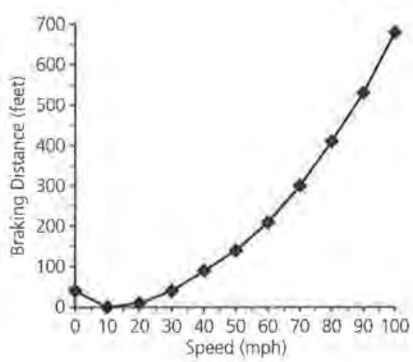
c. No



15. a. $R = 1.1S$; $B = 36.2 - 2.34S + 0.08S^2$

$T = R + B = 1.1S + 36.2 - 2.34S + 0.08S^2 = 36.2 - 1.24S + 0.08S^2$

b. After you reach a speed of about 20 mph, the distance needed to stop increases more and more for each increase in speed.



c. Students' responses should discuss when a relationship is linear and when it is not. Some may include information about the rate of change in stopping distances.

LESSON 10

Trends over Time

Materials: graph paper
Technology: graphing calculators or spreadsheets
Pacing: 1 class period

Overview

Students use contextual situations to develop linear formulas from information conveyed in words or tables. In the process, they revisit and reinforce their understanding of inequalities and how these can be written and expressed graphically. Students are introduced informally to slope when they consider a constant change in one value for a given change in the other, learn to express slope as a ratio, and relate slope to the plot. They begin to think about the relationship between data given in words, in a table and in a plot, and to compare the information they can obtain from each representation.

Teaching Notes

In the first context, students investigate the age of cats in human years. The rule that determines equivalency is a split function, one for ages from 0 to 1, another for ages from 1 to 2, and the rest for ages over 2. This introduces students to the concept of the domain of a function and how it can impact the range or dependent variable. The independent variable produces a response according to the rule or given information. Many situations in the world are functional, where the rule changes for a specified set of domain elements, for instance postal rates and theater tickets. Students should begin to recognize that when a function is linear, the ratio, or slope, will be constant and can be used to continue the established pattern. They should also begin to recognize that the graphical representation will be a straight line, and the difference in dependent values (y values) for a constant difference in the independent values (x values) will be constant.

The concept of disposable personal income, or DPI, may be new to students. Be sure they understand that disposable personal income is the amount of money someone has left after he or she has paid taxes on their earnings. All of their living expenses must be paid with their DPI. Again, students should recognize that when the ratio is constant, the plot will be straight.

Technology

While it is important to do some of the calculations by hand, students should be allowed to use technology such as a spreadsheet or a graphing calculator. The data are available on the data disk that accompanies this module and can be downloaded into a spreadsheet.

Follow-Up

Students might investigate their own disposable personal income and how they spend it over time. They might also check the newspapers or almanacs for plots over time, bring examples to class, and share them orally or through written reports. The reports, along with the original articles, can be used to make bulletin-board displays. These displays not only showcase student work but can be used to promote class discussion of the articles and ways students responded.

STUDENT PAGE 96

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

STUDENT PAGE 97

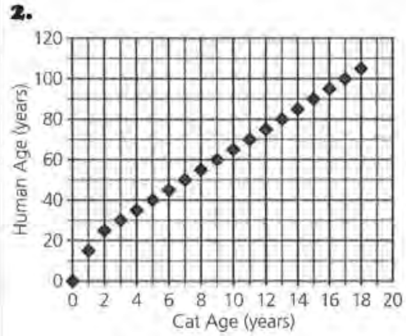
Solution Key

Discussion and Practice

1. The first row contains (0, 0) because it makes sense to think of a newborn cat and a newborn human as the same age.

C (years)	H (years)
0	0
1	15
2	25
3	30
4	35
5	40
6	45
7	50
8	55
9	60
10	65
11	70
12	75
13	80
14	85
15	90
16	95
17	100
18	105

2.



a. After 2 cat years, the graph is a straight line. It curves between 0 and 2 cat years of age.

As students continue to investigate the problem, they should recognize

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?

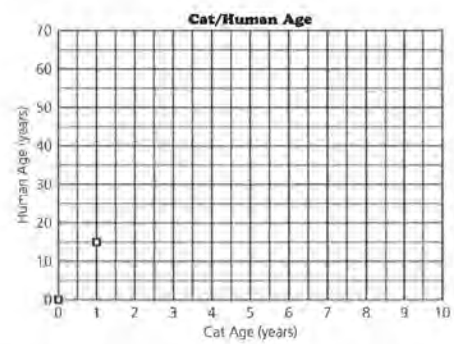
Discussion and Practice

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)?

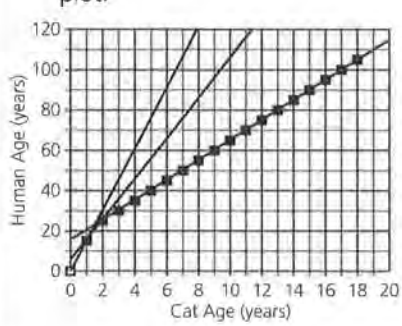
Table 10.1
Cat/Human Age

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.



that there are three distinct rates or ratios as indicated in the following plot.



The graph is three straight line segments: one with slope 15 between ages 0 and 1, one with slope 10 between 1 and 2, and one with slope 5 from age 3 on.

b. The equivalent human age is almost 6 human years. Find 0.375 (which is 4.5 ÷ 12) on the *C*-axis, and follow the graph to read the approximate *H*-value on the *H*-axis.

3. **a.** Students may use the plot, estimate, or think that the ratio of cat to human years is 1 to 15. So, for 1 month it would be 1 to 1.25.

   b. 15:1

   c. 3.75 years, or 45 months (3 times 1.25, the ratio per month)

   d. Values such as 0.167 and 0.250 years (2 and 3 months) are so close together on the *C*-axis that it is hard to find accurate values on the *H*-axis. This is not a problem if ratios are used.

4. **a.** The age in human years will increase by about 10 months, or a little less than one year. No, the ratio between $C = 1$ and $C = 2$ is 1:10.

   b. About 20 human years, 15 from the first cat year and about 5 for half a cat year between $C = 1$ and $C = 2$

5. **a.** The period of time after the cat's second birthday up to and including its third birthday

   b. 10:1

   c. 5:1

   d. If $C > 2$, the H:C ratio is always 5:1.

6. **a.** 85 years

STUDENT PAGE 98

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?

(6) b. 9 years

c. Answers will vary. Students might include that the cat is ill or injured or even a special breed that has a different life span.

7. Perhaps the cost of living has risen at a greater rate than salaries have.

8. Per person

STUDENT PAGE 99

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

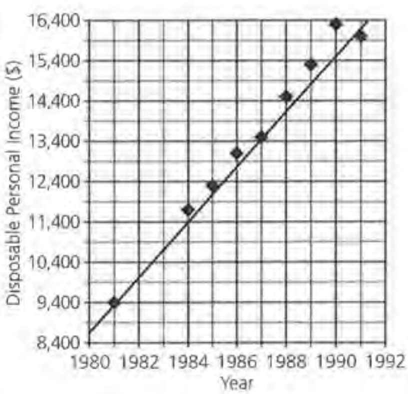
Source: *Statistical Abstract of the United States*, 1994

STUDENT PAGE 100

9. a. It's not very useful. With an x-axis that ranges from year 0 to year 2250, the nine years for which you have data become compacted nearly into a vertical line. It's very difficult to determine which y-value goes with which year, or to get an idea of trend.

b. No; change the x-axis so it covers a much smaller range of years, such as 1980 to 1992. It is not helpful to have a graph from 0 to 1900 with no data points.

10. a. It seems to increase about the same amount every year until 1990 and then it decreases.

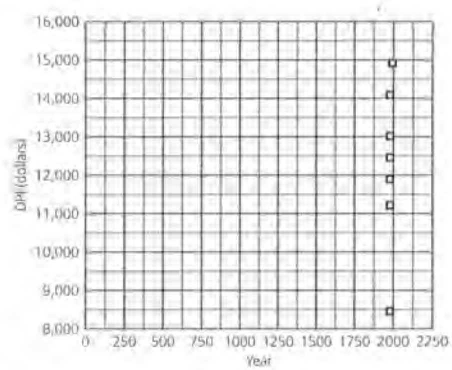


b. About $667 per year; answers will depend on which line was drawn.

11. a. $666, $671, $535, $932, $830, $867, and $-516; the average gain is $569.

b. Answers will vary, but the effects of the negative increase from 1990 to 1991 will place most answers to Problem 11a less than the answers to 10. It seems to be about $100 lower.

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.

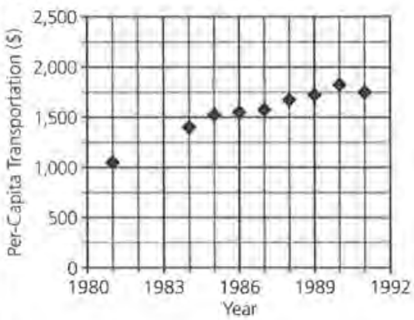
b. Use a straightedge to draw a single straight line close to the data points that you think best represents the trend in these data. Use the line you drew to answer this question: About how much does the DPI tend to increase each year?

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?

Practice and Applications

12. a.

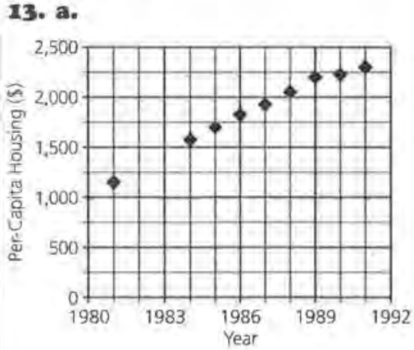


Answers will vary. Be sure that students notice there is a span of 3 years from 1981 to 1984 on the table. They may make the wrong scale on the graph if they do not use this fact.

b. Answers will vary. There is a steady increase, then a flattening-out period, and then a steady increase again. If two different lines are preferred, the break would likely be somewhere in the range 1985–1987.

c. If the break is somewhere in 1985–1987, then the ratio for the first line will be less than the ratio for the second. The ratio for 1981–1985 would be about $114 per year; for 1986–1992 about $50 per year.

13. a.



Per-capita housing expenditures seems to increase by about $125 per year.

STUDENT PAGE 101

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?

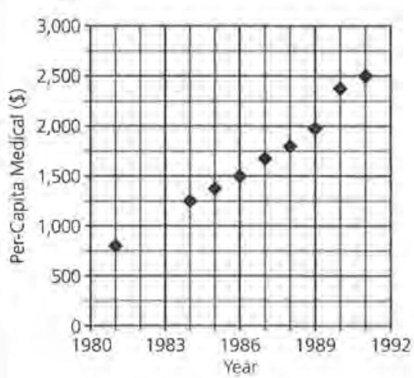
13. Divide the tasks in parts a and b among your group members. Share your results to answer part c.

- **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.

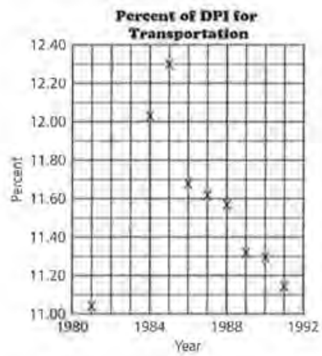


Per-capita medical expenditures increase by about $115 each year from 1981 to 1986. Then the increases were greater and especially high in 1989.

STUDENT PAGE 102

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.

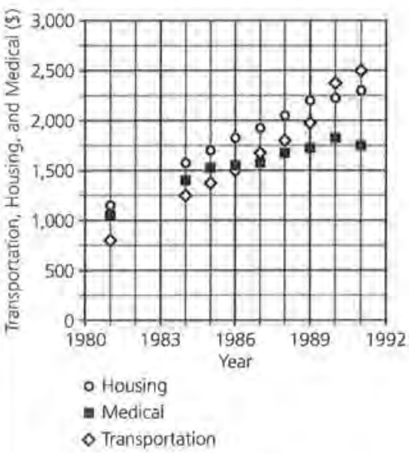
Food expenditures as a percent of DPI over time
Housing expenditures as a percent of DPI over time
Medical expenditures as a percent of DPI over time

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.

c.



Answers will vary. In general, all values increased almost every year, but things become less regular starting around 1990. There seems to be a sudden rise in medical expenses, while transportation has decreased. Housing seems to continue at the same rate of increase.

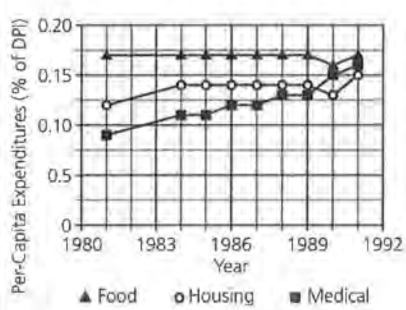
14. a. 17.0%

b. 11.2%

c. The percent spent on transportation rose until 1985 and then decreased sharply, followed by a gradual decrease.

15. a.

Year	Food	Housing	Medical
1981	17%	12%	9%
1982	—	—	—
1983	—	—	—
1984	17%	14%	11%
1985	17%	14%	11%
1986	17%	14%	12%
1987	17%	14%	12%
1988	17%	14%	13%
1989	17%	14%	13%
1990	16%	14%	15%
1991	17%	15%	17%



Possible answer: Food and housing percents seem to have very similar trends, but the graph of the percents for medical costs seems to be different. It rose consistently from 1989 while others decreased and then increased again.

b. Possible answer: The large increases in medical costs since 1989 could be cause for concern, especially if the trends continue. Medical costs are taking a greater and greater percent of DPI. This might be due to the aging population, which will be studied in more detail in Lesson 11.

16. Answers will vary. Be sure students use the data and the percents in their discussion.

LESSON 11

Exponents and Growth

Materials: none
Technology: graphing calculators or spreadsheets
Pacing: 2 class periods

Overview

Students investigate data that are exponential and contrast them to data that are linear. They study graphical and numerical representations using the context of population growth. The focus is on functions that are not linear, in contrast to those in Lesson 10. Students should begin to see that the table values and the graph behave differently for the two lessons. The change in the dependent variable is constant for a linear relationship, while it is not constant for an exponential relationship. Students should also recognize that the rule is different, primarily because the independent variable in an exponential relationship is an exponent. This is what makes the growth change so rapidly. The purpose is not to have students understand exponential functions, but to enable students to see exponential change in contrast to linear change. They should think about the difference in the way the dependent values change for a given change in the independent value. Students should recognize that initially there is often little difference between something that grows linearly and something that grows exponentially, but at some point, exponential growth increases at an increasing rate.

Teaching Notes

Students should compare a ratio from a plot in Lesson 10 with the ratio from the population growth and recognize that the first is constant for a given unit change in time while the second is not. Be sure students understand that you can calculate and compare the ratios for a given plot only if the change in time is equal; for example, the ratio must be unit change per year. This is an introduction to slope, but the concept does not need to be made explicit at this point. Students should, however, understand that the concept involves thinking about the change in the dependent variable per given change in the independent variable; that is, the change in y values divided by the change in the corresponding x values.

The scale used for the graphs can be significant. If the y-axis has very small intervals, data that are curved will look linear. Using the table function of a calculator or spreadsheet to generate table values can help students see the patterns.

Students may have difficulty understanding that a given percent increase is not a constant rate of change. Have students make a table of values and look at the graph for a 5% increase in price, and then contrast it with a constant increase of $5 per year. The 5% increase is an illustration of recursive thinking, in which the next element is defined in terms of the previous one.

Technology

While it is important to do some of the calculations by hand, if possible, students should be allowed to use technology such as a spreadsheet or a graphing calculator.

Follow-Up

Bring in articles about population growth and have students check the stories for accuracy. Are the words used correctly? What conclusions were drawn, and do the facts supplied support them?

STUDENT PAGE 103

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.

STUDENT PAGE 104

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

STUDENT PAGE 105

Solution Key

Discussion and Practice

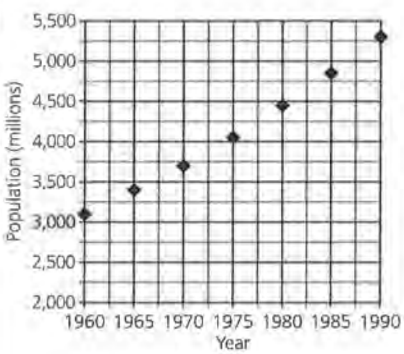
1. **a.** The increase, expressed as a percent, of the population from one year to the next

b. Sub-Saharan Africa

c. It is fast compared with other nations.

d. The population is shrinking.

2. **a.** Yes; it looks straight at first glance, but it actually has a slight curve.



b.

Year	Population (millions)	Amount of Increase (millions)
1960	3049	
1965	3358	309
1970	3721	363
1975	4103	382
1980	4473	370
1985	4865	392
1990	5380	515

The amount of the differences increases each year except for 1975 to 1980. There is an increase from 1985 to 1990 and again from 1965 to 1970.

3. **a.** 10.13%

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?

b.

Year	Population (millions)	Amount of Increase (millions)	Percent of Increase
1960	3049		
1965	3358	309	10.13%
1970	3721	363	10.81%
1975	4103	382	10.27%
1980	4473	370	9.02%
1985	4865	392	8.76%
1990	5380	515	10.59%

There is no pattern in the table.

(3) c. About 10%

d. Yes; the prediction in part c for about a 10% increase would mean that the increase is about 500 million per five years right now, or about 100 million per year.

4. A 10% increase every five years gives the following results:

Year	Population (billions)
1990	5,380.00
1995	5,918.00
2000	6,509.80
2005	7,160.78
2010	7,876.86
2015	8,664.54
2020	9,531.00
2025	10,484.10
2030	11,532.50
2035	12,685.76

8.5 billion is low compared to the prediction in the table, assuming an increase of 10% every five years.

5. **a.** Between 35 and 40 years, that is, between 2025 and 2030

b. Yes, as long as the growth rate is constant. One response might be to continue the sequence of population and year. Others might see a pattern in the numbers; some may use the table function on the graphing calculator.

c. Answers will vary. Students may be surprised that the population is growing so much. They may also be interested in the decreasing amount of time it takes for the population to double.

d. By using different starting populations, students should discover that it will take between 22 and 23 years for the population to double with a 3% growth rate. A different growth rate could affect this projection.

STUDENT PAGE 106

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?

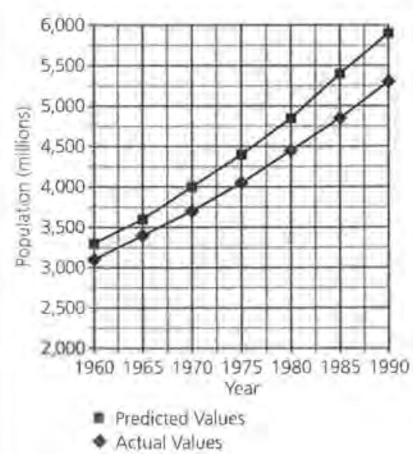
c. What will happen to the function and its graph if the formula is written $P = 993.65\ (1.02^x)$?

d. What does the formula $P = 993.65(1.02^x)$ predict for the population in 2010? Do you think this prediction is reasonable?

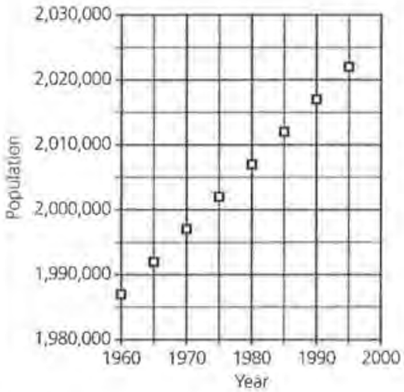
7. Cellular-phone use is growing rapidly. A newspaper article contained the following information: "The number of cellular-phone subscribers has grown in the past decade from fewer than 100,000 in 1984 to more than 19 million. The number is increasing by 17,000 per day." A graph with more specific numbers follows.

6. **a.** 1987, 5564.78 million people; 1995, 6520.03 million people

b. The predicted values appear to be high compared to the actual values, but the shape is similar.



c. The prediction is a straight line rather than a curve, and the line would be far above the actual values.



d. About 8775.11 million

7. **a.** The number of cellular phones is growing dramatically each year.

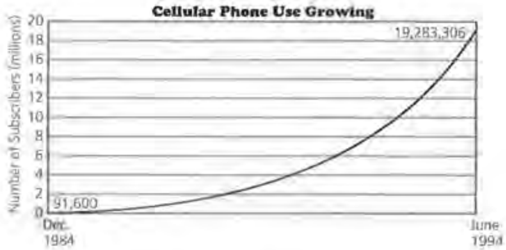
b. Answers will vary. Students could check the formula for a number of the data points that are given. The formula appears to be fairly good.

c. It doesn't seem to fit very well. Using 1994 (94) in the formula gives 27,923,401 subscribers, many more than the 19,283,306 given in the article and graph.

d. It has a similar shape.

e. Using the formula, the number in 1994 is 27,923,401 and in 1995 is 43,560,506. This is an increase of 15,637,105 in one year or about 42,841 per day, which is a lot more than 17,000.

STUDENT PAGE 107



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

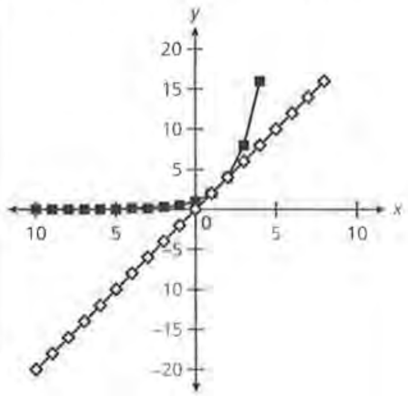
c. Does the data point for 1994 given in the article and graph fit in the formula?

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.

Practice and Applications

8. **a.** The graphs are about the same for only a short interval between $x = 1$ and $x = 2$. For $y = 2x$, the graph is a straight line; the graph for $y = 2^x$ is a curve where y increases rapidly after $x = 2$.



b. For $x < 1$ and for $x > 2$, $2x$ is less than 2^x.

x	$2x$	2^x
–5	–10	.03125
–4	–8	.0625
–3	–6	.125
–1	–2	.5
0	0	1
1	2	2
2	4	4
3	6	8
4	8	16

c. Yes, for $1 < x < 2$; the functions are equivalent at $x = 1$ and $x = 2$; for $x = 1.5$, $2x$ is greater than 2^x. Students can use either a table or a graph to justify their answers.

9. **a.** Both show exponential growth over a certain interval, CDs over the entire period and cassettes until about 1987.

STUDENT PAGE 108

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

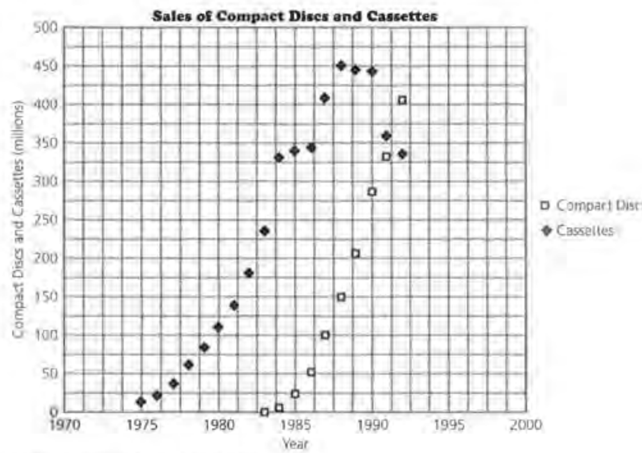
8. Consider the two functions $y = 2x$ and $y = 2^x$.

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.

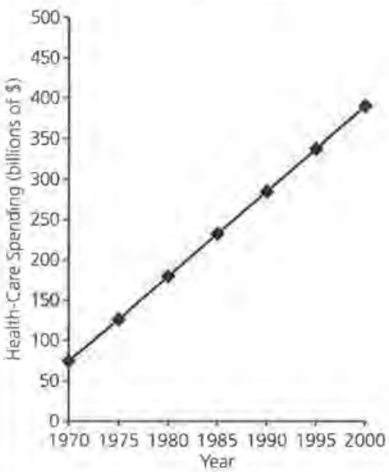
STUDENT PAGE 109

b. Sales of CDs appear to be exponential because the increase is slow at first and then greater and greater each year.

10. a. \$126.9 billion

b.

Year	Health-Care Spending (billions of \$)
1970	74.4
1975	126.9
1980	179.4
1985	231.9
1990	284.4
1995	336.9
2000	389.4



c. No; the distance between the constant-increase line and the actual line continues to increase. The actual and predicted spending increase more each year.

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, . . .).

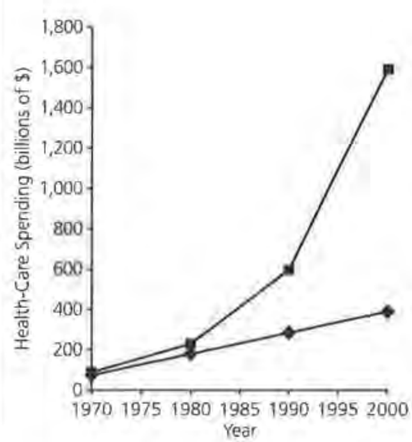
a. How is this rule different from $A = 0.0566(1.1x)$?

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."



11. a. In $A = 0.0566(1.1^x)$, the independent variable is an exponent. In $A = 0.0566(1.1x)$, the independent variable is a factor. As x increases, 1.1^x will get much greater than $1.1x$. Some students may think 1.1^x will remain close to 1.1 because the value 1.1 is so close to 1. Have them experiment with different values for x in order to see what happens. They might even recognize the pattern in powers of 1.1, 1.21, 1.331, ...

(11) **b.** Up until 1990, the value is too low; in 2000 it is too high. It may be an accurate predictor between 1990 and 2000.

Year	Prediction
1970	44.70
1980	115.94
1990	300.72
2000	799.98

c. Possible answer: Because the independent variable is an exponent, or because the difference between the predictions every ten years is not the same

12. Answers will vary.

13. Answers will vary. Some students might use a table such as the one below. Others might think about a 5% increase for a given amount and see whether it is exponential. Some students will think a 5% increase each year is a constant increase.

	Increase	% increase
$2^1 = 2$		
$2^2 = 4$	2	100%
$2^3 = 8$	4	100%
$2^4 = 16$	8	100%

	Increase	% increase
$3^1 = 3$		
$3^2 = 9$	6	200%
$3^3 = 27$	18	200%
$3^4 = 81$	54	200%

LESSON 12

Percents, Proportions, and Graphs

Materials: none
Technology: graphing calculators or spreadsheets
Pacing: 1 class period

Overview

The activities reinforce student work with percents and proportions. In the context of the increase in the elderly population of the United States, students investigate the trend and think about some of the consequences of this trend: the need to increase funding for Social Security, the need more health-care and medical services, an increase in housing and recreation facilities for the elderly, fewer people available to provide services, and so on. Students use inequalities to express relationships and to organize data in a table. They graph cumulative proportions for data according to age distributions and use these graphs to estimate measures of center. These graphs offer a different perspective on functions, as equations are not used to produce the values. The rules use inequalities and the tabular data. Students might consider whether it is appropriate to use a continuous line for the plots and whether they can interpolate for either x or y. Ask students how the functions represented in this lesson are similar to those from Lessons 9 and 10 and how they are different.

The lesson could be used in conjunction with a social-studies unit dealing with trends in the United States population.

Teaching Notes

Students often continue to have difficulty with percents and proportions. An important note: most mathematics books define a proportion as the equality of two ratios. In many real situations, the word is used differently; 0.2 is considered a proportion, designating 2 as a proportion of 10, 20 as a proportion of 100, and so on. It is not necessary to make an issue of this, as mathematical definitions are arbitrary; but be sure students understand what the word means in the given context.

The study of numbers and the way numbers are used to convey information should not end in elementary school. Interpreting data contained in a table is not always simple or obvious. The sophisticated application of number in a variety of contexts, comparing proportions from a base of 3 billion to a proportion from a base of 250,000, or looking at a cumulative percent, makes interpreting numbers complicated. Encourage students to read the numbers carefully, paying attention to the units. If necessary, they should write out some of the relationships. Sometimes making an estimate using rounded values will enable students to decide whether they are thinking correctly. Throughout the module, students have encountered the use of rate and percent as tools to enable them to compare relative frequencies and those ideas are revisited again in this lesson.

The concept of age distribution of a population is new to most students, and graphical representations of such situations may be unfamiliar. You many want to discuss some of the problems an aging population may cause, such as the need for nursing facilities, increased costs of medical assistance, increased market for jobs dealing with older citizens, and a decrease in number of people working to support the elderly.

Technology

Students might do some of the calculations by hand but should be allowed to use technology such as a spreadsheet or a graphing calculator to do most of the calculations.

Follow-Up

Students may bring articles on the population to class and discuss the form in which the information is presented. Are percents, rates, or proportions used? Would another form convey more or different information? You might have students investigate the change in population from 1990 to 1995 and the projected change to 2000.

Have students brainstorm situations where the cumulative proportion is important, for instance, proportion of scores below 90, proportion of incomes below the poverty line, and proportion of cars less than three years old.

STUDENT PAGE 110

LESSON 12

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

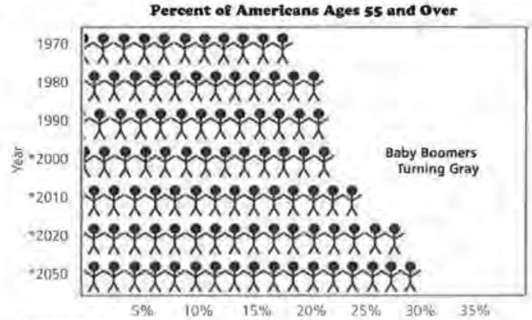
Work with inequalities.

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

Solution Key

Discussion and Practice

1. **a.** In the year 1900, $\frac{121}{1000}$, or 12.1% of people were under age 5.

b. It was 0.196 in 1900 and has fluctuated between that level and 0.135 since then. In 1992 it was 0.142 and is projected to continue to decrease.

c. It has risen rapidly since 1920 or so from 0.041 to 0.127 in 1992 and is projected to reach 0.130 in 2000.

2. **a.** Answers will vary. You must be careful because there is no way to tell the specific ages of the people within a group of ages. There is also no way to find the number of people who are a given age. Some schemes can be employed to make a plot similar to a number-line plot that does present the data in a useful way.

b. Not really; there is no way to tell if most people are in the younger half of their respective age groups or if they are in the older half, but either situation would affect the mean. A typical technique is to assume the people are equally distributed throughout the interval. Rough approximations can be obtained by using the mean age for each group or by calculating values for each end of the age interval.

STUDENT PAGE 111

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	0.067	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

3. **a.** 0.121 in 1900; 0.072 in 1980

b. 0.356 in 1900; 0.353 in 1980

c. 0.178 in 1900; 0.308 in 1980

d. Percent of those over 45 has risen, percent under 5 has fallen, and percent of those ages 15–35 has remained fairly level.

4. **a.** 1900; 344 per 1000 people were under age 15. This counts people from both the under-5 and the 5–14 age groups.

b. 0.447

c. Subtract the proportion under age 65 from 1.000.

d. Subtract the proportion under age 45 from the proportion under 25.

STUDENT PAGE 112

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?

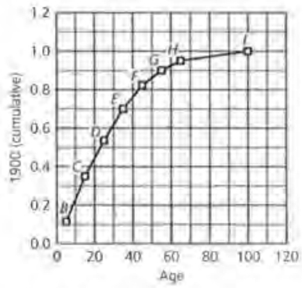
b. Find the entry that corresponds to $A < 25$ for 1960.

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?

STUDENT PAGE 113

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?

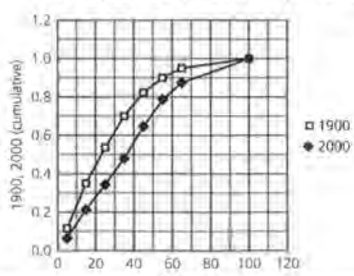
6. What age corresponds to

   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.



5. a. Approximately 0.7 of the population is under age 35.

   b. The older the age, the smaller the proportion that is added on. There are not many older people proportionally.

   c. The change between points B and C is more than 4 times as large as the change between points G and H.

   d. Answers will vary. In general, drawing straight segments between points tends to imply that the people within that age group are distributed evenly. It is reasonable to do this if you remember that it is only an estimate.

6. a. Approximately those younger than 13; this can be determined by finding the x value corresponding to a y value of 0.25.

   b. Approximately those younger than 23. this can be determined by finding the x value corresponding to a y value of 0.50.

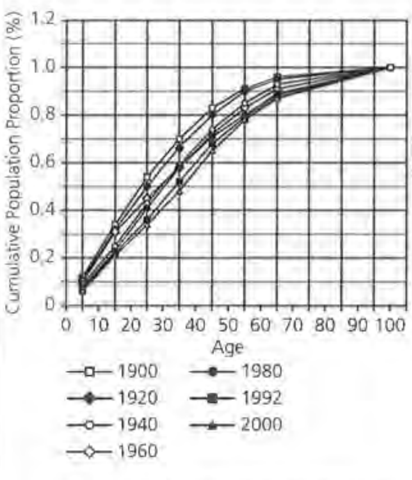
   c. The median age is about 23. This follows from part b, as 50% of the population is younger than 23; therefore, 50% of the population is at least as old as 23.

STUDENT PAGE 114

7. a. A smaller proportion of the population will be younger than 5 years old in 2000 than was younger than 5 years old in 1900.

b. Answers will vary. The proportion for a given age and younger is consistently predicted to be smaller in 2000 than it was in 1900, until age 100 (or very old). For example, in 1900 about 90% of the population was 55 or younger. In 2000 the prediction is for only 70% to be 55 or younger. 30% of the population will be older than 55, a large increase from the 10% in 1900.

8.



a. Answers will vary. In general, the plot for earlier years lies above plots for later years. This demonstrates the increasing proportions of older age groups.

b. No; the data are from a cumulative table, so every value is at least as large as the value in the cell directly above it.

c. Each cumulative proportion is the sum of all the cells above it in the noncumulative table. Since there are no negative values in the noncumulative table, the least a cumulative proportion can be is equal to the cumulative proportion above it. This can only happen when there is a value of 0 in the noncumulative table.

9. a. 23 in 1900; 25 in 1920; 29 in 1940; 29 in 1960; 30 in 1980; 33 in 1992; 36 in 2000. The vertical axis is the cumulative proportion. 0.5 will give you the fiftieth percentile, which is the median.

b. The median age has increased over time, especially since 1980.

10. a. Answers will vary. Students may use the midpoint of the interval and find the product of the midpoint and the proportion. Some students will not remember how to find the mean using the proportion, that is, find the sum of the products. They may need to see a worked-out example, for instance, using 100 people and figuring the number for a set of proportions and comparing the two ways to find the mean.

For the year 1992, using 65 to 100 as the top interval for an average

---

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.

of 82.5: Mean = $0.076 \cdot 2.5 + 0.143 \cdot 9.5 + 0.142 \cdot 19.5 + 0.166 \cdot 29.5 + 0.156 \cdot 39.5 + 0.107 \cdot 49.5 + 0.082 \cdot 59.5 + 0.127 \cdot 82.5 = 36.02$ years. Replacing the proportions with those for 2000, the projected mean will be 37.45 years.

(10) b. Answers will vary. The medians are lower in each case because the mean is skewed by the greater proportions of older people.

11. Answers will vary. Many students may conclude that there is at least *some* evidence to *suggest* problems of this sort. Be sure students use the data in their arguments, not just a verbal generalization.

Practice and Applications

12. a. 0.68

b. Portugal

c. Some of these countries have large populations and some have small populations. Having just the absolute numbers of people with different levels of education would make it difficult to compare relative education levels.

d. Answers will vary. The United States (0.36) and Canada (0.40) have the greatest proportion of those with a postsecondary degree. The United States has the least (0.17) with no high-school diploma, followed by Germany (0.18).

STUDENT PAGE 115

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.

STUDENT PAGE 116

13. a. 0.11 between 5 and 14; 0.15 older than 65

b. South Korea. Decision processes will vary. Some students may make a table of differences for the countries (0.19 to 0.03 for South Korea using those under age 5; 0.67 to 0.03 using those under age 15).

c. Answers will vary. The distribution of ages in the United States, Argentina, and Germany are not that different. South Korea is very different with many more younger people. You might have students check with a social-studies teacher for a possible reason.

14. Be sure that students realize that Table 12.5 is not cumulative as was Table 12.4.

a. 0.18

b. Answers will vary. Some students might suggest aerobic shoes, because the proportion of those in that age group (0.08 + 0.26 = 0.34) that purchased aerobic shoes is the greatest of all the shoe categories.

c. No; each column is a proportion based on the total number of shoes sold for that shoe category. If more shoes sell in one category than in another, then adding proportions in the categories has no meaning at all because the base numbers are different.

d. Answers will vary. In general, the younger age groups prefer gym shoes/sneakers; the older age groups prefer walking shoes; and the middle groups prefer jogging/running shoes and aerobic shoes.

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?

$$5 < A \le 14; \ A \ge 65$$

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
$14 \le A \le 17$ years	.04	.17	.09	.02
$18 \le A \le 24$ years	.08	.05	.11	.03
$25 \le A \le 34$ years	.26	.09	.19	.10
$35 \le A \le 44$ years	.25	.11	.21	.18
$45 \le A \le 64$ years	.22	.12	.25	.36
$A \ge 65$ 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.

STUDENT PAGE 117

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	8.74	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.

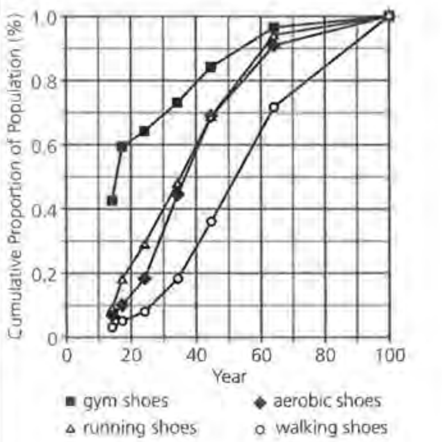
a. How do the two distributions of heights differ?

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.

15. a.

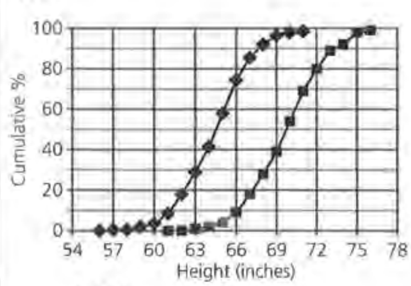


The plot, which begins at 0 and ends at 100, is slightly inaccurate, but it can be used for approximations.

Answers will vary. The proportion of walking shoes purchased by each age category is consistently lower than the other categories. The proportion of those who buy sneakers is greater until people are slightly over 60.

b. The median age is 15 for gym shoes/sneakers, 35 for jogging/running shoes, 36 for aerobic shoes, and 52 for walking shoes. Half of the people buying any category of shoes is younger than the median age; half of the people are at least as old as the median age.

16.



a. Answers will vary. Students might miss the fact that the table gives percents rather than proportions. This is a good place to reinforce the difference. A percent is always per 100, while a proportion can be part of any base, often written as a decimal where the base is a multiple of 10. Aside from the 5-inch displacement (males started in a 0.5-inch-taller category), the distributions are similar. Be sure students recognize the horizontal shift or transformation caused by the 0.5 inch from the male heights. The female differences from 63–65 inches are greater than the male differences from 68–70 inches, and the male differences between 71–73 inches are greater than the female differences from 66–68 inches.

b. About 65 inches for females and 70 inches for males; you can read the fiftieth percentile (0.5) from the plot.

c. Answers will vary. Approximately 64.45 inches for females and 69.42 inches for males. One method might be to find the exact proportion for each height, and then find the product of the heights and the proportion and use the average of that set of products as an estimate for the mean.

d. Answers will vary.

ASSESSMENT

Driving Records

Materials: graph paper
Technology: the same used in lessons
Pacing: 1/2 class period or as take-home assignment

Overview

This assessment includes concepts from both units of the module: expressions, formulas, and functions. Students make line plots, create formulas to calculate and interpret rates, find percents, make and interpret cumulative plots, and find the mean and median for data. They also have to use the data and their work to make decisions about driving records, that is, which age group has the worst driving record and whether males or females have a poorer record.

Teaching Notes

Many students feel that older people have worse driving records than younger ones. They can use this data to investigate such a claim. Some students will feel that they need information in addition to the crash data, such as the number of accidents caused by the age groups. This information may be difficult to obtain, but students might refer to their local police departments or to the state traffic bureau. The data sets in the assessment were obtained from *1994 Wisconsin Traffic Crash Facts,* a yearly publication from the Wisconsin Department of Transportation, Highway Safety Strategies/Analysis Section, Division of Planning, P.O. Box 7913, Madison, WI, 53707. Most states should have similar publications that can be obtained free by contacting the state department.

The problems in the first set of data, crashes according to age, are quite structured. Students are specifically asked to carry out investigations similar to those in the lessons. The question posed for the male/female data is open-ended, and students may approach a solution in a variety of ways. There is no correct approach, but you should be sure that students support their conclusions with graphs and numerical arguments. Suggest to students that this is their opportunity to show you what they have learned and how well they are able to apply the concepts from the module. You might choose to have students work in pairs or individually on the problems. Allow students several days in which to complete the assessment.

Some of the answers may be slightly different because students might choose to use drivers from the two younger categories, 14 and under and 15; or they may leave them out of the calculations. Do not be too concerned about which they choose to do, but concentrate on how students used the data to answer the questions.

Technology

Students should be allowed to use the same forms of technology they have been using in the lessons. If they do not have access to this technology, the work will be tedious, time-consuming, and susceptible to computational errors. One of the objectives of the entire module is to provide students with experiences that necessitate the use of spreadsheets and procedures for performing a repeated series of calculations. They should be able to demonstrate their proficiency in this respect in the way they carry out the assessment.

STUDENT PAGE 118

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

STUDENT PAGE 119

Solution Key

1. **a.** The 35–44 age group has the greatest percent of drivers, 22.4%, followed by the 25–34 age group with 21.7%. The 25–34 age group has the greatest percent of drivers involved in crashes, 22.4%, followed by the 35–44 age group with 18.1%. The 16-year-olds have the greatest percent of drivers in their age group in crashes, 21.0%, followed by the 17-year-olds with 15.6%.

b. Possible answer: There are more drivers in the 35–44 age group. By a slight difference, more of the crashes involve drivers in the 25–34 age group. The 16- and 17-year-olds, however, show the most crashes for their age groups; more than 1 in 5 of the 16-year-olds and 3 out of 20 17-year-olds are involved in a crash.

2. **a.** Possible answer: $R = \frac{TC}{D} \cdot 1000$, where R = the total crash rate per 1000 drivers, TC = the number of drivers in each age category in all of the crashes, and D = the total number of drivers in that age category.

Age	Total Crash Rate
16–19	157.29
20–24	101.20
25–34	72.45
35–44	56.59
45–54	47.81
55–64	41.17
65–74	34.70
75–84	36.85
85 +	38.65

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.

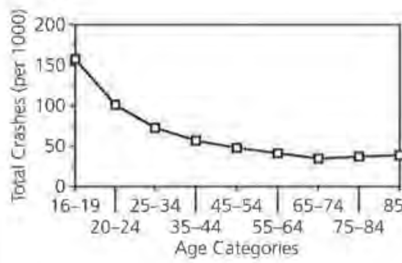
c. Estimate the mean age of drivers involved in crashes. Describe the procedure you used. How does your estimate for the mean age compare to that for the median age?

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.

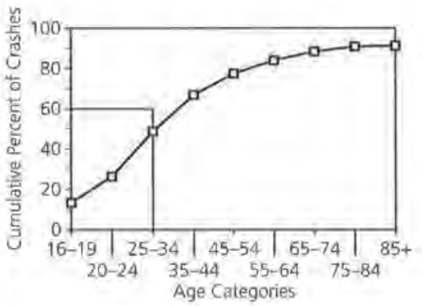
b.



The crash rate is much greater for drivers in the 16–19 age category, and the rate decreases as the drivers get older.

3. **a.** The cumulative percent can be found by dividing the cumulative sum of the drivers in the total number of crashes for each age by the cumulative sum of the total number of drivers for that age.

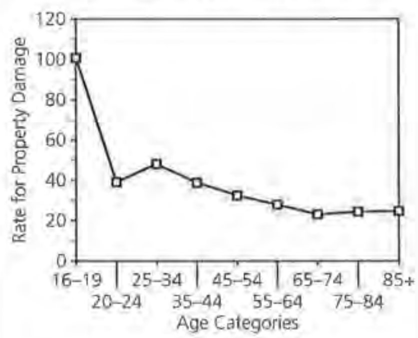
Age	Cumulative %
16–19	13.500
16–24	26.400
16–34	48.800
16–44	66.900
16–54	77.700
16–64	84.000
16–74	88.400
16–84	90.900
16–85+	91.300

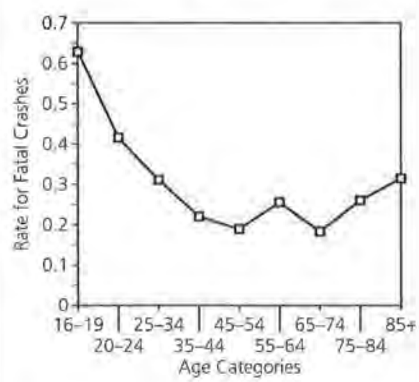


b. The median age is around 29 or 30. The cumulative plot does not reach 100% because there are a significant number of people involved in crashes whose ages were unknown. This will affect the estimate of the median age of drivers involved in crashes. Students can estimate the median age by drawing a line from the 50th percentile to the plot, and then finding the corresponding age on the horizontal axis.

c. Answers will vary. Students may estimate the mean age by using the midpoint of the intervals as a representative age for that interval, then taking the sum of the products of age and frequency. A possible answer could be 37 using this technique and counting 85+ as 89.5.

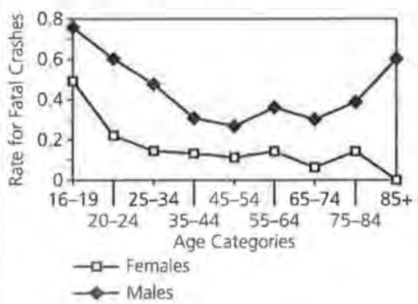
4. Students may choose different ways to approach the question. One way would be to calculate the rates for each and look at the two plots. The rates per 1000 drivers for the property-damage crashes are much higher than the rates for fatal crashes, but you can see from the plots that there seems to be a slight increase in the fatal-crash rate for older people. The property-damage-crash rate is the least and relatively constant for drivers in the 65-and-over age categories.





5. Answers will vary. Most students will find evidence that supports the fact that teenagers, particularly 16- and 17-year-olds, have the highest crash rate and should have the highest insurance premiums. Some students may choose to explore all of the different crashes, while others may be content to look at only several. Encourage students to investigate all of the information and to make graphs that will support their claims. If students offer only one statistic such as a percent as their evidence, give them only a small amount of credit. This is their opportunity to show you what they know and can do.

6. Answers will vary. Some students may suggest that males are involved in more crashes than females based on comparing rates and looking at the plots such as the one below. In every age group, the rate for males is higher than that for females. Other students may offer different evidence for their conclusions.



STUDENT PAGE 120

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.

Teacher Resources

Total Average

NAME ______________________________

The *Total Average* is another measure for rating a baseball hitter. The Total Average *TA* is the ratio of the *total bases* a player accumulates for his team to the *total outs* he costs his club. *Total bases* accumulated include all bases earned by hits (a single counts as 1 base, a double as 2, and so on), as well as walks, stolen bases, and the number of times the batter is hit by a pitched ball. If a player is thrown out stealing, one base is subtracted from the player's accumulated bases. *Total outs* accumulated include the number of times a player makes an out at bat *(Outs = At bats – Hits)* and the number of times thrown out stealing. A player is charged an additional out each time he hits into a double play.

(Source: *Inside Sports,* February 1996)

Edgar Martinez, the 1995 *TA* leader, had 101 singles, 52 doubles, 29 home runs, 116 walks, and 4 stolen bases and was hit by 8 pitches. He had 511 at-bats and got 182 hits. He was caught stealing 3 times and hit into a double play 11 times.

1. Martinez's *TA* for 1995 was 1.3. Show how his *TA* was calculated.

2. Remember that *TA* is the ratio of *Total bases* to *Total outs.*

a. Write a formula you could use to calculate the *TA* for any baseball or softball player.

b. Would it be possible for a player to have a *TA* of 1? Explain your answer.

3. The statistics for two more players are given in the table below. Notice that the number of Bases on Hits has already been computed for you.

a. Finish the table for each player.

Player	Bases on Hits	Walks	Hit by Pitch	Stolen Bases	Caught Stealing	Total Bases	At-bats	Hits	Outs	Double Plays	Caught Stealing	Total Outs
Frank Thomas	299	136	6	3	2		493	152	341	14		
Barry Bonds	292	120	5	31	10		506	149	357	12		

b. Use your formula to find the *TA* for the two players. (Assume a minimum of 3.1 plate appearances per game played.)

Writing Equations

NAME ______________________________

The following information was in the October 1995 *World Trade* magazine.

"Looking for an on ramp to the internet? Here are five national internet access providers that offer toll-free and/or local-number access."

- Alternet: Price \$20 per month plus \$3 per hour for local access, \$9 per hour for 800 access
- Cerfnet: Price \$20 per month plus \$10 per hour (\$8 per hour on weekends)
- Netcom Online Communications Services: Price \$19.50 per month plus \$20 sign-up fee
- Performance Systems International's InterRamp Service: Price \$116 (includes \$29 registration fee plus three months of unlimited usage at \$29 per month); local area network (LAN) access costs \$1.25 to \$6.50 per hour depending on modem speed and time of day plus a \$19 setup fee
- WinNet Communications: Price \$19.95 per month for four hours of toll-free access plus 9 cents a minute for each additional minute spent online

1. Write an equation that could represent the total cost to use each of the access providers. List any assumptions you made or additional information you would like.

2. According to a survey conducted by the Emerging Technologies Research Group in November and December, 1995, the average amount of time spent on the internet was 6.5 hours per week.

a. Which of the internet providers seems to be the most economical for a six-month trial if you are a typical user and you plan to spend one hour using the internet on Saturday and one hour on Sunday?

b. Which of the price packages seem suspicious or incomplete? Why?

c. Would your answer to Problem 2a change if you planned to be online only 15 minutes per day from Monday to Friday and a half hour on each of Saturday and Sunday? Show the work you did to make your decision.

LESSON 7 QUIZ

Summation, Standard Deviation, and Z-Scores

NAME ______________________________

1. The amount of money Tore earned each day over 14 working days was given by the following:

Day	Earnings E
1	19.50
2	26.00
3	13.00
4	32.50
5	19.50
6	45.50
7	65.00
8	13.00
9	32.50
10	32.50
11	13.00
12	52.00
13	45.50
14	19.50

a. Calculate $\sum_{i=4}^{8} E_i$

b. If $\sum_{i=1}^{15} E_i = \$461.50$, how much did Tore earn on the fifteenth day?

c. Tore's earnings were taxed at a 7.2% rate. Which of the following can he use to find the amount he has to pay in taxes?

$\sum_{i=1}^{14} 0.072\, E_i$ or $0.072 \sum_{i=1}^{14} E_i$

2. Consider Tore's earnings for the 14 days.

a. Find the average amount of money Tore earned each day for that time period.

b. Explain the difference between using the median and the mean as a measure of center.

(2) c. What was the standard deviation? How does this help you understand something about his earnings?

3. Over the same time period, Lissa had an average salary of \$25.75 per day with a standard deviation of 0.

a. Did Lissa make more or less than Tore on the tenth day? Explain how you know.

b. Explain what the following means in terms of Lissa's salary: $\sum_{i=1}^{14} 25.75$

4. The following table gives the birth rate and the life expectancy for a random set of countries in Africa. The mean for the birth rate of all of the African countries in 1995 is 43 births per 1000 people, and the standard deviation is 7.3. The mean life expectancy for all African countries is 52.6 years, and the standard deviation is 7.4 years.

Country	Birth Rate	Life Expectancy
Kenya	44	59
Mauritius	18	70
Uganda	51	42
Liberia	47	55
Zaire	47	52
Algeria	34	66
Egypt	32	62
Malawi	54	44
Ghana	42	56
Congo	45	52

Source: *The Universal Almanac,* 1996

Calculate the z-scores for the birth rate and the life expectancy in each country. Then use the z-scores to answer these questions.

a. Which countries are above the mean in both categories?

b. Which country is closest to the mean in both categories?

c. In which of the two categories does Kenya have the higher rating when compared to the other countries?

d. Which country seems to have the highest life expectancy and the highest birth rate compared to the other countries?

EXTENSION: QUIZ AFTER LESSON 1, 4, 5, 7, OR 8

More on Rating Quarterbacks

NAME ______________________________

In Lesson 1 you were given a formula to rate passing by the NFL quarterbacks. According to the NFL, there were four categories: percentage of touchdown passes per attempt, percentage of completions per attempt, percentage of interceptions per attempt, average yards gained per attempt. The average standard is 1.000. To earn a 2.000 rating, a passer must perform at exceptional levels which they define as 70% in completions, 10% in touchdowns, 1.5% in interceptions, and 11 yards average gain per pass attempt.

The passing records of the NFL quarterbacks in the American Federation Conference (AFC) for the 1994 season are given in the table below.

Table E.1: Passing Records of NFL Quarterbacks

AFC Quarterbacks	Attempts	Completions	Yards Gained	TDs	Interceptions
Jeff Blake	306	156	2154	14	9
Drew Bledsoe	691	400	4555	25	27
Boomer Esiason	440	255	2782	17	13
John Elway	494	307	3490	16	10
Jeff Hostetler	454	263	3334	20	16
Stan Humphries	453	264	3209	17	12
Jim Kelly	448	285	3114	22	11
Dan Marino	615	385	4453	30	17
Joe Montana	493	299	3283	16	9
Neil O'Donnell	370	212	2443	13	9

Source: *Sporting Almanac* 1995

1. Do *any* of the quarterbacks seem to be exceptional in any of the four categories? Explain your answer.

2. Review the method outlined below that was provided by the Minnesota Vikings. Is this formula different from the one given in Lesson 1? Explain why or why not.

I. Determine each of the four categories to as many decimal places as possible:

fraction of passes completed per attempt;
average per attempt;
touchdowns per attempt;
interceptions per attempt.

II. Subtract 0.3 from the *fraction of passes completed per attempt* and 3.0 from *average yardage per attempt.*

III. Multiply each of the numbers except *average yardage per attempt* by 100.

IV. Take the numbers from step III and multiply each by the values indicated:

completion per attempt by 0.05;
touchdown per attempt by 0.2;
average yardage per attempt by 0.25;
interceptions per attempt by 0.25

V. Subtract the *interceptions per attempt* value from 2.375.

VI. Add all four categories together.

VII. Divide the total reached in step VI by 6.

VIII. Multiply the total from step VII by 100 and round to the nearest whole number.

Exploring Symbols: An Introduction to Expressions and Functions

NAME ______________________________

Part I

1. In 1995 a Chicago law firm, Small Business Advocate, charged some clients an annual fee of \$575 plus a percent of the firm's hourly billing rate of \$135.

Source: *Inc*, November 1995, p. 99

a. Identify the variables involved in determining how much a lawyer from that firm would have charged you in 1995.

b. If you had to pay \$67.50 per hour, describe the discount you received.

2. **a.** For most legal services, clients pay an annual fee and 20% of the hourly billing rate. Write an equation that would help you determine the cost of a lawyer.

b. How much would you have paid for 10 hours of work?

c. How many hours of work could you expect if you only had \$1,000 to spend at the 20% rate.

3. Another law firm charged \$250 per hour.

a. How are the variables used to find the cost for this firm different from the variables in determining cost for the first firm, Small Business Advocate?

b. Would it ever have been more economical to use this firm rather than Small Business Advocate? Explain your answer.

4. Suppose another firm charged fees according to the following formula:
Fee = $\$150H + C + 50\ L \cdot H$ where H is the number of hours of work, C is an additional charge determined by the category of legal help (for example, divorce case, traffic offense, felony), and L is a class fee charged by lawyers at different levels in the firm.

a. If a top-rated lawyer is at level 4, how much more per hour does that lawyer charge?

(4) b. Write an expression for the hourly fee you would expect to pay.

c. Which would be less expensive: to have a lawyer from the highest level (4) for three hours or to have a lawyer from the lowest level (1) for six hours? Is your answer dependent on the category of legal help? Explain.

Part II

5. The table below gives the number of AIDS cases for ten countries and the population of each of the countries. The total number of reported AIDS cases in 1995 in the world was 1,169,811. In which country does the AIDS epidemic seem to be the worst? Explain how you made your choice.

Table T.1: AIDS Cases by Countries

Country	Cases	Population
United States	441,528	260,713,585
Brazil	62,314	158,739,257
Kenya	56,573	28,240,658
Uganda	46,120	19,121,924
Tanzania	45,968	27,985,660
Zimbabwe	38,552	10,975,078
Malawi	37,673	9,732,409
France	35,773	57,840,445
Spain	31,221	39,302,665
Zambia	29,734	9,188,190

Source: *The Universal Almanac*, 1996

Part III

6. The following table contains data collected on the number of hours studied by students in a typical week for students with jobs and for students without jobs.

Table T.2: Number of Hours Studied

	No. Hours	Mean	Median	St. Dev.	Min.	Max.	Q_1	Q_3
With Jobs	25	11.22	10	5.28.	3	21	6.5	15.0
Without Jobs	37	10.75	10	6.78	3	40	5.0	14.5

a. Write a paragraph comparing the number of hours spent studying for students with and without jobs. Use graphs and statistics to illustrate your discussion.

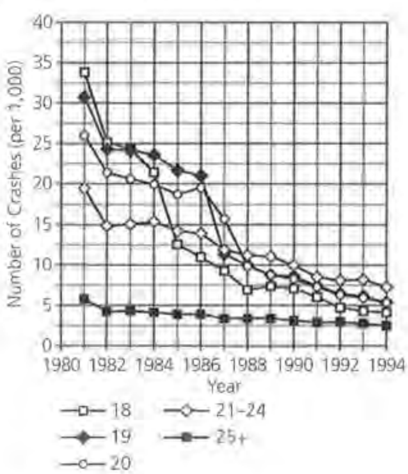
b. Find the mean number of hours spent studying for all 62 of the students. Explain your answer.

7. Shana studied for two tests. On the first she received an 81; the class mean was 75 and the standard deviation was 4.3. On the second test she received a 52; the class mean was 40 and the standard deviation was 7.5. On which test did she do better compared to the other scores?

Part IV

The data on driver crash rates as of 1994 from the state of Wisconsin are plotted below.

8.



Crash Rates by Age per 1,000 Licensed Drivers—14 Year Summary

Age	18	19	20	21–24	25+
No. licensed drivers	56,965	56,802	59,116	257,773	3,028,698
1994 crash rate	4.09	5.21	5.26	7.25	2.43

Source: Wisconsin Traffic Crash Facts, 1994

a. What was the crash rate for drivers between ages 21 and 25 in 1994? Explain what this indicates in terms of the number of licensed drivers.

b. What was the crash rate for drivers between the ages of 18 and 25?

9. Refer to the table on crash rates.

a. For which age group was the decrease in the crash rate the greatest? Explain your answer.

b. The drinking age was raised to 19 on July 1, 1984, and to 21 on September 1, 1986. Draw a vertical line to mark these spots on the graph. Is there any relation between these dates and the plot? Explain.

c. Write at least four observations about the crash rate by age over the 14-year period.

Part V

The data in the table contains the cumulative proportion of licensed drivers (in millions) by age for Wisconsin in 1994.

Table T.4: Proportion of Licensed Drivers

Age	Licensed Drivers (millions)	Cumulative Proportion
16–18	1.4	0.01
19–23	7.3	0.05
24–28	16.7	0.15
29–33	19.5	0.27
34–38	20.8	0.39
39–43	19.5	0.50
44–48	17.5	0.61
49–53	13.7	0.69
54–58	10.8	0.75
59–63	9.5	0.81
64–68	9.2	0.86
69 and over	8.5	0.91

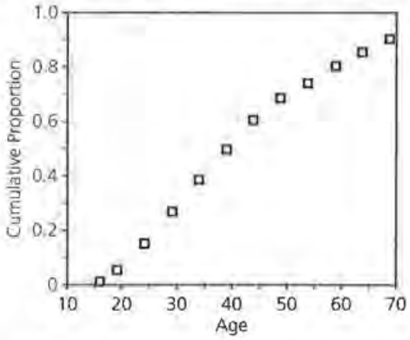
Source: *Wisconsin Traffic Crash Facts*, 1994

10. Use table T.4.

a. What proportion of the licensed drivers are from 24 to 28 years old?

b. If A_i is the number of licensed drivers for a given age i, find $\sum_{i=16}^{28} A_i$.

11. Study the plot below of the cumulative proportion.



a. Which age group(s) has the smallest proportion of licensed drivers? How can you find this answer from the graph?

b. Describe how the proportion of licensed drivers changes with age.

c. Approximate the median age of licensed drivers. Show how you made your approximation.

d. Approximate the mean age of licensed drivers. Show how you made your approximation.

Part VI

The table contains data about the number of motor-vehicle miles of travel and the driver fatality rate in Wisconsin. The fatality rate is the number of motor-vehicle deaths per 100 million miles traveled.

Table T.5: Fatality Rates

Year	U.S. Fatality Rate	Wisconsin Fatality Rate	Motor-Vehicle Miles of Travel (millions)
1980	3.3	3.16	31,165
1981	3.2	2.76	33,611
1982	2.8	2.36	32,795
1983	2.6	2.16	34,107
1984	2.6	2.35	35,456
1985	2.5	2.05	36,680
1986	2.5	1.97	38,428
1987	2.4	2.03	40,194
1988	2.3	1.92	42,339
1989	2.2	1.90	43,087
1990	2.1	1.72	44,276
1991	1.9	1.75	45,456
1992	1.8	1.36	47,495
1993	1.7	1.44	48,805
1994	1.7	1.40	50,253

Source: *Wisconsin Traffic Crash Facts*, 1994

12. Refer to Table T.5.

a. Would the number of motor-vehicle deaths in 1990 necessarily have to be greater than the number of motor-vehicle deaths in 1994? Explain why or why not.

b. Write a formula to determine the number of motor-vehicle deaths in any given year.

c. Use your formula to determine in which year there were the greatest number of deaths.

13. Make a plot that will enable you to compare the fatality rate for the United States to that for the state of Wisconsin over the 15-year period. Describe how the rates are alike and how they are different.

14. Consider the plot of the year and the number of motor-vehicle miles (in millions) driven.

a. Based on the data and the plot, how would you describe the change in the number of miles driven per year?

b. Describe how the graph will differ if the number of miles driven turns out to be an exponential function of time rather than a linear function.

c. Predict the number of miles you think will be driven in the year 2000. Explain why you think your prediction might be reasonable.

Total Average

Students find a formula from a verbal description and use the formula to determine the total average.

This quiz is provided as an option for students who are interested in a sports situation. The quiz should take about $\frac{1}{2}$ class period. It might also be given as a take-home exercise or for extra credit.

1. *Total bases:*

TB = singles + 2 • doubles + 3 • triples + 4 • home runs + walks + hits by pitch + stolen bases – caught stealing
= 101 + 2 • 52 + 3 • 0 + 4 • 29 + 116 + 8 + 4 – 3 = 446

Total outs: (where *outs* are *at-bats* minus *hits*)

TO = outs + hits into double plays + caught stealing
= (511 – 182) + 11 + 3 = 343

$TA = \frac{TB}{TO} = \frac{446}{343} = 1.3$

2. **a.** Possible formulas:

TA = (singles + 2 • doubles + 3 • triples + 4 • home runs + walks + hits by pitch + stolen bases – times caught stealing)÷ (outs + hits into double play + caught stealing), where outs are at-bats minus hits.

TA = $\frac{TB}{TO}$, where TB and TO are defined as in Problem 1.

b. A player could have a TA of 1 if Total bases were equal to Total outs. It seems reasonable that this could occur.

3. **a.**

Player	Bases on Hits	Walks	Hit by Pitch	Stolen Bases	Caught Stealing	Total Bases	At-bats	Hits	Outs	Double Plays	Caught Stealing	Total Outs
Frank Thomas	299	136	6	3	2	446	493	152	341	14	2	357
Barry Bonds	292	120	5	31	10	438	506	149	357	12	10	379

b. TA for Frank Thomas: 1.238; TA for Barry Bonds: 1.156

Writing Equations

Students write symbolic formulas from verbal descriptions for the fee structures of several different internet companies and then compare the results.

This quiz will enable you to see whether students can translate a formula from words into symbols and evaluate that formula for given values. Students will also have to make choices among different formulas and make assumptions when setting them up. Be sure students indicate what the symbols represent and use the correct units.

Allow students to use a graphing calculator or a computer to do their work. They should be able to use the spreadsheet function or the repeat key to perform the calculations several times without typing all of the values over again. If some students do not understand how the calculator can help them, have students who do understand demonstrate different procedures for the class.

The information is not explicit for several of the companies, and one reason may be the confusion about long-distance versus local-access calls. Netcom seems particularly inexpensive and not dependent on the amount of time spent per day or week. This is so different from the other companies that it seems suspicious.

As extra credit or an extension, students can experiment with different amounts of time to see whether it makes a difference in choosing a company.

The quiz should take 20 to 30 minutes to complete. It can be given as a quiz or assigned as an assignment, collected the next day and scored.

1. *Alternet:* $20M + (3h)W = C$ if local, where C is total cost, M is the months, h is hours per week, and W is the number of weeks in M months; $20M + (9h)W = C$ if using an 800 number (and the variables represent the same quantities as above)

Cerfnet: $20M + (10h_{WD})W + (8h_{WE})W = C$, where M is the number of months, h_{WD} is the number of weekday hours, h_{WE} is the number of weekend hours, and W is the number of weeks in M months

Netcom: $19.50M + 20 = C$, where M is the number of months

Performance Systems:
$116 + 29(M - 3) + 19 + 1.25hW \leq C$ and
$C \leq 116 + 29(M - 3) + 19 + 6.50hW$, where M is the number of months and $M > 3$, h is the number of hours per week, and W is the number of weeks in M months

WinNet: $19.95M + 0.09(mW - 240M)$, where M is the number of months, m is the number of minutes per week, and W is the number of weeks in M months

2. **a.** *Alternet:* Cost will be between
$20(6) + 3(6.5)(26) = \$627$ and
$20(6) + 9(6.5)(26) = \$1,641$.

Cerfnet: $20(6) + 10(4.5)(26) + 8(2)(26) = \$1,706$

Netcom: $19.50(6) + 20 = \$137$

Performance Systems: Cost will be between
$116 + 29(3) + 19 + 1.25(6.5)(26) = \$433.25$ and
$116 + 29(3) + 19 + 6.50(6.5)(26) = \$1,320.50$.

WinNet: $19.95(6) + 0.09[(6.5)(60)(26) - 240(6)] = \$902.70$

Possible answer: The most economical internet access company seems to be Netcom. However, the cost is so much lower than the others that it seems that this might be for local access rather than for long distance or 800 numbers.

b. Possible answer: The price for Netcom seems unreasonable. It does not indicate whether it is for an 800 number or for local access. The price is much lower than for the other access providers. In general, the prices vary so much that it seems necessary to find other information about each of the companies and their offers.

c. If you were online for 15 minutes each weekday and for 30 minutes each Saturday and Sunday, Netcom at \$137 would still be the cheapest.
Alternet: $\$295.50 \leq C \leq \$646.50$
Cerfnet: \$653
Performance Systems: $\$295.13 \leq C \leq 602.25$
WinNet: \$306

LESSON 7 QUIZ: SOLUTION KEY

Summation, Standard Deviation, and Z-Scores

This quiz covers the use of the summation symbol and some of its properties, and finding the mean and the standard deviation and interpreting them in a given situation. Students also have to find the z-scores for two sets of data and use the results to answer several questions. The focus is on understanding the symbolism and on the ability to do the calculations either by hand or with a calculator.

The quiz can be used after Lesson 7. Problem 1 is covered in Lesson 5, Problems 2 and 3 in Lesson 6, and Problem 4 in Lesson 7. You may choose to give it in two parts, with Problems 1–3 following Lesson 6 and Problem 4 after Lesson 7. Ask students to find the z-scores and interpret the results. It is important to make sure that students can use their calculators to find the statistics and to catch misconceptions about notation and vocabulary before they become part of the student's operating knowledge.

1. a. \$175.50

b. $\sum_{i=1}^{15} E_i = \sum_{i=1}^{14} E_i + \sum_{i=1}^{1} E_i$,

so $\$429 + \sum_{i=1}^{1} E_i = \$461.50$

and $\sum_{i=1}^{1} E_i = \$32.50$.

c. Both of the formulas will give the correct amount of tax. The first finds the tax for each day and adds the individual taxes. The second finds the total amount of earnings and finds the tax on that. In either case, the tax is \$32.175.

2. a. About \$30.64 per day

b. Answers will vary. Using the mean involves knowing exactly how much was involved, the weight of each piece of data or day's salary. The median indicates only where the middle number occurs and does not yield much information about the values themselves. The data can be in any order to calculate the mean, but must be ordered according to size to find the median.

c. The standard deviation was \$15.62, which indicates a considerable variation.

3. a. Lissa earned \$25.75 each day, because her standard deviation was zero and her average salary was \$25.75. On the tenth day, Tore earned \$32.50 so Lissa earned less.

b. Lissa's earnings of \$25.75 were constant for 14 days.

4.

Country	Birth Rate	Life Expectancy	Birth-Rate z-score	Life Expectancy z-score
Kenya	44	59	0.137	0.865
Mauritius	18	70	–3.425	2.351
Uganda	51	42	1.096	–1.432
Liberia	47	55	0.324	0.548
Zaire	47	52	0.548	–0.081
Algeria	34	66	–1.233	1.181
Egypt	32	62	–1.507	1.270
Malawi	54	44	1.507	–1.162
Ghana	42	56	–0.137	0.459
Congo	45	52	0.274	–0.081

a. The countries that are above the mean in both categories are Kenya and Liberia. Students can answer this either from looking at the two means and the data, or by using the z-scores for the two categories.

b. The countries that are closest to the mean in both categories are those whose z-scores are close to 0. Congo seems to be the closest, although it has a higher z-score than Ghana and Kenya for birth rate and it is much closer to the mean than either in terms of life expectancy, 0.081 as opposed to 0.865 and 0.459.

(4) c. Kenya's rating is higher in terms of life expectancy when compared to the African mean with a z-score of 0.865. Its birth rate is closer to the mean.

d. Answers will vary. Students could indicate Liberia has the greatest positive z-score for both categories, 0.324 and 0.548; but others may argue that Kenya has a higher life-expectancy z-score, 0.865, and by adding the two, a higher total for the two z-scores. Both positions have merit, particularly since students are using z-scores to help in their arguments.

EXTENSION QUIZ: SOLUTION KEY

More on Rating Quarterbacks

In this activity, students revisit the rating of quarterbacks, first inspecting data on a set of quarterbacks to see if there are any extreme values. They then work through a new formula for rating quarterbacks; one that is given as a step-by-step procedure and that, at first glance, does not seem to have any relation to the one presented in Lesson 1. After some arithmetic and symbolic manipulation, the two turn out to be equivalent.

The Extension may be assigned after Lesson 1; after Lesson 4 or Lesson 5 when students have had more experience working with variables and formulas; after Lesson 7 when they have studied averages and variation; or after Lesson 8 when they have studied z-scores. It is probably best given as a homework assignment, perhaps covering several days.

If students do the Extension after Lesson 7, they should use their knowledge of mean and standard deviation or of median and interquartile ranges to determine how "different" the data are in each category. If they have studied z-scores, they should use them in their analysis.

Comparing the formulas is a good place to introduce students to the concept of a mathematical proof. Some will just try both formulas to see if they give the same result. Stress that this is not a proof. The manipulation seems complicated, but it is not as problematic as it seems and provides a nice proof that the two approaches are equivalent. Essentially, the proof depends on an accurate translation, the distributive property and combining constants. Multiplying the numerator and denominator by 4 is prompted by knowing what the final result should look like from the rule given in Lesson 1.

1. Answers will vary. If they simply inspect the numbers, students might draw the following conclusions: Drew Bledsoe is way ahead of the others in attempts, completions, and interceptions (by 76 attempts over Marino, 115 completions over Marino, and 10 interceptions over Marino and Kelly). He and Dan Marino are ahead of the others in yards gained. Jeff Blake is low in attempts (64 below O'Donnell) and completions (56 below O'Donnell) but not too far behind the others in yards gained.

The computed values for measures of center and variability are shown in the first table on page 178. Using measures of center and variability, students might recognize that there was a relatively large standard deviation in the attempts compared to the IQR which indicates there is probably an outlier. The upper limit for an outlier is 602. Both Marino and Bledsoe would be considered outliers.

If the Extension is assigned after Lesson 8, students may use z-scores to answer the question. In that case, a z-score table would give the values listed in the second table on page 178.

Here, it is apparent that Drew Bledsoe's 27 interceptions are quite different with a z-score of 2.447; Dan Marino's 30 touchdowns and Bledsoe's attempts are also quite different from the others with z-scores of 2.071 and 1.947. These are the greatest z-scores in any category. This is a good place to point out the difficulty of dealing with relative differences: a difference of 70 or 100 in the attempts and completions is not as significant as a difference of 10 in interceptions that have lower numbers. This illustrates again the importance of having a standard way to think about the variability in data that are presented in different scales.

AFC Quarterbacks	Attempts	Completions	Yards Gained	TDs	Interceptions
Mean	476	282	3282	19	13.3
Standard Dev.	110	73	767	5.3	5.6
Median	453	274.5	3246	17	11.5
IQR	(494, 440) = 54	(307, 255) = 52	(3490, 2782) = 708	(22, 16) = 6	(16, 9) = 7

AFC Quarterbacks	Attempts	Completions	Yards Gained	TDs	Inter- ceptions
Jeff Blake	−1.546	−1.740	−1.470	−0.941	−0.768
Drew Bledsoe	1.947	1.614	1.660	1.130	2.447
Boomer Esiason	−0.330	−0.379	−0.651	−0.377	−0.054
John Elway	0.160	0.335	0.272	−0.565	−0.589
Jeff Hostetler	−0.203	−0.269	0.068	0.188	0.482
Stan Humphries	−0.212	−0.256	−0.095	−0.377	−0.232
Jim Kelly	−0.258	0.033	−0.219	0.565	−0.411
Dan Marino	1.257	1.407	1.527	2.071	0.661
Joe Montana	0.151	0.225	0.002	−0.565	−0.768
Neil O'Donnell	−0.965	−0.970	−1.093	−1.129	−0.768

2. Answers will vary. The ratings formulas are the same. Using the steps given, we get:

$$R = (100)\cdot \frac{100(.05)(\frac{C}{A} - .3) + .25(\frac{Y}{A} - 3) + 100(.2)(\frac{TD}{A}) + [2.375 - 100(.25)(\frac{I}{A})]}{6}$$

$$= (100)\,\frac{5\frac{C}{A} - 1.5 + .25\frac{Y}{A} - .75 + 20\frac{TD}{A} + 2.375 - 25\frac{I}{A}}{6}$$

$$= \frac{500\frac{C}{A} - 150 + 25\frac{Y}{A} - 75 + 2{,}000\frac{TD}{A} + 237.5 - 2{,}500\frac{I}{A}}{6}$$

$$= \frac{500\frac{C}{A} + 25\frac{Y}{A} + 2{,}000\frac{TD}{A} + 12.5 - 2{,}500\frac{I}{A}}{6}$$

$$= \frac{4}{4}\,\frac{12.5 + 500\frac{C}{A} + 25\frac{Y}{A} + 2{,}000\frac{TD}{A} - 2{,}500\frac{I}{A}}{6}$$

$$= \frac{50 + 2{,}000\frac{C}{A} + 100\frac{Y}{A} + 8{,}000\frac{TD}{A} - 10{,}000\frac{I}{A}}{24}$$

This is the formula given in Lesson 1.

The resulting ratings with either formula are shown in the table at the right.

AFC Quarterbacks	Rating	Rank
Jeff Blake	76.9	9
Drew Bledsoe	73.6	10
Boomer Esiason	77.3	8
John Elway	85.7	2
Jeff Hostetler	80.8	6
Stan Humphries	81.6	5
Jim Kelly	84.6	3
Dan Marino	89.2	1
Joe Montana	83.6	4
Neil O'Donnell	78.9	7

Exploring Symbols: Introduction to Expressions and Functions

This test covers concepts introduced in many of the lessons but is not all-inclusive. Students are asked to demonstrate what they know about writing formulas given a verbal description and using and interpreting formulas given in symbols. They are expected to take statistical summaries of several data sets, construct graphs from the data, and make comparisons. Students have to understand rates in order to compare plots over time for different age categories and interpret cumulative proportions given in a table and a plot. They also have to demonstrate that they understand to some degree the difference between the plots of exponential and linear relationships.

The test can be given in class or as a take-home test. Most of the plots are provided and the statistical summaries have, in many cases, already been calculated so that students can focus on interpretations rather than on the calculations. They do have to calculate the number of deaths predicted for a number of years and make a plot. To facilitate the process, you may want to allow them to link the data for Problem 11 before the test begins.

1. **a.** Number of hours billed; the percent of the \$135 billing rate

b. 50%

2. **a.** Possible answers: $C = 575 + 0.2(135)H$ or $C = 575 + 27H$, where H is the number of hours worked and C is the total cost

b. $27(10) + 575 = \$845$

c. 15 hours; $1000 = 575 + 27H$, $H \approx 15.74$

3. **a.** There is only one variable here, the number of hours. For the first firm, both the rate and the number of hours varied.

b. The first firm is cheaper for one or two hours. Student approaches will differ. Some may use a table and notice that the fee at the second firm is greater than that at Small Business Advocates for 3 hours. The assumption here is that any part of an hour is charged at the full-hour rate.

(3b)

Hours	Small Business	Second Firm
1	602	250
2	629	500
3	656	750

4. **a.** \$200 more per hour

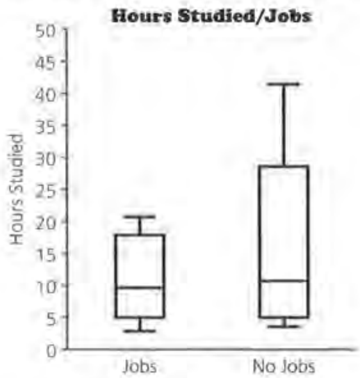
b. The hourly fee is $150 + 50L$, where L is the class of the lawyer.

c. The top-rated lawyer would charge \$1050 + C for 3 hours of work, while the lowest-rated lawyer would charge \$1200 + C for 6 hours of work. It would be cheaper to hire the top-rated lawyer for 3 hours. The answer does not depend on the type of service because the charge is the same in both cases.

5. Answers will vary. Some students may indicate that the United States has the greatest percent of the cases with 38%, while others may investigate the rate per unit of population. The table below contains the rate per 1,000 people. In this case, Malawi has the highest rate per 1,000 people in the country, followed by Zimbabwe. The United States is sixth when the total population is considered.

Nation	AIDS Cases	Population	Percent of Total	Rate per 1,000
U.S.	441,528	260,713,585	37.74	1.7
Brazil	62,314	158,739,257	5.33	0.4
Kenya	56,573	28,240,658	4.84	2.0
Uganda	46,120	19,121,924	3.94	2.4
Tanzania	45,968	27,985,660	3.93	1.6
Zimbabwe	38,552	10,975,078	3.30	3.5
Malawi	37,673	9,732,409	3.22	3.9
France	35,773	57,840,445	3.06	0.6
Spain	31,221	39,302,665	2.67	0.8
Zambia	29,734	9,188,190	2.54	3.2

6. a. Answers will vary. Half of the students in each category study about the same amount of time, with 3 hours per week as the minimum and 10 hours as the median. Half of the students with and without jobs study less than 10 hours per week. There is much more variability in the number of hours that students without jobs study, and most of the difference occurs for those without jobs who study more than 10 hours per week; some study as much as 40 hours, while the maximum for those with jobs is 21 hours per week. The standard deviation reflects this difference; it is 1.5 times as great as for those without jobs. A box plot can be constructed from the given data which makes the comments about variability more obvious.



b. Be sure students do not find the mean of the two means. The unequal number of students in the sets would make this value lower than the actual mean. The mean can be found by finding the product of the mean for those with jobs, 11.22, and the number of students, 25, and the product of the mean for those without jobs, 10.75, and the number of students, 37. The sum of these two values can be divided by the total number of people, 62, to give the overall mean.

$$\frac{11.22(25) + 10.75(37)}{62} \approx 10.94$$

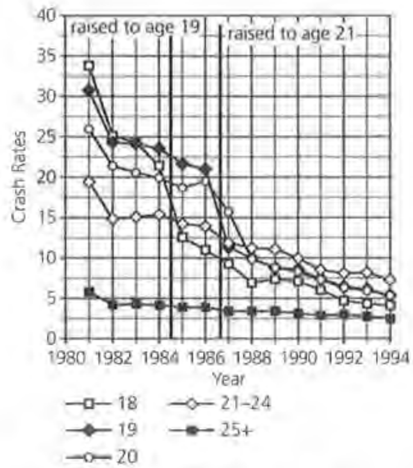
7. The second test; using z-scores, she was 1.4 standard deviations above the mean on the first test and 1.6 standard deviations above the mean on the second.

8. a. With a rate of 7.25 per 1,000 drivers, the number of crashes is $7.25\left(\frac{257,773}{1,000}\right) \approx 1869$.

b. The number of crashes for the four age groups is $233 + 296 + 311 + 1,869 = 2,709$. The combined rate is then $1,000\left(\frac{2,709}{430,656}\right) \approx 6.29$.

9. a. 18-year-olds; their rate fell from nearly 34 per 1,000 to about 7.5 per 1,000 between 1981 and 1988; this is the greatest drop in the plot for that time span.

b.



Possible answer: In 1984, the crash rate fell from about 22 to 13 per 1,000 for those aged 18 and then continued to fall. There was not that much change for the other ages until 1986, when there was a sharp drop for those aged 19 and 20.

c. Answers will vary. Be sure students use the numerical values as well as words when they describe the plot. Students might indicate that the crash rate for those over the age of 25 has changed very little over the years, although there has been a slight drop. The crash rate for those aged 21–24 dropped less dramatically from 15 to about 7.5 per 1,000; for those aged 18, the drop was from 34 to 4 per 1,000. This age group now has the highest crash rate, around 12.5 per 1,000. The other three age groups dropped significantly, particularly in the years 1984 and 1986 when there was a change in the drinking age.

10. a. 10% are from 18–28 years old.

b. 25.4 million drivers

11. a. The least proportion of drivers exists in the groups younger than 19 or older than 54, where it is around 5%.

b. The proportion of drivers is small, about 5%, for those younger than 19. It increases to about 10% or 11% for the intermediate age groups and then decreases for those older than 54 where it is again about 5%.

c. About 38; students can draw a horizontal line from 0.50 to the plot and read the corresponding x-value.

d. Answers will vary. Using the median of each interval and finding the sum of the products of the midpoint and the proportion for each age interval gives 39.9 or 40 years as the mean age of licensed drivers.

12. a. Possible answer: No; the number of deaths is a function of both the number of miles traveled and the death rate. You can have a large death rate and a small number of miles traveled and not have very many deaths. You can also have a large death rate and a large number of miles traveled, so there would be many deaths. In this case, however, the number of deaths in 1990 was 761.55 and the number of deaths in 1994 was 703.54

b. Possible answer: $R \cdot \frac{M}{100} = D$, where M is the number in millions of motor-vehicle miles driven, R is the Wisconsin fatality rate, and D is the number of deaths.

Year	Wisconsin Fatality Rate R	Motor-Vehicle Miles M (millions)	Deaths D
1980	3.16	31,165	984.814
1981	2.76	33,611	927.664
1982	2.36	32,795	773.962
1983	2.16	34,107	736.711
1984	2.35	35,456	833.216
1985	2.05	36,680	751.940
1986	1.97	38,428	757.032
1987	2.03	40,194	815.938
1988	1.92	42,339	812.909
1989	1.90	43,087	818.653
1990	1.72	44,276	761.547
1991	1.75	45,456	795.480
1992	1.36	47,495	645.932
1993	1.44	48,805	702.792
1994	1.40	50,253	703.542

c. 1980

13. Answers will vary. Some students may make a scatter plot as indicated below, but others may use box plots or back-to-back stem-and-leaf plots. Both types of plots, however, will not show the comparison over time.

Both rates decreased. The Wisconsin rate was consistently lower than the United States rate, but it decreased and then increased several times, while the U.S. rate either decreased or stayed almost the same.

14. a. Answers will vary. Some students might use the plot and generalize from the linear pattern they observe. Others might make a statement based on analyzing the difference in the number of miles driven each year in the table. Somewhere around 1,400 million more miles driven each year would be reasonable, but check student work. Be sure they remember to include millions as the unit.

b. Possible answer: If the trend continues to be linear, the number of miles driven each year will continue to increase at about the same rate, 1,400 million miles. If the trend is exponential, the rate will increase and continue to increase more each year. In terms of the plot, if the relationship is linear, the plot will continue to look like a straight line. If the relationship is exponential, the plot will begin to curve up rapidly.

c. Answers will vary. If students believe the trend will remain linear, they may estimate around 55,800 million miles. Students can justify their answer by using the plot and past trend, or by offering some other valid reason.

ACTIVITY SHEET 1

Lesson 1: Variables and Formulas (Problems 2 and 13)

NAME ______________________________

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*

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

Lesson 2: Formulas That Manage Money (Problems 6, 7, and 9)

NAME ______________________________

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*
______	______	______	______	______	______	______
______	______	______	______	______	______	______
______	______	______	______	______	______	______
______	______	______	______	______	______	______

Table 2.3:
Gross Salary

Employee	Hourly Pay Rate *R*	Hours Worked *H*	Gross Pay *G*	Federal Income Tax *FI*	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	______	______	______	______

Lesson 2: Formulas That Manage Money (Problems 6, 7, and 9)

NAME ______________________________

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	________	________

ACTIVITY SHEET 3

Lesson 4: Expressions and Rates (Problem 15)

NAME ____

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

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
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

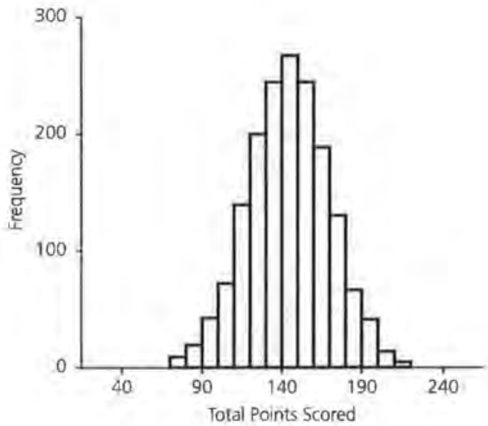
Source: *The American Almanac*, 1994–1995

ACTIVITY SHEET 4

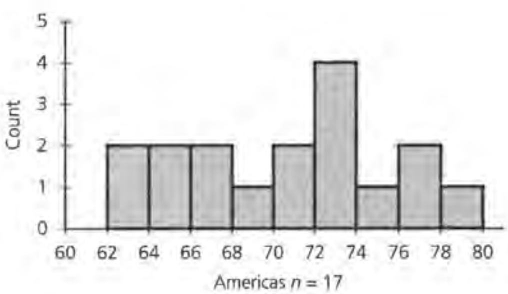
Lesson 7: Formulas That Summarize Variation in Data (Problems 15, 18, 19)

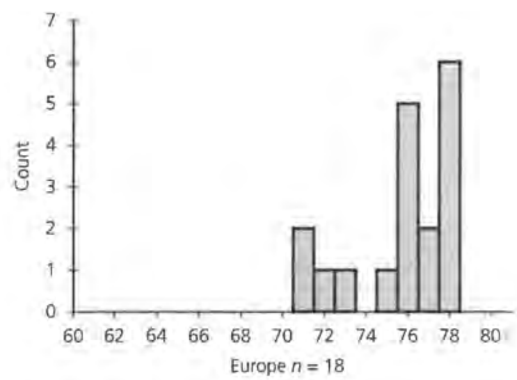
NAME ______________________________

15.

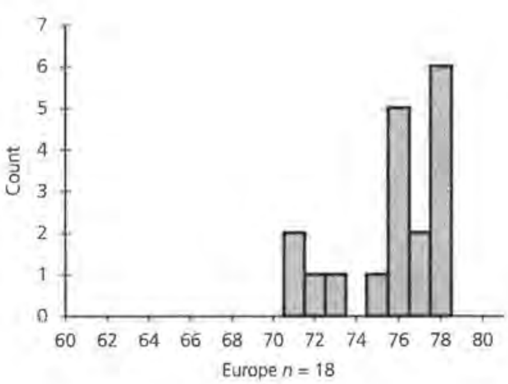


18.





19.



Lesson 8: Comparing Measurements (Problem 6a)

NAME ______________________________

z-Scores for SAT and ACT

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

Source: *American Almanac,* 1994–95

ACTIVITY SHEET 6

Lesson 8: Comparing Measurements (Problem 11c)

NAME ______________________________

NBA Basketball Hall of Fame, 1990–1994

Year	Player	Field Goal %	*z*-Score	Rebounds	*z*-Score
1991	Archibald, Nate	.467	________	2,046	________
1993	Bellamy, Walt	.516	________	14,241	________
1990	Bing, Dave	.441	________	3,420	________
1994	Blazejowski, Carol	.546	________	1,015	________
1991	Cowens, Dave	.460	________	10,444	________
1993	Erving, Julius (Dr. J.)	.507	________	5,601	________
1991	Gallatin, Harry	.398	________	6,684	________
1992	Hawkins, Connie	.467	________	3,971	________
1990	Hayes, Elvin	.452	________	16,279	________
1993	Issel, Dan	.506	________	5,707	________
1990	Johnston, Neil	.444	________	5,856	________
1992	Lanier, Bob	.514	________	9,698	________
1993	McGuire, Dick	.389	________	2,784	________
1993	Meyers, Ann	.500	________	819	________
1990	Monroe, Earl	.464	________	2,796	________
1993	Murphy, Calvin	.482	________	2,103	________
1993	Walton, Bill	.521	________	4,923	________

Source: *The Universal Almanac,* 1995

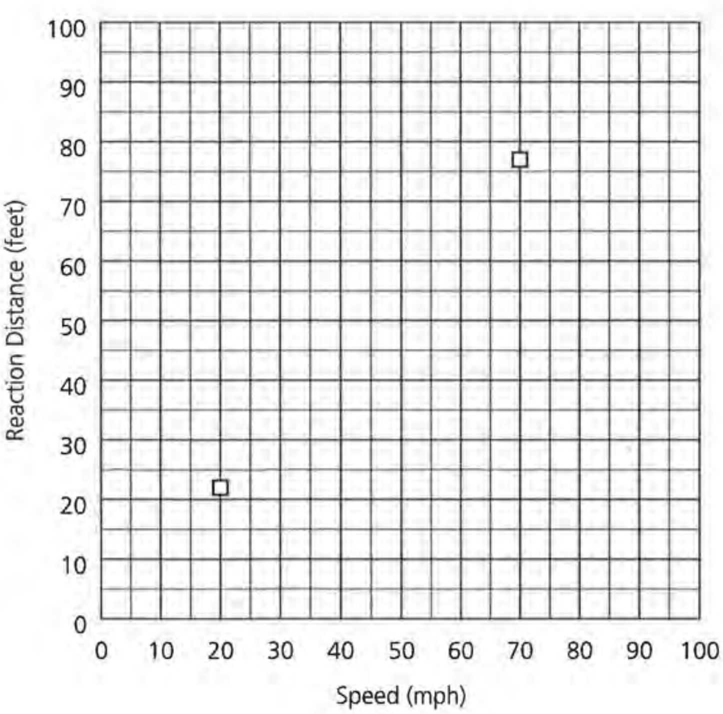
Lesson 9: An Introduction to Functions (Problem 4)

NAME ______________________________

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



Data-Driven Mathematics
Procedures for Using the TI-83

I. Clear menus

ENTER will execute any command or selection. Before beginning a new problem, previous instructions or data should be cleared. Press **ENTER** after each step below.

1. To clear the function menu, **Y=**, place the curser anyplace in each expression, **CLEAR**
2. To clear the list menu, **2nd MEM ClrAllLists**
3. To clear the draw menu, **2nd Draw ClrDraw**
4. To turn off any statistics plots, **2nd STATPLOT PlotsOff**
5. To remove user created lists from the Editor, **STAT SetUpEditor**

II. Basic information

1. A rule is active if there is a dark rectangle over the option. See Figure 1.

```
Plot1 Plot2 Plot3
\Y1 = 170 + 7(X–72)
\Y2 = –.35X – 22
\Y3 =
\Y4 =
\Y5 =
\Y6 =
\Y7 =
```

FIGURE 1

On the screen above, Y1 and Plot1 are active; Y2 is not. You may toggle Y1 or Y2 from active to inactive by putting the cursor over the = and pressing **ENTER**. Arrow up to **Plot1** and press **ENTER** to turn it off; arrow right to **Plot2** and press **ENTER** to turn it on, etc.

2. The Home Screen (Figure 2) is available when the blinking cursor is on the left as in the diagram below. There may be other writing on the screen. To get to the Home Screen, press **2nd QUIT**. You may also clear the screen completely by pressing **CLEAR**.

```
7 → X
                    7
Y1
                   11
█
```

FIGURE 2

3. The variable x is accessed by the **X, T, Θ, *n*** key.
4. Replay option: **2nd ENTER** allows you to back up to an earlier command. Repeated use of **2nd ENTER** continues to replay earlier commands.
5. Under **MATH**, the **MATH** menu has options for fractions to decimals and decimals to fractions, for taking *n*th roots, and for other mathematical operations. **NUM** contains the absolute value function as well as other numerical operations (Figure 3).

```
MATH NUM CPX PRB
1: abs(
2: round(
3: iPart(
4: fPart(
5: int(
6: min(
7↓max(
```

FIGURE 3

III. The STAT Menus

1. There are three basic menus under the **STAT** key: **EDIT**, **CALC**, and **TESTS**. Data are entered and modified in the **EDIT** mode; all numerical calculations are made in the **CALC** mode; statistical tests are run in the **TEST** mode.
2. **Lists and Data Entry**
Data is entered and stored in Lists (Figure 4). Data will remain in a list until the list is cleared. Data can be cleared using **ClrLIST L_i** or (List name), or by placing the cursor over the List heading and selecting **ClrLIST ENTER**. To enter data, select **STAT EDIT** and with the arrow keys move the cursor to the list you want to use.

Type in a numerical value and press **ENTER.** Note that the bottom of the screen indicates the List you are in and the list element you have highlighted. 275 is the first entry in L1. (It is sometimes easier to enter a complete list before beginning another.)

L1	L2	L3
275	67	190
5311	144	120
114	64	238
2838	111	153
15	90	179
332	68	207
3828	94	153
L1 (1) = 275		

FIGURE 4

For data with varying frequencies, one list can be used for the data, and a second for the frequency of the data. In Figure 5 below, the L5(7) can be used to indicate that the seventh element in list 5 is 4, and that 25 is a value that occurs 4 times.

L4	L5	L6
55	1	-------
50	3	
45	6	
40	14	
35	12	
30	9	
25	4	
L5 (7) = 4		

FIGURE 5

3. **Naming Lists**
Six lists are supplied to begin with. L1, L2, L3, L4, L5, and L6 can be accessed also as **2nd** $\mathbf{L_i}$. Other lists can be named using words as follows. Put the cursor at the top of one of the lists. Press **2nd INS** and the screen will look like that in Figure 6.

	L1	L2 1
	-------	-------
Name =		

FIGURE 6

The alpha key is on, so type in the name (up to five characters) and press **ENTER.** (Figure 7)

PRICE	L1	L2 2
	-------	-------
PRICE(1) =		

FIGURE 7

Then enter the data as before. (If you do not press **ENTER,** the cursor will remain at the top and the screen will say "error: data type.") The newly named list and the data will remain until you go to Memory and delete the list from the memory. To access the list for later use, press **2nd LIST** and use the arrow key to locate the list you want under the **NAMES** menu. You can accelerate the process by typing **ALPHA P** (for price). (Figure 8) Remember, to delete all but the standard set of lists from the editor, use SetUp Editor from the **STAT** menu.

NAMES OPS MATH
↑PRICE
: RATIO
: RECT
: RED
: RESID
: SATM
↓SATV

FIGURE 8

4. Graphing Statistical Data

General Comments

- Any graphing uses the **GRAPH** key.
- Any function entered in Y1 will be graphed if it is active. The graph will be visible only if a suitable viewing window is selected.
- The appropriate x- and y-scale can be selected in **WINDOW**. Be sure to select a scale that is suitable to the range of the variables.

Statistical Graphs

To make a statistical plot, select **2nd Y=** for the **STAT PLOT** option. It is possible to make three plots concurrently if the viewing windows are identical. In Figure 9, Plots 2 and 3 are off, Plot 1 is a scatter plot of the data (Costs, Seats), Plot 2 is a scatter plot of (L3,L4), and Plot 3 is a box plot of the data in L3.

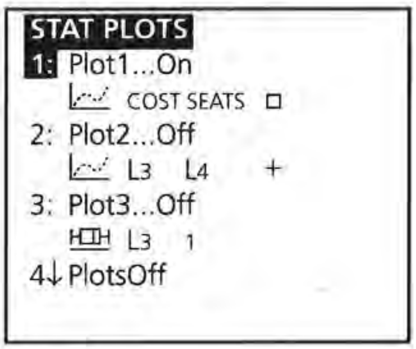


FIGURE 9

Activate one of the plots by selecting that plot and pressing **ENTER.**

Choose **ON**, then use the arrow keys to select the type of plot (scatter, line, histogram, box plot with outliers, box plot, or normal probability plot). (In a line plot, the points are connected by segments in the order in which they are entered. It is best used with data over time.) Choose the lists you wish to use for the plot. In the window below, a scatter plot has been selected with the x-coordinate data from COSTS, and the y-coordinate data from SEATS. (Figure 10) (When pasting in list names, press **2nd LIST,** press **ENTER** to activate the name and **ENTER** again to locate the name in that position.)

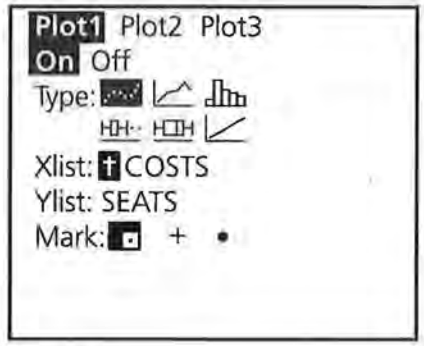


FIGURE 10

For a histogram or box plot, you will need to select the list containing the data and indicate whether you used another list for the frequency or are using 1 for the frequency of each value. The x-scale selected under **WINDOW** determines the width of the bars in the histogram. It is important to specify a scale that makes sense with the data being plotted.

5. Statistical Calculations

One-variable calculations such as mean, median, maximum value of the list, standard deviation, and quartiles can be found by selecting **STAT CALC 1-Var Stats** followed by the list in which you are interested. Use the arrow to continue reading the statistics. (Figures 11, 12, 13)

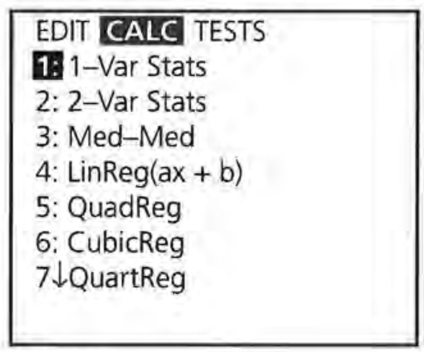


FIGURE 11

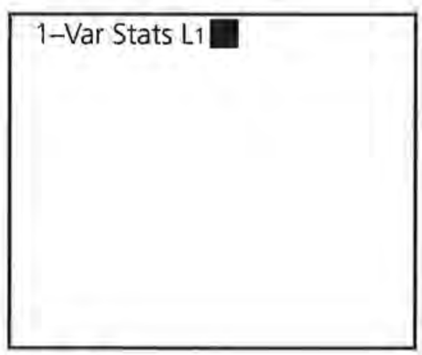


FIGURE 12

```
1–Var Stats
x̄ = 1556.20833
Σx = 37349
Σx² = 135261515
Sx = 1831.353621
σx = 1792.79449
↓n = 24
```

FIGURE 13

Calculations of numerical statistics for bivariate data can be made by selecting two variable statistics. Specific lists must be selected after choosing the **2-Var Stats** option. (Figure 14)

```
2–Var Stats L1, L
2
```

FIGURE 14

Individual statistics for one- or two-data sets can be obtained by selecting **VARS Statistics,** but you must first have calculated either 1-Var or 2-Var Statistics. (Figure 15)

```
X/Y Σ EQ TEST PTS
1: n
2: x̄
3: Sx
4: σx
5: ȳ
6: Sy
7↓σy
```

FIGURE 15

6. **Fitting Lines and Drawing their graphs**
Calculations for fitting lines can be made by selecting the appropriate model under **STAT CALC: Med-Med** gives the median fit regression; **LinReg** the least-squares linear regression, and so on. (Note the only difference between **LinReg (ax+b)** and **LinReg (a+bx)** is the assignment of the letters a and b.) Be sure to specify the appropriate lists for *x* and *y*. (Figure 16)

```
Med–Med LCal, LFA
CAL
```

FIGURE 16

To graph a regression line on a scatter plot, follow the steps below:

- Enter your data into the Lists.
- Select an appropriate viewing window and set up the plot of the data as above.
- Select a regression line followed by the lists for *x* and *y*, **VARS Y-VARS Function** (Figures 17, 18) and the Yi you want to use for the equation, followed by **ENTER.**

```
VARS Y-VARS
1: Function...
2: Parametric...
3: Polar...
4: On/Off...
```

FIGURE 17

```
Med–Med _CAL, LFA
CAL, Y1
```

FIGURE 18

The result will be the regression equation pasted into the function Y1. Press **GRAPH** and both the scatter plot and the regression line will appear in the viewing window. (Figures 19, 20).

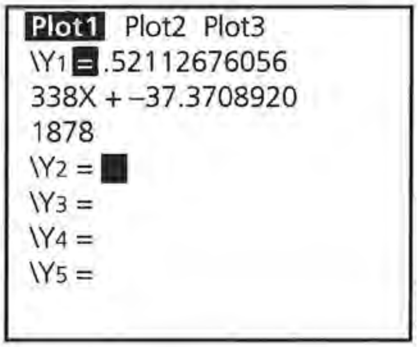


FIGURE 19

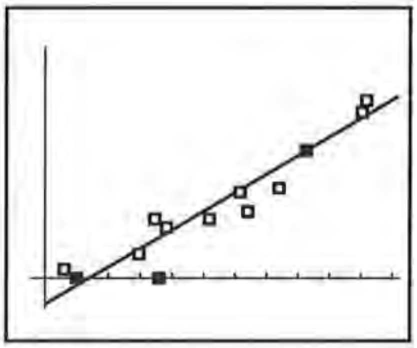


FIGURE 20

- There are two cursors that can be used in the graphing screen.

TRACE activates a cursor that moves along either the data (Figure 21) or the function entered in the Y-variable menu (Figure 22). The coordinates of the point located by the cursor are given at the bottom of the screen.

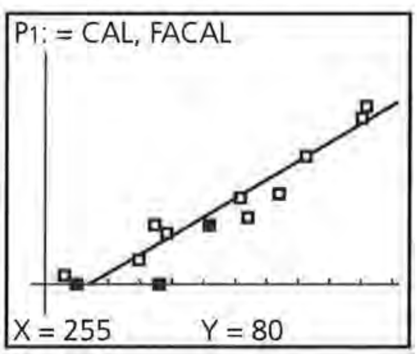


FIGURE 21

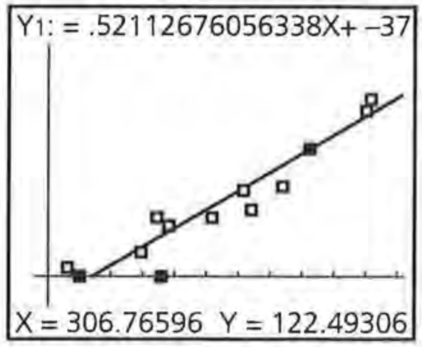


FIGURE 22

Pressing **GRAPH** returns the screen to the original plot. The up arrow key activates a cross cursor that can be moved freely about the screen using the arrow keys. See Figure 23.

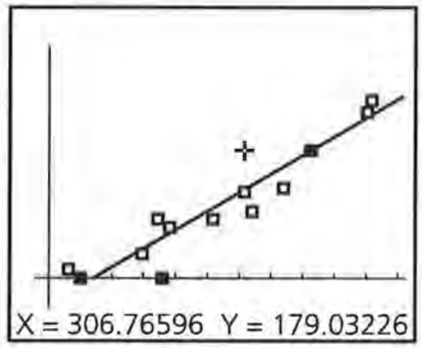


FIGURE 23

Exact values can be obtained from the plot by selecting **2nd CALC** Value. Select **2nd CALC Value ENTER**. Type in the value of x you would like to use, and the exact ordered pair will appear on the screen with the cursor located at that point on the line. (Figure 24)

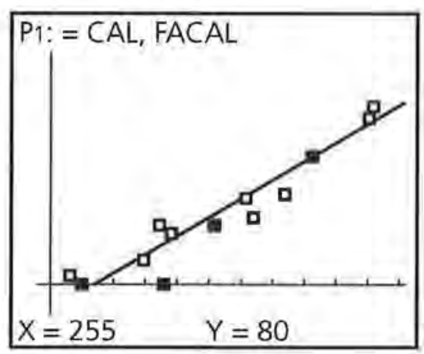


FIGURE 24

IV. Evaluating an expression:

To evaluate $y = .225\,x - 15.6$ for $x = 17$, 20, and 24, you can:

1. Type the expression in Y1, return to the home screen, 17 **STO X,T,Θ,n ENTER, VARS Y1=1VARS Function Y1 ENTER ENTER.** (Figure 25)

```
17 → X
                 17
Y1
       −11.75833333
█
```

FIGURE 25

Repeat the process for $x = 20$ and 24.

2. Type $17^2 - 4$ for $x = 17$, **ENTER.** (Figure 26) Then use **2nd ENTRY** to return to the arithmetic line. Use the arrows to return to the value 17 and type over to enter 20.

```
.225*17 − 15.6
              −11.775
.225*2█ − 15.6
```

FIGURE 26

You can also find the value of x by using the table command. Select **2nd TblSet** (Figure 27). (Y1 must be turned on.) Let **TBlMin** = 17, and the increment **ΔTbl** = 1.

```
TABLE SETUP
  TblMin = 17
  ΔTbl = █
Indpnt:  Auto  Ask
Depend:  Auto  Ask
```

FIGURE 27

Select **2nd TABLE** and the x- and y-values generated by the equation in Y1 will be displayed. (Figure 28)

X	Y1	
17	-11.76	
18	-11.53	
19	-11.31	
20	-11.08	
21	-10.86	
22	-10.63	
23	-10.41	

X = 17

FIGURE 28

V. Operating with Lists

1. A list can be treated as a function and defined by placing the cursor on the label above the list entries. List 2 can be defined as L1 + 5. (Figure 29)

L1	L2	L3
275	-------	190
5311		120
114		238
2838		153
15		179
332		207
3828		153

L2 = L1 + 5

FIGURE 29

Pressing **ENTER** will fill List 2 with the values defined by L1+5. (Figure 30)

L1	L2	L3
275	280	190
5311	5316	120
114	119	238
2838	2843	153
15	20	179
332	337	207
3828	3833	153

L2(1) = 280

FIGURE 30

2. List entries can be cleared by putting the cursor on the heading above the list, and selecting **CLEAR** and **ENTER.**

3. A list can be generated by an equation from Y= over a domain specified by the values in L1 by putting the cursor on the heading above the list entries. Select **VARS Y-VARS Function Y1 ENTER (L1) ENTER.** (Figure 31)

L1	L2	L3
120	12	-------
110	14	
110	12	
110	11	
100	7	
100	6	
120	9	

L3 = Y1(L1)

FIGURE 31

4. The rule for generating a list can be attached to the list and retrieved by using quotation marks (**ALPHA** +) around the rule. (Figure 32) Any change in the rule (Y1 in the illustration) will result in a change in the values for L1. To delete the rule, put the cursor on the heading at the top of the list, press **ENTER**, and then use the delete key. (Because L1 is defined in terms of **CAL**, if you delete **CAL** without deleting the rule for L1 you will cause an error.)

CAL	FACAL	L1 5
255	80	-------
305	120	
410	180	
510	250	
320	90	
370	125	
500	235	

L1 = "Y1(LCAL)"

FIGURE 32

VI. Using the DRAW Command

To draw line segments, start from the graph of a plot, press **2ND DRAW**, and select **Line(.** (Figure 33)

DRAW POINTS STO
1: ClrDraw
2: Line(
3: Horizontal
4: Vertical
5: Tangent(
6: DrawF
7↓Shade(

FIGURE 33

This will activate a cursor that can be used to mark the beginning and ending of a line segment. Move the cursor to the beginning point and press **ENTER**; use the cursor to mark the end of the segment, and press **ENTER** again. To draw a second segment, repeat the process. (Figure 34)

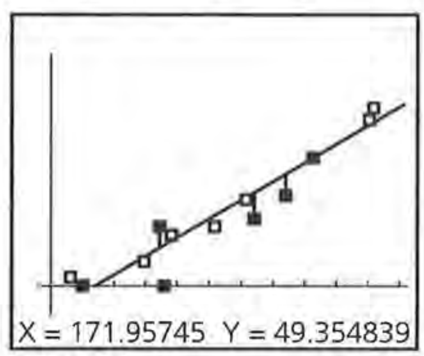


FIGURE 34

VII. Random Numbers

To generate random numbers, press **MATH** and **PRB**. This will give you access to a random number function, **rand**, that will generate random numbers between 0 and 1 or **randInt(** that will generate random numbers from a beginning integer to an ending integer for a specified set of numbers. (Figure 35) In Figure 36, five random numbers from 1 to 6 were generated.

```
MATH  NUM  CPX  PRB
1: rand
2: nPr
3: nCr
4: !
5: randInt(
6: randNorm(
7: randBin(
```

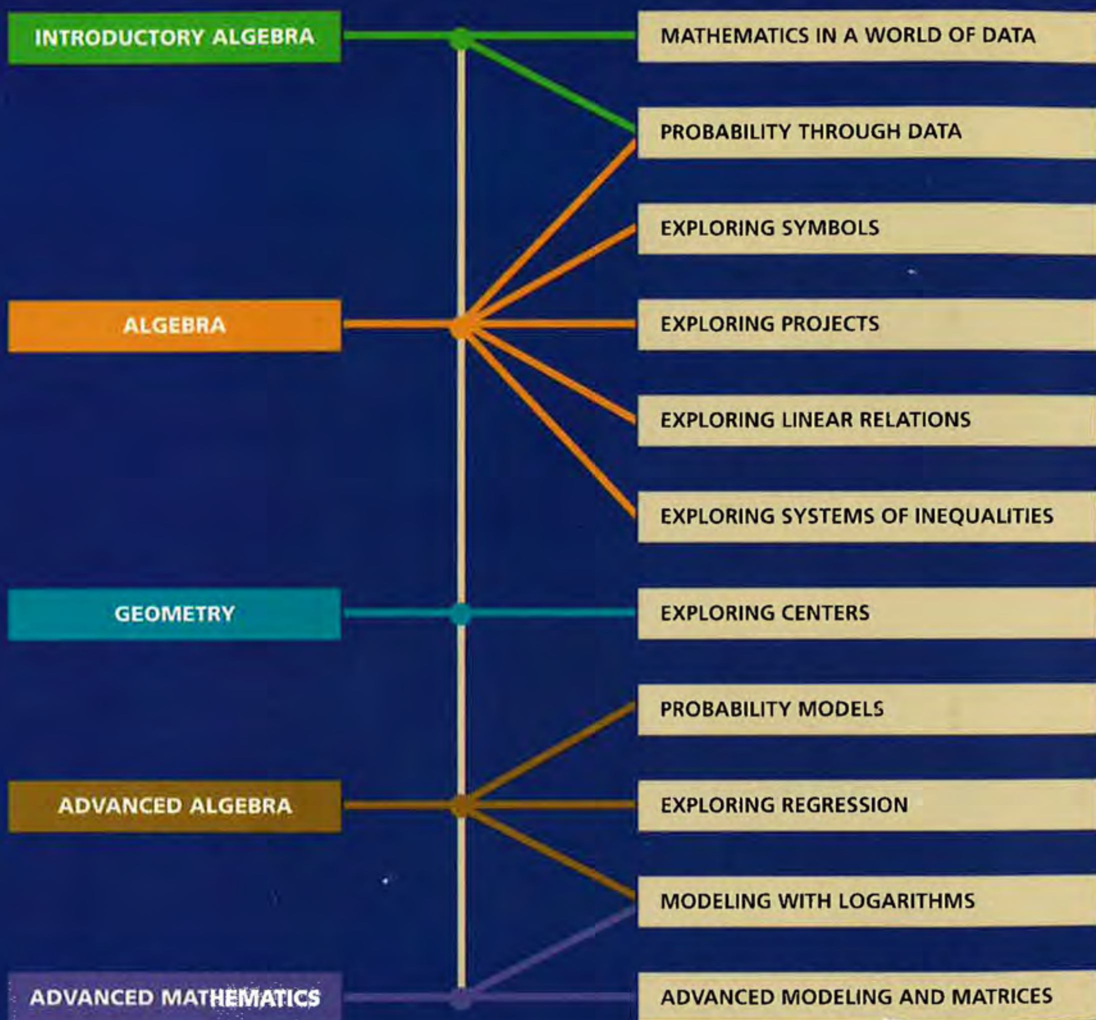
FIGURE 35

```
randInt(1, 6, 5,)
      (2 4 2 5 3)
```

FIGURE 36

Pressing **ENTER** will generate a second set of random numbers.

Data-Driven Mathematics is a series of modules written by teachers and statisticians that focuses on the use of real data and statistics to motivate traditional mathematics topics. This chart suggests which modules could be used to supplement specific middle-school and high-school mathematics courses.



Dale Seymour Publications® is a leading publisher of K–12 educational materials in mathematics, thinking skills, science, language arts, and art education.

90000
9 781572 322349
ISBN 1-57232-234-9
21175

DALE SEYMOUR PUBLICATIONS®