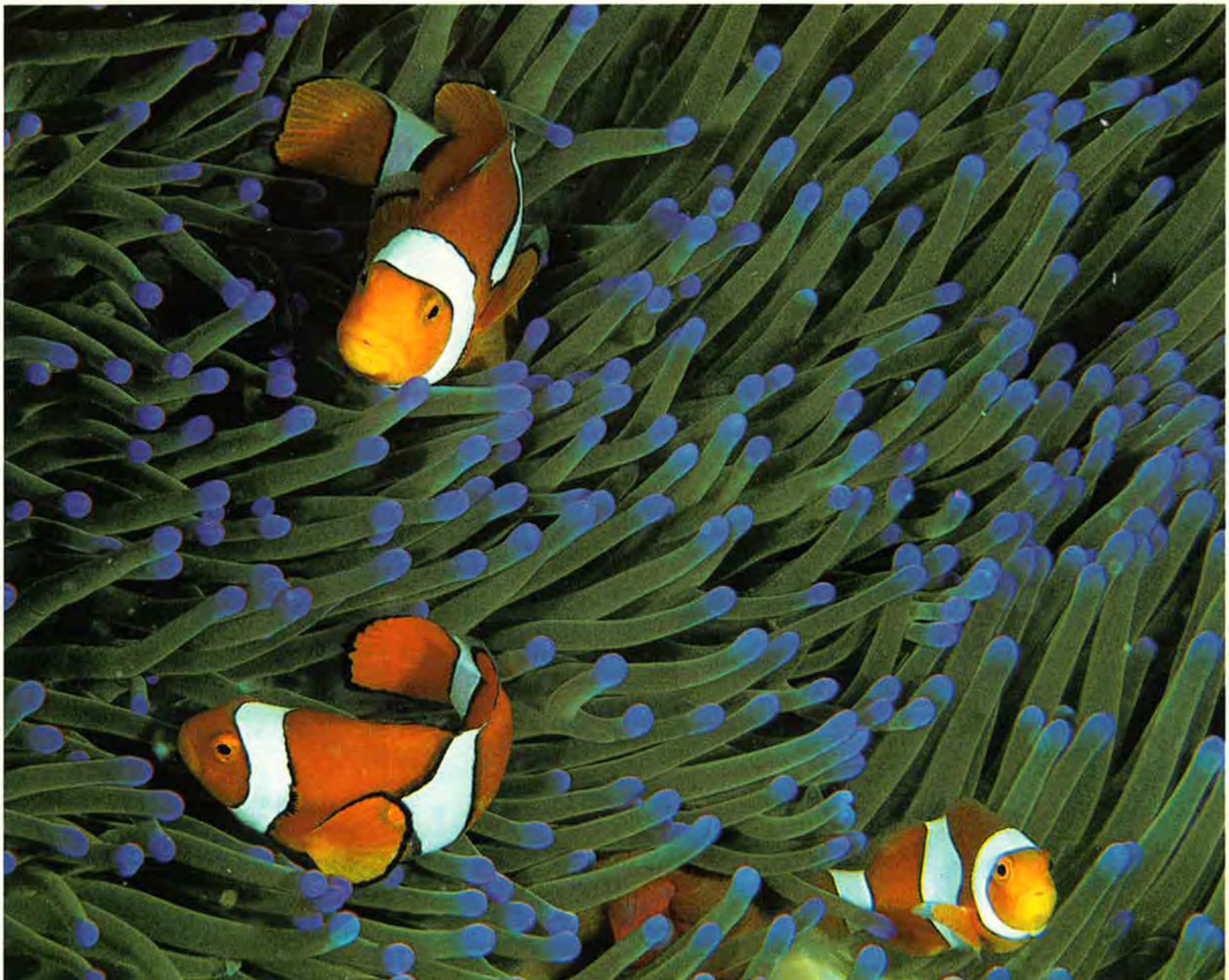
ADVANCED ALGEBRA ADVANCED MATHEMATICS

Modeling with Logarithms

JACK BURRILL, MIRIAM CLIFFORD, JAMES LANDWEHR

DATA-DRIVEN MATHEMATICS



DALE SEYMOUR PUBLICATIONS®

Modeling with Logarithms

DATA-DRIVEN MATHEMATICS

Jack Burrill, Miriam Clifford, and James M. Landwehr

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 Mathematics Editor: Nancy R. Anderson

Project Editor: John Sullivan

Production/Manufacturing Director: Janet Yearian

Production/Manufacturing Manager: Karen Edmonds

Production Coordinator: Roxanne Knoll

Design Manager: Jeff Kelly

Cover and Text Design: Christy Butterfield

Cover Photo: Louisa Preston

Photo, page 3: William Means/Tony Stone Images, Inc.

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

Printed in the United States of America.

Order number DS21185

ISBN 1-57232-250-0

1 2 3 4 5 6 7 8 9 10-ML-03 02 01 00 99 98

This Book Is Printed On Recycled Paper

DALE SEYMOUR PUBLICATIONS®

Authors

Jack Burrill
National Center for Mathematics Sciences Education
University of Wisconsin-Madison
Madison, Wisconsin

Miriam Clifford
Nicolet High School
Glendale, Wisconsin

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

Consultants

Patrick Hopfensperger
Homestead High School
Mequon, Wisconsin

Christine Lucas
Whitefish Bay High School
Whitefish Bay, Wisconsin

Vince O'Connor
Milwaukee Public Schools
Milwaukee, Wisconsin

Matthew Parlier
Whitnall High School
Greenfield, Wisconsin
Vincent High School
Milwaukee, Wisconsin

Data-Driven Mathematics Leadership Team

Gail F. Burrill
National Center for Mathematics Sciences Education
University of Wisconsin-Madison
Madison, Wisconsin

Miriam Clifford
Nicolet High School
Glendale, Wisconsin

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

Richard Scheaffer
University of Florida
Gainesville, Florida

Kenneth Sherrick
Berlin High School
Berlin, Connecticut

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
- 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
- Kay Williams and Judith O'Fallon for advice and suggestions in the early stages of the writing
- Richard Crowe, Christine Lucas, and Matthew Parlier for their thoughtful and careful review of the early drafts
- The many students and teachers from Nicolet High School and Whitnall High School who helped shape the ideas as they were being developed

Table of Contents

About *Data-Driven Mathematics*

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

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

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

Using This Module

There are many patterns in the world that can be described by mathematics. Mathematical modeling is the process of finding, describing, analyzing, and evaluating such patterns using mathematics. The first step in building such a model is to recognize different categories of patterns and to understand the underlying mathematical structure within those categories that can help in the search for an appropriate mathematical model.

In this module, you will explore ways to find a mathematical model for problems involving bivariate data. You will use data sets such as the federal debt over time, decibel measures from various sounds, and the number of motor vehicles registered in the United States to investigate similarities and differences among patterns. You will study the effects of scale changes and transformations on data plots and on the graphs of various mathematical functions. Logarithms are introduced graphically and numerically in a nontraditional way that emphasizes their role in mathematical modeling. You will use algebra skills and concepts developed in the module to create mathematical models. These models are used to answer questions, summarize results, and make predictions about variables. Correlation is introduced as an assessment of the linear relationship between two variables and as an aid in the modeling process.

Modeling with Logarithms is divided into three units.

Unit I: Patterns and Scale Changes

Mathematicians and statisticians represent and examine data patterns in different forms: numeric, geometric (graphs or pictures), and symbolic (formulas). Each representation yields different information and aids in understanding. The interpretation of a data pattern may also be affected by the scale or units. Changing the units from centimeters to meters in a data set changes the appearance of the number pattern or graph, which can influence the message the data set conveys. Lesson 1 is devoted to the study of patterns in data and their representations. Lesson 2 examines the effects of unit or scale change upon the graphic representation of the data.

Unit II: Functions and Transformations

There are some fundamental functions one should be familiar with in both symbolic and graphic form. Often the graph of a function can be altered in a way that would make the process of mathematical modeling simpler. As you relate the shape of the graph to the equation of a function, you will learn to use functions to transform data that alter the graphic representation. Lesson 3 reviews the relationships among some very useful functions, their symbolic expressions, and their graphs. Lesson 4 examines patterns in graphs and deals with the concepts of increasing, decreasing, linear, and nonlinear functions. Lesson 5 is concerned with the transformation of data to linearize a scatter plot. Lesson 6 investigates the changes in a graph relative to inverse functions.

Unit III: Mathematical Models from Data

Several tools can be used to create a mathematical model and analyze how well a model describes a data set. These include the ideas already studied: looking for patterns in functions, transforming data, and considering scale changes. The mathematical model is the most appropriate equation that fits a data set. In searching for a model, you may find more than one that seem appropriate. It is therefore necessary to develop some skills to help determine which is the best one. Lessons 7–10 address the concepts involved in determining "best fit." In Lesson 11, all the modeling skills must be used in an application.

Each lesson begins with an Investigate section to pique the interest with leading questions to direct your thinking and set the stage for the lesson. The material following the Discussion and Practice heading is designed to help you discover the particular knowledge put forth in the unit and may be done in groups when appropriate or as individuals. It will be necessary, however, for you to participate in whole-class discussions to ensure that you are exposed to all the approaches that were used in the solutions. The lesson ends with a summary and a Practice and Applications section which may add an additional challenge. Note: It is important to do all of the problems in a lesson consecutively to follow the development of the concepts.

Unit I

Patterns and Scale Changes

LESSON 1

Patterns

What information can be learned about a pattern expressed numerically?

What information can be learned about a pattern expressed geometrically with a graph or picture?

What information can be learned about a pattern expressed symbolically?

Recognition of patterns is an integral part of the work done by mathematicians and statisticians. Patterns are all around us, and you need to develop an ability to find them in shapes, symbols, and data. This unit will reinforce finding and representing patterns in the world around us.

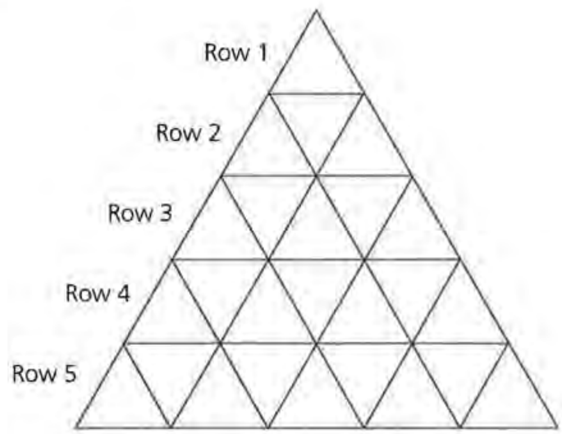
OBJECTIVE

Understand how ordered pairs, graphs, equations, and tables can be used to describe patterns.

"Spaceship Earth Epcot," by William Means

INVESTIGATE

At Epcot Center in Orlando, Florida, *Spaceship Earth,* a 180-foot high geosphere, was constructed with triangles. Can you think of other places you may have seen triangle patterns? Consider the triangle pattern below. How many triangles do you see?



Some relationships can be modeled and studied further. When you model information, it may be helpful to first write it with numbers or symbols. For example, ordered pairs can be used to represent patterns.

Discussion and Practice

In the triangle pattern above, the visual pattern of the row and number of triangles in that row, (row, triangles), can be represented by the ordered pairs (1, 1), (2, 3), (3, 5), (4, 7), ... In symbols this can be written (r, t).

1. Examine each set of ordered pairs below. In each example the ordered pairs represent a pattern in the triangle above. Explain how they relate to a visual pattern in the triangle.
   a. (1, 1), (2, 4), (3, 9), (4, 16), . . .
   b. (1, 0), (2, 1), (3, 2), (4, 3), . . .
2. Write two other ordered-pair patterns that can be generated from the triangle picture. Explain how the numbers in your ordered pairs are related.
3. Graph each set of ordered pairs in Question 1. Write a sentence to describe the patterns you see in the graph. As the first number in the ordered pair increases, what happens to the second number?

4. Graph the ordered pairs you found in Question 2. Write a sentence to describe the patterns you see in the graph. As the first number in the ordered pair increases, what happens to the second number?

An equation may be used to describe the relationship between the first and second number in an ordered pair. Recall that for the ordered pairs (r, t) in the triangle pattern on page 4, r = row number and t = number of triangles in that row.

5. How do the ordered pairs (1, 1), (2, 3), (3, 5), (4, 7), . . . relate to the equation $t = 2r - 1$, where r represents the first number in the ordered pair and t represents the second number in the ordered pair?

6. Write an equation to describe the pattern in at least one of the other sets of ordered pairs in Questions 1 and 2.

7. These data relate to common objects that you probably have in your home. They were collected at a department store. Each row in this table can be considered an ordered pair. Plot the following ordered pairs (length, width) and describe the pattern.

Length	Width
8	10
5	7
4	6
22	28
3.5	5
8.5	11
11	14
16	20
20	30
24	36

Recall from your previous work that lines such as least squares or median-fit lines are used to show the linear trend of a graph. These lines are often used to determine values *between* those given on a table and *beyond* the values given on a table. In most cases, the straight line will not pass through all the points on the graph but is used to summarize the linear relation between the variables, just as mean or median is used to summarize the center of a univariate set of data. Equations of straight lines can be quickly determined from ordered pairs and then be used to make predictions.

8. Draw a line on your graph.

   a. Use the line drawn to determine three ordered pairs that could have also been in the data.

   b. Write an equation for your line.

9. What do you think the data on the table in Question 7 represent? What might be the appropriate units for these data?

Example: Ancestor Patterns

In Salt Lake City, Utah, there is a genealogy library that helps people searching for information about their ancestors. Books containing information such as birth, death, and immigration records sometimes make it possible to locate the names of ancestors who lived several hundred years ago. The number of ancestors you have in past generations forms a mathematical pattern. For example, you have 2 parents and 4 grandparents.

10. Write the information for 10 generations in a table like this.

Generations Ago	Number of Ancestors
1	2
2	______
3	______
4	______
5	______
6	______
7	______
8	______
9	______
10	______

11. Make a scatter plot of the ordered pairs.

12. Write a few sentences to describe the patterns on the graph.

13. Draw a straight line through your scatter plot that appears to come closest to all of the data points, and use it to make some predictions.

   a. How do this graph and its line compare to the data set and line in Question 8?

b. Use your predictions to determine if a linear equation would be a good summary of the pattern. Explain why or why not.

14. Study the relationship between the first and second variables in each of your ordered pairs from the table in Question 10.

a. Write an equation that can be used to describe the relationship.

b. Determine how many ancestors you had 12 generations ago.

c. Suppose you had 33,554,432 ancestors 25 generations ago. How many did you have 24 generations ago? 26 generations ago? Explain how you determined your answers.

Summary

Studying the mathematical properties of patterns helps you make sense out of data. In this module, you will continue to study patterns and their graphs. Notice that some graph patterns are straight and some are curved. All of the data points in a set do not have to lie exactly on a line for the trend to be considered a straight line. A linear equation is used to model straight-line trends. When data follow curved patterns, equations that are not linear may be used to describe their trends.

Practice and Applications

15. List at least two different ways mathematics can be used to show a pattern.

16. Write at least three different words or phrases that can be used to describe trends on the graphs you made.

LESSON 2

Changes in Units on the Axes

What effect will the change of unit or scale have on the numeric representation?

What effect will the change of unit or scale have on the geometric representation?

What effect will the change of unit or scale have on the symbolic representation?

OBJECTIVE

Understand how changes in units affect tables and graphs.

INVESTIGATE

Often the effect of changing the units of measure for the items being graphed or being represented in a table is completely overlooked. For instance, if you wanted to conduct a survey to determine about how much loose change people carry in their pockets or purses, what units could you use?

Discussion and Practice

Collect information from students in class to answer the question "How much loose change are you carrying?"

1. Record the information in a table like the one shown on the next page. Show each amount four different ways, expressing answers in decimal form.

Person	Total Amount Expressed in Number of Pennies	Total Amount Expressed in Number of Dimes	Total Amount Expressed in Number of Quarters*	Total Amount Expressed in Number of Dollars
Example	**57**	**5.7**	**2.28**	**0.57**
1	______	______	______	______
2	______	______	______	______
3	______	______	______	______
4	______	______	______	______
5	______	______	______	______
6	______	______	______	______
7	______	______	______	______
8	______	______	______	______
•	______	______	______	______
•	______	______	______	______

*To express the amount in quarters, divide the number of cents by 25.

2. Make four scatter plots on one coordinate plane using the ordered pairs (person, total amount) for each amount.
   a. Money people carry expressed in number of pennies
   b. Money people carry expressed in number of dimes
   c. Money people carry expressed in number of quarters
   d. Money people carry expressed in number of dollars
3. Find the mean and the median for each column. Would the mean or the median better describe how to represent the typical amount of change a person in the class has? Explain why you made that choice.
4. Use mathematical symbols to describe the relationship between the amount in number of quarters and the amount in number of dimes.
5. Write a summary paragraph to explain the relationship between any two of the units used in Question 4. Discuss why either unit could be used.

The federal debt is the amount of money the federal government owes. Most of it is owed to citizens who lend the government money by buying bonds or treasury bills. The debt increased steadily from 1980–1995 because of deficit spending and increases in interest owed on the debt. Deficit spending occurs when the government spends more in a year than it takes in through taxes and other revenues. The federal deficit is added to the federal debt each year.

6. The federal debt is usually described in trillion dollars, written as \$1,000,000,000,000, with \$1 as the unit. It is hard to grasp how much one trillion dollars is.

   a. A billion seconds is 31.7 years. How long is a trillion seconds?

   b. If a million dollars in \$1,000 bills would make a stack four and one-half inches high, how high would a trillion dollars in the same currency stack?

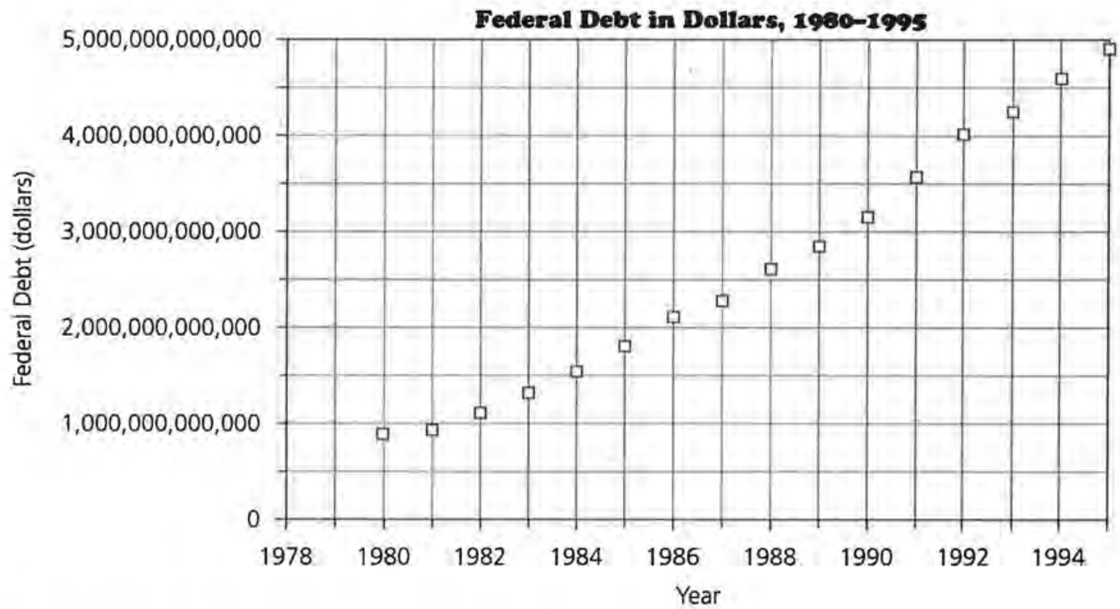
How much is the federal debt increasing each year? Study these data to help you answer the questions.

Year	Federal Debt	Federal Debt (dollars)	Federal Debt (billion dollars)
1980	0.9091 trillion dollars	______	______
1981	0.9949 trillion dollars	______	______
1982	1.137 trillion dollars	______	______
1983	1.372 trillion dollars	______	______
1984	1.565 trillion dollars	______	______
1985	1.818 trillion dollars	______	______
1986	2.121 trillion dollars	______	______
1987	2.346 trillion dollars	2,346,000,000,000	2,346
1988	2.601 trillion dollars	______	______
1989	2.868 trillion dollars	______	______
1990	3.207 trillion dollars	______	______
1991	3.598 trillion dollars	______	______
1992	4.002 trillion dollars	______	______
1993	4.351 trillion dollars	______	______
1994	4.644 trillion dollars	______	______
1995	4.921 trillion dollars	______	______

Source: U.S. Treasury Department

7. Rewrite the numbers in the "Federal Debt" columns in dollars and billion dollars as shown in the row for 1987.

8. The following is a graph with the years on the x-axis and the debt in dollars on the y-axis. Make a conjecture as to what the graph might look like if the debt in the trillion-dollars column had been plotted on the y-axis.



9. Plot (year, federal debt in billion dollars). Label the axes. Describe in a short paragraph what, if any, changes occur in the graph when the units on the y-axis are changed from dollars to billion dollars.

10. The population of the United States in 1980 was about 230,000 thousand people. What was the amount of the federal debt per thousand people in 1980? What was the dollar amount of the federal debt per person in 1980?

11. Our taxes support the federal government. When the federal government spends more money than it receives in taxes and other revenues, it is called "deficit spending," that is, spending money the government does not have. This amount, referred to as the "deficit," must be borrowed and is added to the federal debt each year.

Year	Deficit (billion dollars)
1980	73.4
1981	79.3
1982	128.5
1983	208.7
1984	186.8
1985	213.3
1986	223.1
1987	152.0
1988	153.6
1989	149.9
1990	221.7
1991	269.5
1992	288.7
1993	252.5
1994	205.4
1995	165.5

Source: U.S. Government Printing Office

a. Make a scatter plot of (year, deficit in billion dollars).

b. Compare your Federal Deficit and Federal Debt graphs. Identify the similarities.

Summary

Paying attention to the units attached to a number is important when interpreting data. A unit change or scale change, such as from an amount of money in number of nickels to an amount of money in number of quarters, does not affect the value of the amount. However, a unit change or scale change may affect numeric representations, graphs, and the symbols used to represent numbers.

Practice and Applications

12. If you had to choose a single unit to represent the amount of loose change a person carries, would you prefer to use pennies, dimes, quarters, or dollars? Write an argument supporting your choice.

13. What additional information can you gain from a graph of the federal debt that might not be apparent in the table?

14. Is the size of the debt affected by the units? What units do you prefer to use to describe the federal debt? Explain why.

15. This table contains information about the United States population.

Year	Population (thousands)
1980	226,546
1981	229,466
1982	231,664
1983	233,792
1984	235,825
1985	237,924
1986	240,133
1987	242,289
1988	244,499
1989	246,819
1990	248,718
1991	252,138
1992	255,039
1993	257,800
1994	260,350
1995	262,755

Source: *Statistical Abstract of the United States,* 1997

a. How much was the federal deficit in dollars per person in 1995?

b. Working at \$10.00 per hour, 40 hours per week, how many weeks would it take you to pay your share of the deficit in 1995? How many days?

16. Create a table that contains the amount of federal debt per person for the years 1980–1995.

a. Make a scatter plot of the debt-per-person data for the years 1980–1995 and describe its shape.

b. Use your graph to estimate the federal debt per person for 1996.

c. Is it reasonable to calculate the debt per state? Explain.

Speed Versus Stopping Distance and Height Versus Weight

In driver-education classes, students are usually taught to allow, under normal driving conditions, one car length for every ten miles of speed and more distance in adverse weather or road conditions. The faster a car is traveling, the longer it takes the driver to stop the car. The stopping distance (the total distance required to bring an automobile to a complete stop) depends on the driver-reaction distance (the distance traveled between deciding to stop and actually engaging the brake) and the braking distance (the distance required to bring the automobile to a complete stop once the brake has been applied). The data below represent average driver-reaction distances based on tests conducted by the U.S. Bureau of Public Roads. Average total stopping distances will be investigated later in the module.

Speed (mph)	Driver-Reaction Distance (ft)
20	22
25	28
30	33
35	39
40	44
45	50
50	55
55	61
60	66
65	72
70	77
75	83
80	88

Source: U.S. Bureau of Public Roads

1. Describe any pattern in the data table using the knowledge you gained in this unit.
2. Make a scatter plot of (speed, reaction distance). Can a straight line be used to summarize the trend of the graph? Explain. Draw a line and find its equation.
3. Use the equation to make at least 3 predictions for reaction distances at speeds not included in the table. How accurate do you think your predictions are?

The following data sets come from the Mayo Clinic Family Healthbook relating average height to average weight for both males and females.

Women's Height (inches)	Women's Weight (pounds)
57	117
58	119
59	121
60	123
61	126
62	129
63	133
64	136
65	140
66	143
67	147
68	150
69	153
70	156
71	159

Source: *Mayo Clinic Family Healthbook*

Men's Height (inches)	Men's Weight (pounds)
61	139
62	142
63	144
64	147
65	150
66	153
67	156
68	159
69	162
70	165
71	169
72	172
73	176
74	180
75	185

Source: *Mayo Clinic Family Healthbook*

4. Use the data to graph (women's height, women's weight) or (men's height, men's weight). Draw a straight line that seems to come closest to all the data points on your graph. Find the equation of your line.
5. What does the slope of the line tell you, and how does it relate to the table?
6. The federal deficit per person in 1992 was $1131.98. Write an estimate of this amount with hundred dollars per person as the unit. Write this amount with thousand dollars per person as the unit.

Unit II

Functions and Transformations

LESSON 3

Functions

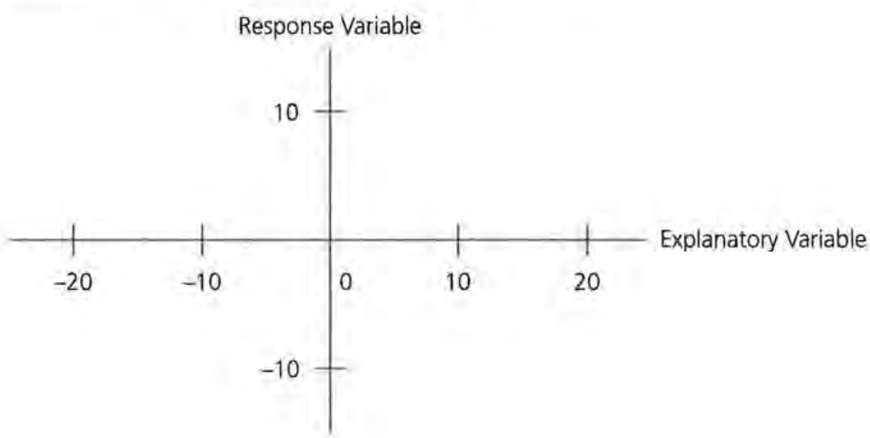
What does the graph of each specific function type look like?

What is the relationship between the coefficients of a function's expression in symbolic form and the graph of that function?

OBJECTIVE

Recognize graphs and equations for different functions.

In modeling, statisticians and mathematicians look for patterns that can be used to explain and/or understand a data set. The process of *mathematical modeling* consists of examining a data set for patterns and looking for a function whose properties most closely represent the data's properties. This lesson will review the properties and shapes of specific functions so that they may be used in this process. Important terms in the study of mathematics and statistics used in mathematical modeling are *response variable, explanatory variable, relations,* and *functions.* When examining data and determining the dependence of one variable upon another, you can identify the dependent variable as the ***response variable.*** The other variable is then referred to as the ***explanatory variable.***



Relations are sets of ordered pairs of the two variables. Within the set of relations is a subset called "functions." ***Functions*** are relations in which every instance of the explanatory variable is paired with a single instance of the response variable. These terms will be used throughout the remainder of this module.

INVESTIGATE

There are many specific functions that are useful in mathematical modeling: *linear functions, logarithmic functions, exponential functions, power functions, quadratic functions, reciprocal functions,* and *square-root functions.*

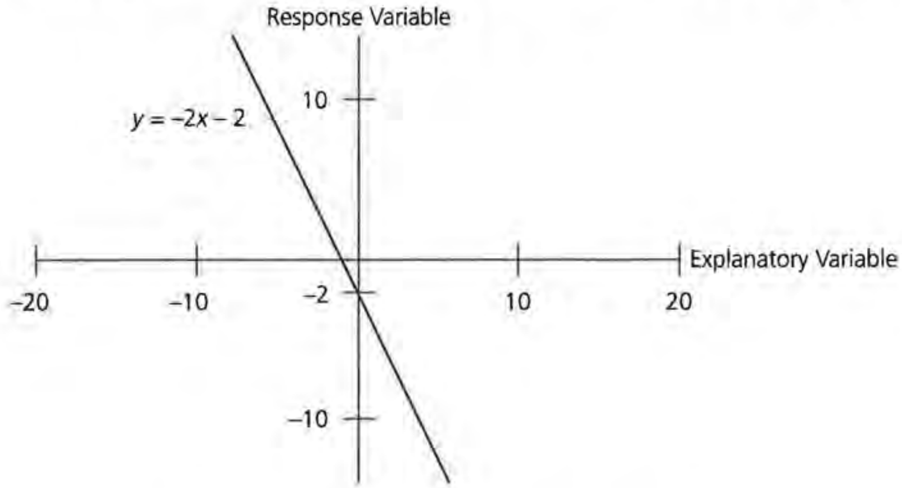
It is helpful to know how the appearance of the graph of a function relates to the data it represents. For instance, what will be the appearance of a graph when the function is increasing? decreasing? constant? What information is gained about the graph of a function by knowing it has an asymptote? And how does changing the rate of change affect the appearance of the graph of a function?

Discussion and Practice

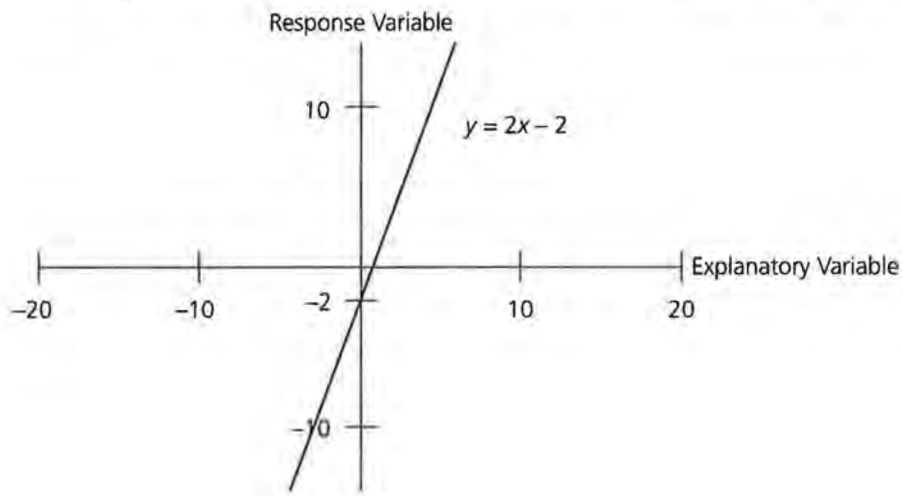
It is important for you to be able to recognize functions by their graphs and equations.

Linear Function A general form of the equation of a linear function is $y = bx + a$ or $y = b(x - c) + a$, where a, b, and c are constants. The graph appears as a straight line. The slope is determined by the numerical value of the constant b in the equation. If b is positive, the line slopes up to the right; and if b is negative, the line slopes down to the right.

This graph intercepts the y-axis at –2 and has a negative slope.



This graph intercepts the y-axis at –2 and has a positive slope.



1. For each of the following linear functions, determine the slope of the line and the coordinates of a point on the line. Then sketch the graph. You may use a graphing utility.

   a. $y = 2x + 50$

   b. $y = -2x + 50$

   c. $y = 2x - 50$

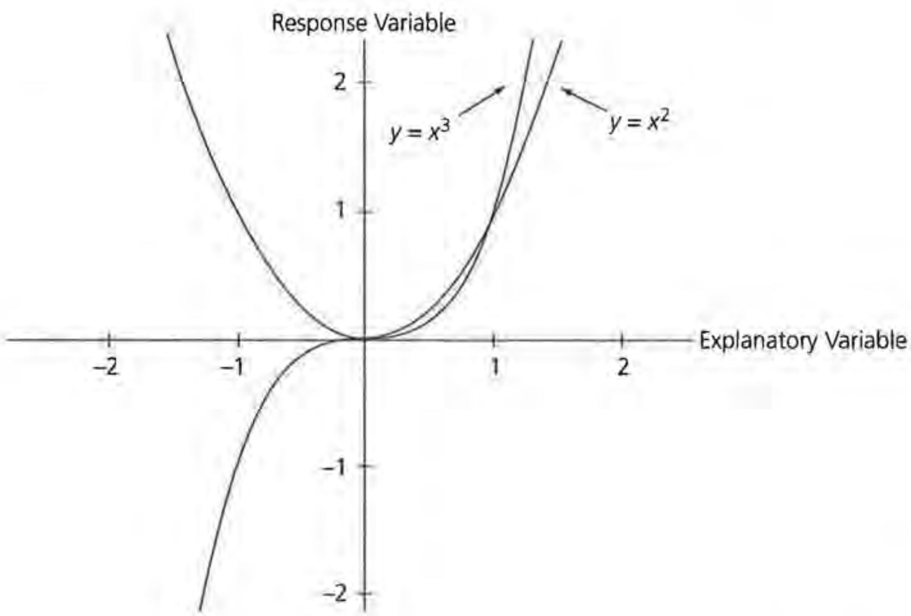
   d. $y = 2(x - 3) + 50$

   e. $y = -2(x - 3) + 50$

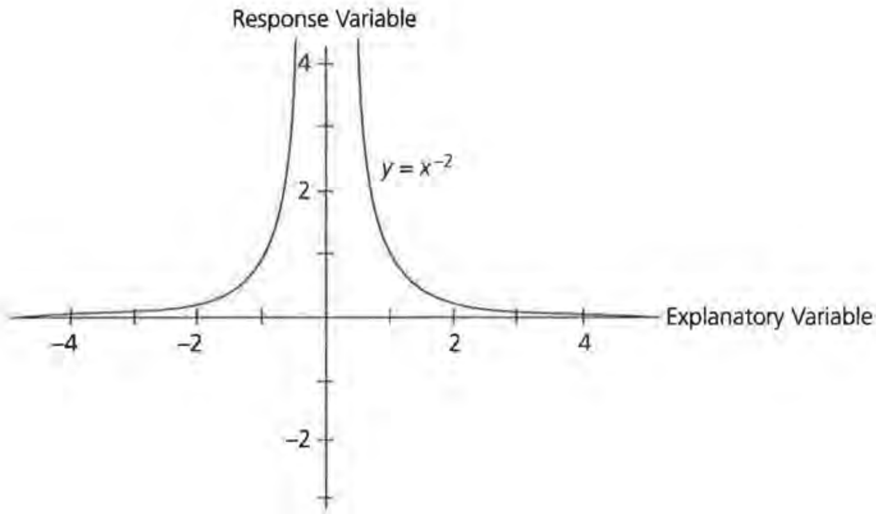
   f. $y = 2(x - 3) - 50$

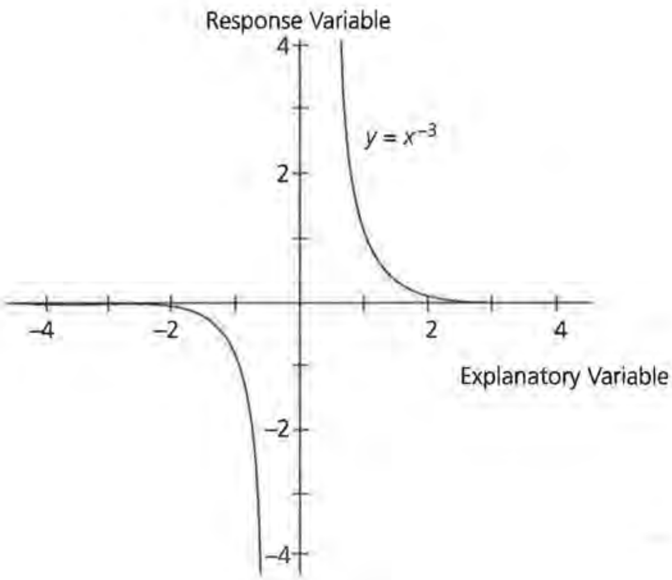
2. Describe the effect each of the constants a, b, and c has on the graph of the equation $y = b(x - c) + a$.

Power Function The general form of the equation of a power function is $y = ax^b$ where a and b are constants. The graph appears as a smooth curve with the amount, and sometimes the direction, of the curvature influenced by the value of b.



The graph above represents two power functions with positive exponents. The next two graphs represent two power functions with negative exponents. In contrast to the power functions with positive exponents, the power functions with negative exponents are decreasing functions for positive values of the explanatory variable. For negative values of the explanatory variable, the power functions with even negative exponents are increasing and those with odd negative exponents are decreasing. The rate of increase or decrease is related to the absolute value of the exponent.





3. Sketch a graph of each function. You may use a graphing utility.

a. $y = x^1$ **b.** $y = x^2$

c. $y = x^3$ **d.** $y = x^4$

e. $y = x^5$

4. Sketch a graph of each function. You may use a graphing utility.

a. $y = x^{-2}$ **b.** $y = x^{-3}$

c. $y = x^{-4}$

5. Describe the general effect the constant b has on the graph $y = x^b$. Let the constant b take on both positive and negative values.

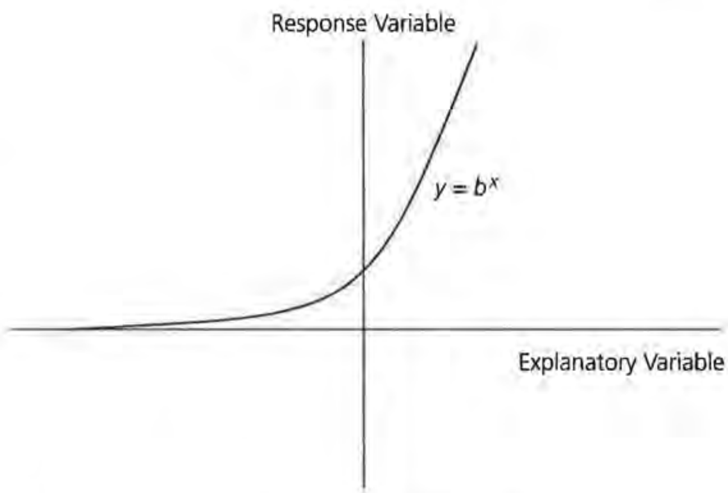
6. Make a conjecture about the effect the constant a has on the graph of $y = ax^b$. With your graphing utility, graph the following to test your conjecture.

a. $y = 2x^2$ b. $y = 3x^2$

c. $y = -2x^2$ d. $y = -3x^2$

e. $y = 5x^2$

Exponential Function The general form of the equation of an exponential function is $y = ab^x$, where a and b are constants and $b > 0$ and $b \neq 1$. The graph appears as a smooth curve that increases or decreases and has a horizontal asymptote as the value of the explanatory variable approaches negative infinity.



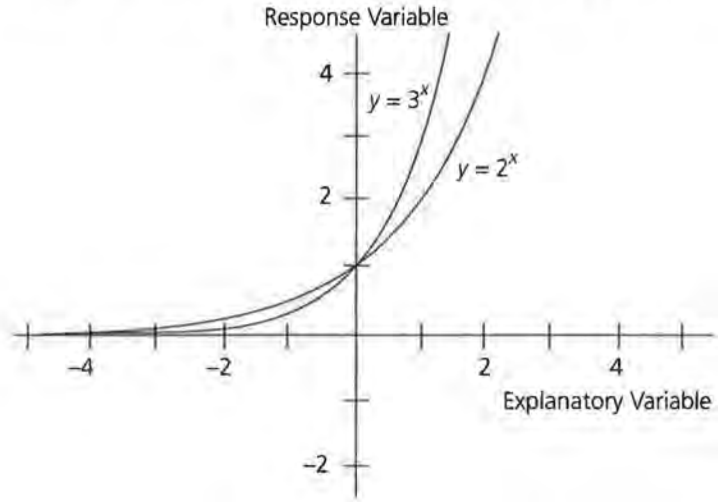
7. Sketch a graph of each function. You may use a graphing utility.

a. $y = 2^x$ b. $y = 3^x$

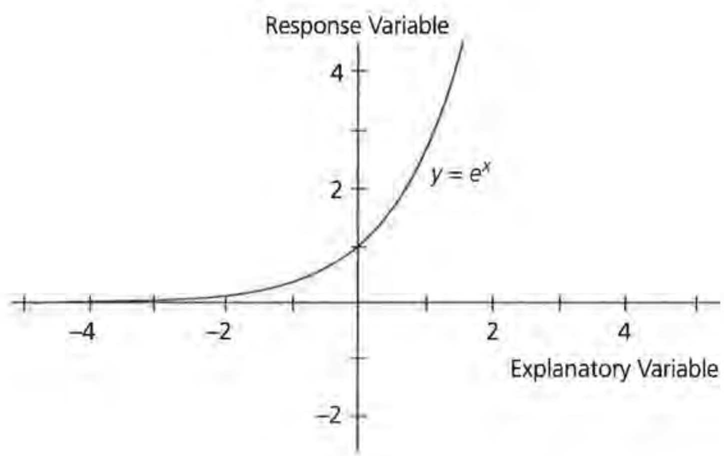
c. $y = 4^x$ d. $y = 5^x$

8. Describe how the graph of an exponential function changes as the base b changes.

Two exponential functions are graphed simultaneously on the graph below to display the likeness in shape and the difference in the rate of increase. Although both functions increase to infinity as the explanatory variable increases, you will notice that $y = 2^x$ increases at a slower rate than $y = 3^x$ does.



The symbol e represents the numerical constant 2.7182818... .
It is an irrational number defined as the limit of $(1 + \frac{1}{x})^x$ as x gets larger and larger and approaches infinity. The letter e was chosen for its discoverer, Leonard Euler. The graph of $y = e^x$ is shown below. You can see that its graph lies between the graphs of $y = 2^x$ and $y = 3^x$, since the value of e lies between 2 and 3.



9. Sketch a graph of each function. You may use a graphing utility.

   a. $y = 2^{-x}$

   b. $y = 3^{-x}$

   c. $y = 4^{-x}$

10. What is the relationship between the graphs of the functions $y = 2^{-x}$ and $y = (\frac{1}{2})^{-x} = 2^{x}$?

In mathematical modeling, you must distinguish between a power function and an exponential function.

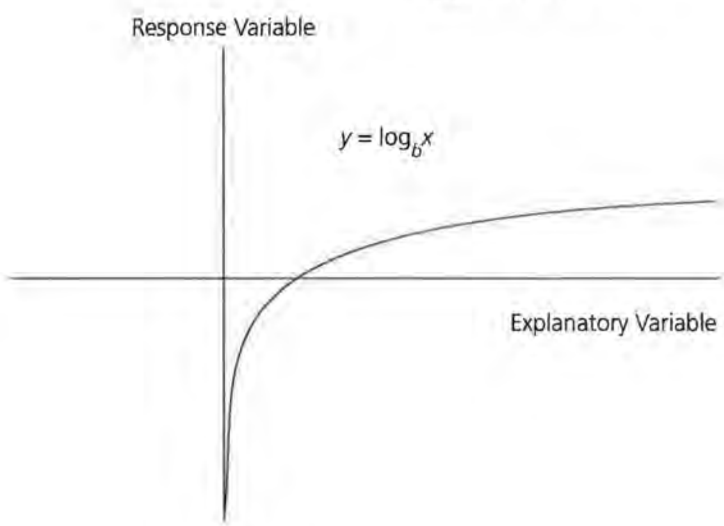
11. Graph the power function $y = x^2$ and the exponential function $y = 2^x$ on the same set of axes.

    a. Describe in a short paragraph the differences between the graphs of these two functions.

    b. Describe a situation that could be modeled by $y = x^2$ and a situation that could be modeled by $y = 2^x$.

12. Without graphing, describe differences between $y = x^3$ and $y = 3^x$.

Logarithmic Function The general form of the equation of a logarithmic function is $y = \log_b x$, where $b > 0$ and $b \neq 1$. The graph appears as a smooth curve and has a vertical asymptote as the value of the explanatory variable approaches zero.

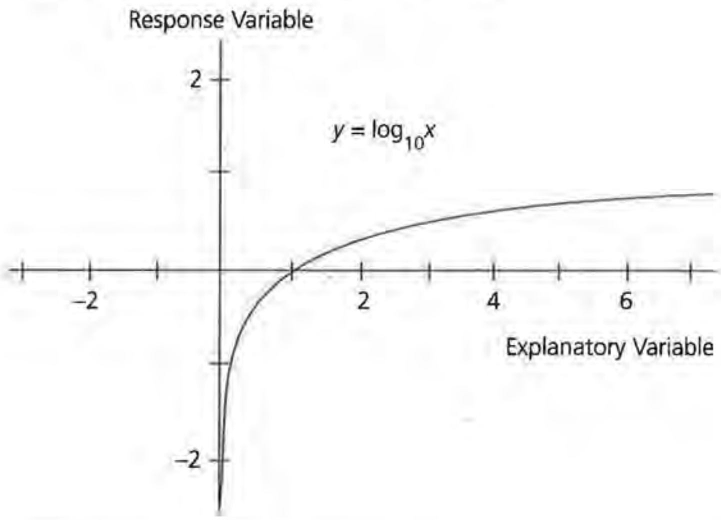


The general rule for converting a logarithmic function from one base to another is:

$$\log_b x = \frac{\log_c x}{\log_c b}$$

For example, $\log_2 6 = \frac{\log_{10} 6}{\log_{10} 2}$

When no base is indicated, such as in $\log x$, the base is understood to be 10; that is, $\log x = \log_{10} x$.



Another logarithmic function is so important in theoretical mathematics that it has its own symbol. It is the logarithm base e, written symbolically as $\ln x$: $\ln x = \log_e x$.

13. Use a graphing utility to graph each function.

a. $y = \log x$

b. $y = \log_5 x$

c. $y = \log_3 x$

d. $y = \log_{0.5} x$

e. $y = \log_{0.75} x$

f. $y = \ln x$

14. Describe the effect that changing the base has on the graph of a logarithmic function.

15. Use your graphing utility to graph each function.

a. $y = \log x + 3$

b. $y = \log(x + 3)$

c. $y = \ln x - 2$

d. $y = \ln(x - 2)$

e. $y = \log(x - 4) - 2$

f. $y = 2\log x$

g. $y = \log 2x$

h. $y = \log 2(x - 2)$

16. Explain how each constant *a, b, c,* and *d* causes the graph of $y = a \log b(x - c) + d$ to differ from the graph of $y = \log x$.

Summary

The knowledge that the equation of a function and its graph are different representations of the same data set is very helpful in the process of modeling. A further knowledge of the effect the various constants have upon the graph of the function is helpful. In this unit, you investigated those items with respect to linear, power, exponential, and logarithmic functions. In the remainder of this module, you will use this knowledge to determine what function might be the best model for the data with which you are working.

Practice and Applications

For each of the following equations make a sketch of its graph. This is a mental exercise, and the graphing utility should be used only to check your results and relative accuracy.

17. $y = 3e^x + 1$

18. $y = 2 \log (x - 3) + 2$

19. $y = 3x + -2$

20. $y = 5(2)^{x+6} - 4$

21. $y = 2(3)^{2-x} + 1$

LESSON 4

Patterns in Graphs

Why are mathematical models used to describe data?

Are there any common patterns that appear in graphs of functions?

A mathematical model is an equation used to describe the response variable in terms of the explanatory variable. If the data set to be modeled has the specific characteristic that a straight line would best describe the pattern formed by its points, it calls for a ***linear model.*** If the pattern seemed more curved, the data set would be said to be calling for a ***nonlinear model.*** If the data set's response variable increases in value while the explanatory variable increases in value, the model being called for is an ***increasing model.*** If the response variable decreases in value while the explanatory variable increases in value, the model called for is a ***decreasing model.***

OBJECTIVES

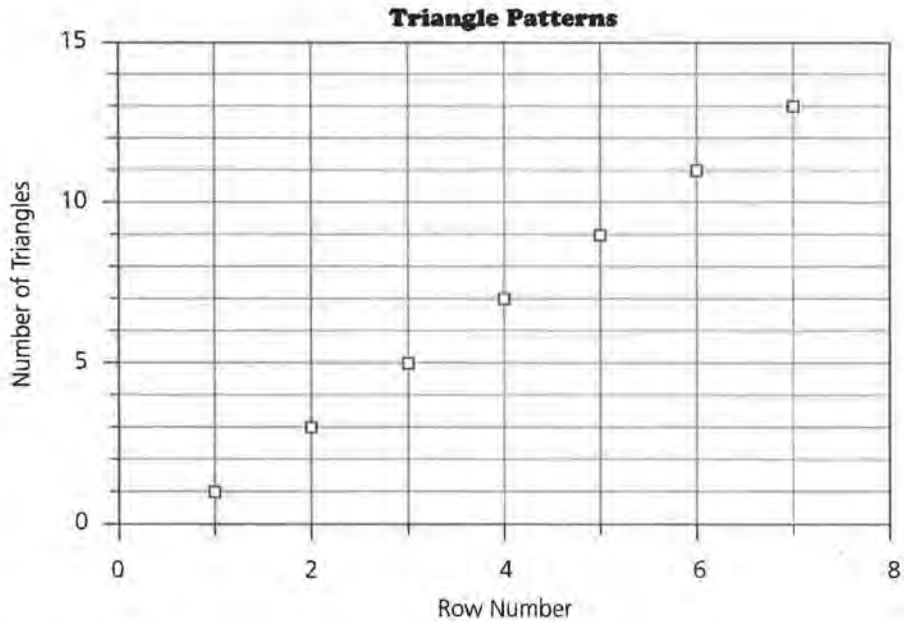
Define a mathematical model and explore different data sets.

Identify which data sets can be represented by linear and nonlinear models.

Make suggestions regarding probable models.

INVESTIGATE

When you investigated the patterns in the triangle in Lesson 1, you considered (row number, number of triangles). The results looked like the graph on the following page.



In this case, the graph appears to represent a data set that would call for a linear model. Is it obvious that it also calls for an increasing model?

When reading information in newspapers and magazines or watching the news on TV, you will often encounter information in the form of a table. This information may be used to answer a question, describe a trend, or tell a story.

Discussion and Practice

1. Is it possible for a data set to be represented by a nonlinear model and also be an increasing model? Explain.
2. Is it possible for a data set to be represented by a linear model and also be a decreasing model? Explain.

Summary

Recognition of linear and nonlinear models as well as increasing and decreasing models is part of the mathematical modeling process.

Practice and Applications

The following tables and graphs provide information that may follow a pattern. For each table or graph in Questions 3–14, look for patterns following this procedure:

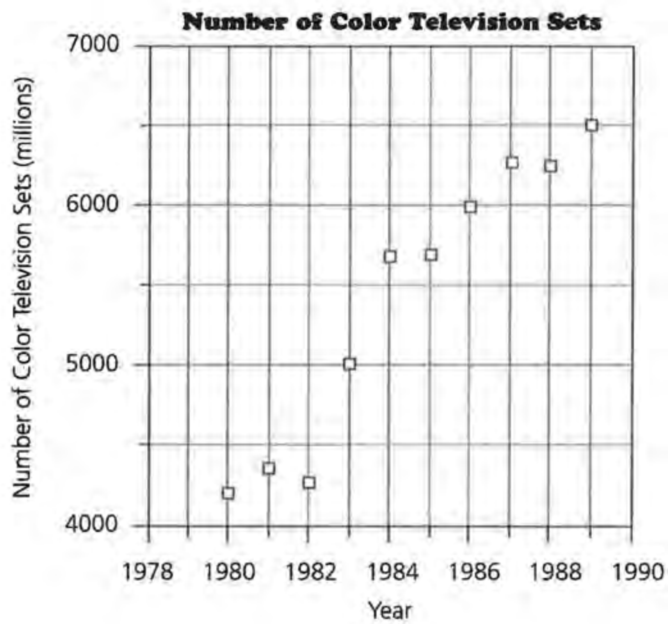
a. Create a scatter plot for any data set that does not already have a graph.

b. Examine each scatter plot and identify linear and non-linear models. You may use a graphing utility to make your scatter plots.

c. In your examination, determine which of the model functions are classified as increasing and which are classified as decreasing.

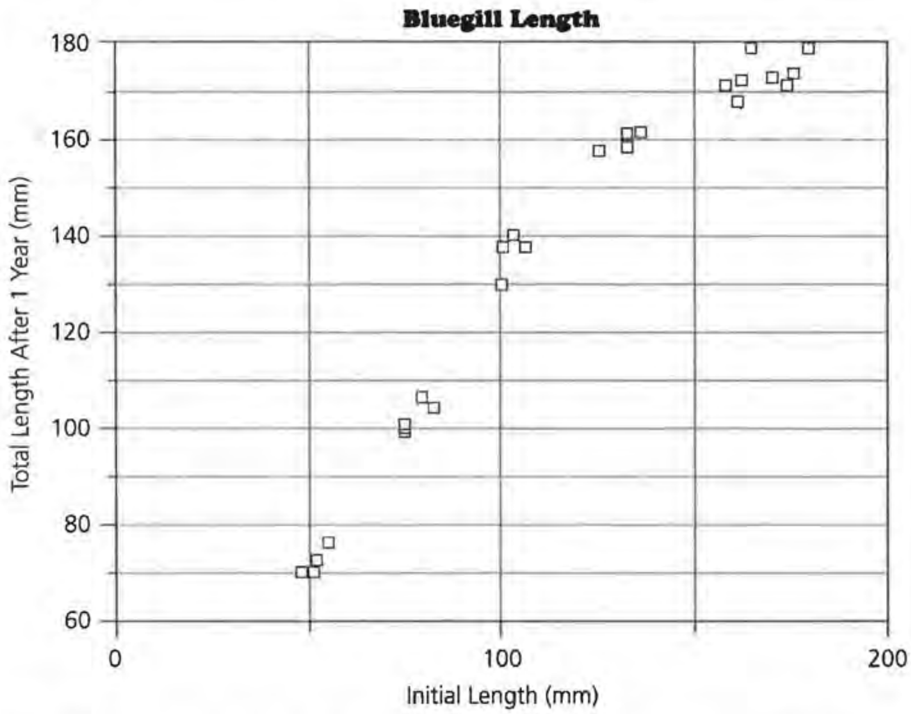
d. After you have described the characteristics, use the information you gained in Lesson 3 to suggest one or two types of functions that could be possible models for that data set.

3.



4.

5.



6. The data below show the amount of suspended particulate matter in millions of metric tons emitted into the air for various years in the United States.

Year	Metric Tons of Suspended Particles in the Air (millions)
1940	22.8
1950	24.5
1960	21.1
1970	18.1
1971	16.7
1972	15.2
1973	14.1
1974	12.4
1975	10.4
1976	9.7
1977	9.1
1978	9.2
1979	9.0
1980	8.5
1981	7.9
1982	7.0
1983	6.7
1984	7.0
1985	7.3

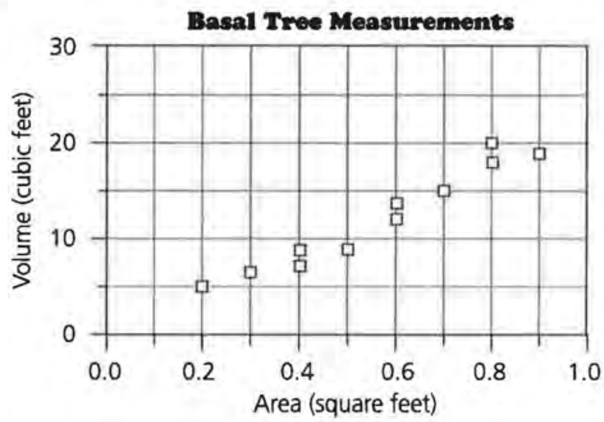
Source: *USA by Numbers, Zero Population Growth, Inc.*, 1988

7.

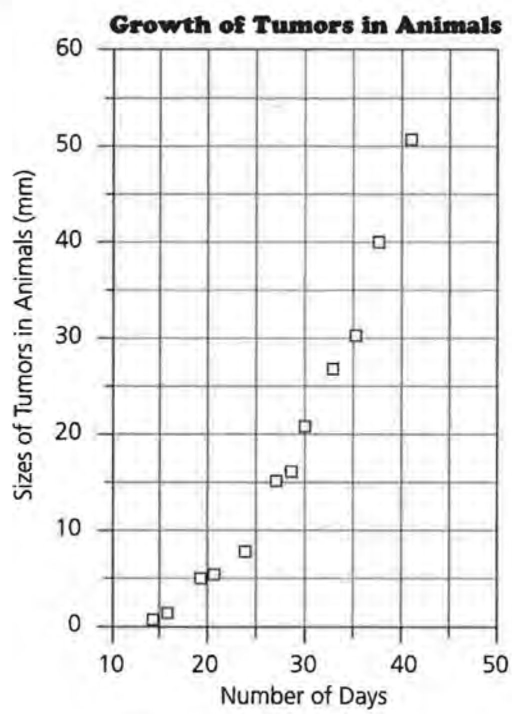
Year	World Silver Production (metric tons)
1986	12,970
1987	14,019
1988	15,484
1989	16,041
1990	16,216
1991	15,692
1992	15,345

Source: *World Almanac*, 1995

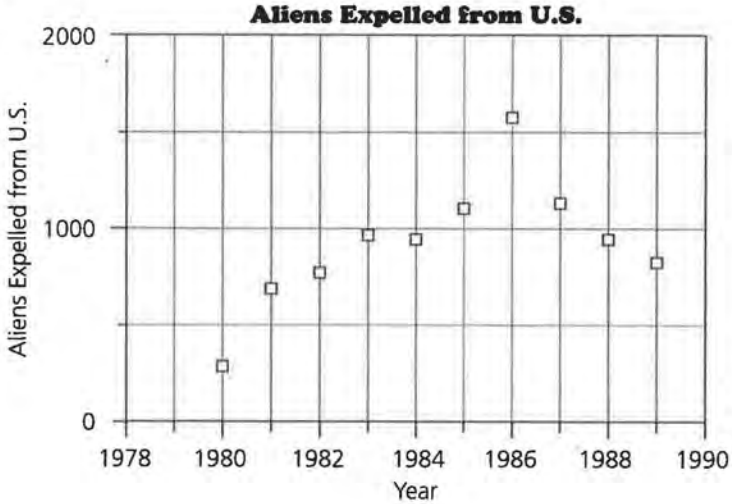
8.



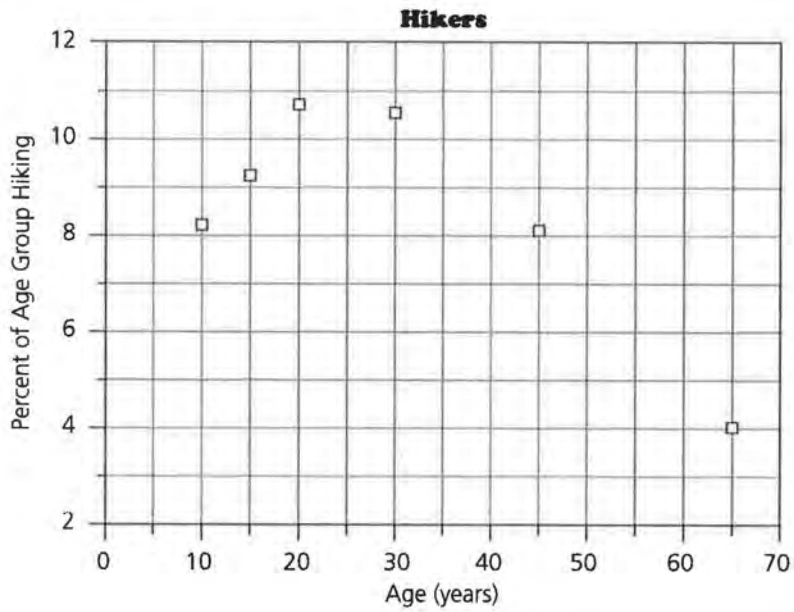
9.



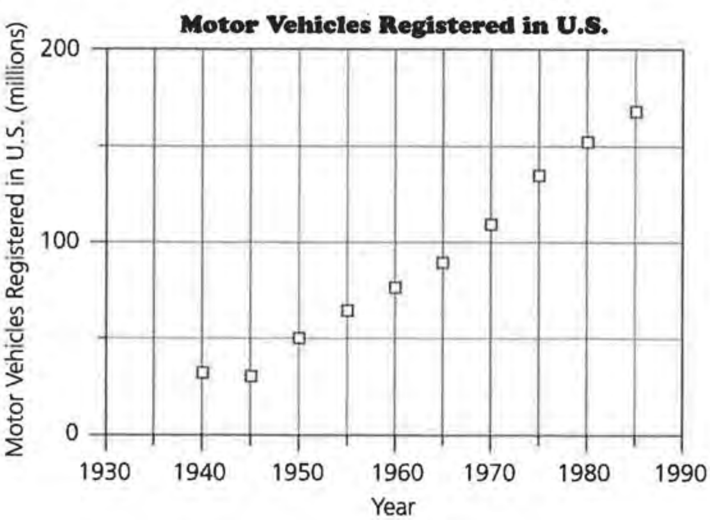
10.


Aliens Expelled from U.S.
Aliens Expelled from U.S.
2000
1000
0
1978
1980
1982
1984
1986
1988
1990
Year

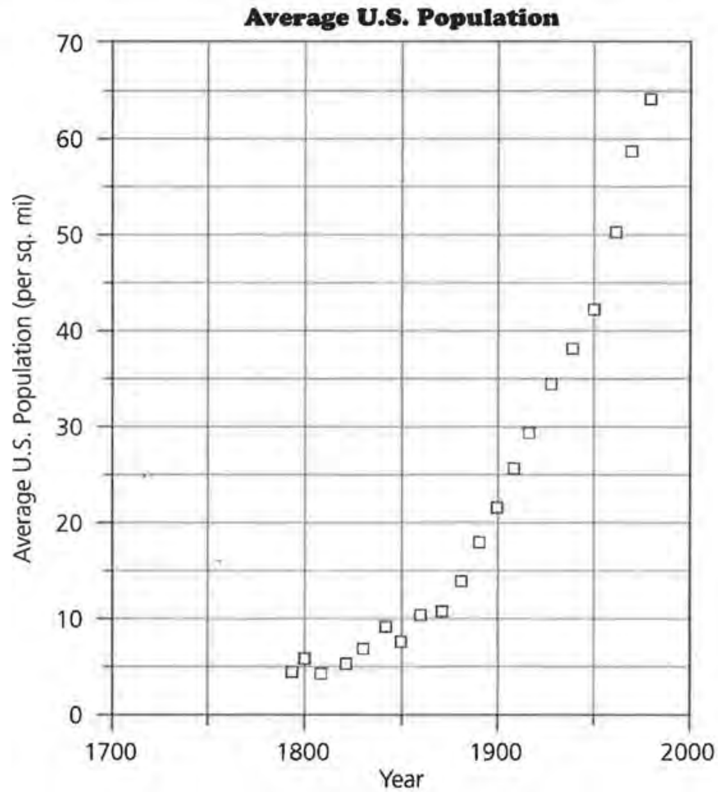
11.


Hikers
Percent of Age Group Hiking
12
10
8
6
4
2
0
10
20
30
40
50
60
70
Age (years)

12.


Motor Vehicles Registered in U.S.
Motor Vehicles Registered in U.S. (millions)
200
100
0
1930
1940
1950
1960
1970
1980
1990
Year

13.



14.

Year	Cumulative Numbers of U.S. Stamps Issued
1868	88
1878	181
1888	218
1898	293
1908	341
1918	529
1928	647
1938	838
1948	980
1958	1123
1968	1364
1978	1769
1988	2400

Source: *Scotts Standard Postage Stamp Catalog,* 1989

LESSON 5

Transforming Data

What are some things that will affect the shape of a graph?

What effect would transforming data have on the graph of the data?

As you know, *pi* (π) is the ratio of the circumference of a circle to its diameter, or 3.14159.... It is used in formulas that describe circular figures.

OBJECTIVE

Transform specific data sets to make them appear linear when plotted in a scatter plot.

INVESTIGATE

In the Old Testament of the *Bible* (II Chronicles 4:2), it is stated, "Then he made the molten sea; it was round, ten cubits from brim to brim, and five cubits high, and a line of thirty cubits measured its circumference." The circumference was, therefore, 6 times the radius or 3 times the diameter. The Hebrews used 3 for π. The Egyptians used $\sqrt{10}$ or 3.16. Which value of π do you commonly use?

Discussion and Practice

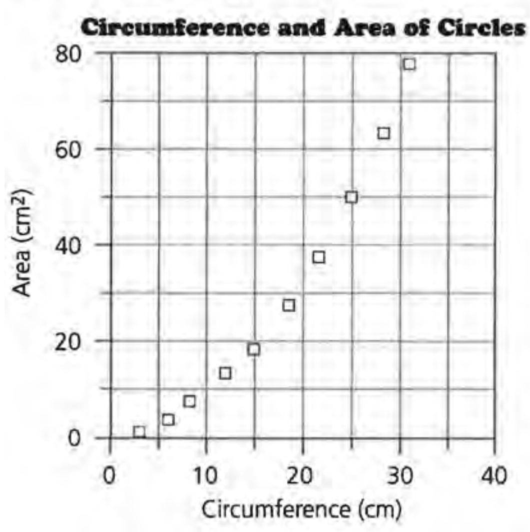
1. Use $\pi = 3.14159$ to complete the table on page 37 about circles.

Diameter (cm)	Area (cm^2)	Circumference (cm)
1	______	______
2	______	______
3	______	______
4	______	______
5	______	______
6	______	______
7	______	______
8	______	______
9	______	______
10	______	______

a. Make a scatter plot of (diameter, area).

b. Make a scatter plot of (diameter, circumference).

2. The data from the table above were used to make the (circumference, area) scatter plot below. Compare the (circumference, area) scatter plot to your (diameter, area) and (diameter, circumference) scatter plots. How are they alike or different? Explain.



In previous lessons, we analyzed the changes that occurred in graphs when the unit measure of either or both of the variables was changed.

3. Explain in a short paragraph what effect a unit change can have on a graph.

In this lesson, you will analyze the changes on a graph produced by a *transformation* performed on a variable. A transformation on a variable is accomplished by using a function to change its value. You are familiar with the effect of unit changes on a graph. For example, changing from quarters to dollars or changing from people to thousands of people are examples of unit changes. Such changes can expand or contract a graph, but the distances between tick marks on the scale remain uniform.

Some situations force us to consider other kinds of scale changes. For example, consider the data of the intensity of sound (number of times as loud as the softest sound) in decibels.

Sample Sound	Decibels	Number of Times as Loud as Softest Sound
Jet airplane	140 decibels	100,000,000,000,000
Air raid siren	130 decibels	10,000,000,000,000
Pneumatic hammer	120 decibels	1,000,000,000,000
Bass drum	110 decibels	100,000,000,000
Thunderclap	100 decibels	10,000,000,000
Niagara Falls	90 decibels	1,000,000,000
Loud radio	80 decibels	100,000,000
Busy street	70 decibels	10,000,000
Hotel lobby	60 decibels	1,000,000
Quiet automobile	50 decibels	100,000
Average residence	40 decibels	10,000
Average whisper	30 decibels	1,000
Faint whisper	20 decibels	100
Rustling leaves or paper	10 decibels	10
Softest sound heard	0 decibels	1

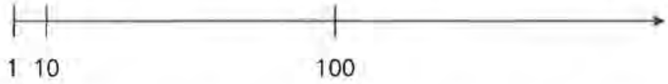
Source: *Definitions of Integrated Circuits, Logic, and Microelectronics Terms*

4. Use the data above to make a graph of (number of times as loud as softest sound, decibels). Determine your own scale and label the axes.

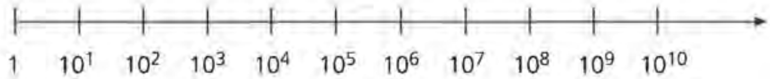
5. What problems did you have in choosing the scale for the graph in Question 4?

What suggestions do you have for changing the scale to make the graph easier to read?

To put all of the points on the graph, the graph had to range from 0 to 100,000,000,000,000 on the horizontal axis. This is impractical because of the relative size of the largest and smallest intervals in the x-range. On the first number line on page 39, there are 9 units between 1 and 10, and 90 units between 10 and 100. The distance between the x-values 1 and 10, 10 and 100, 100 and 1000, and so on, increases as x increases.

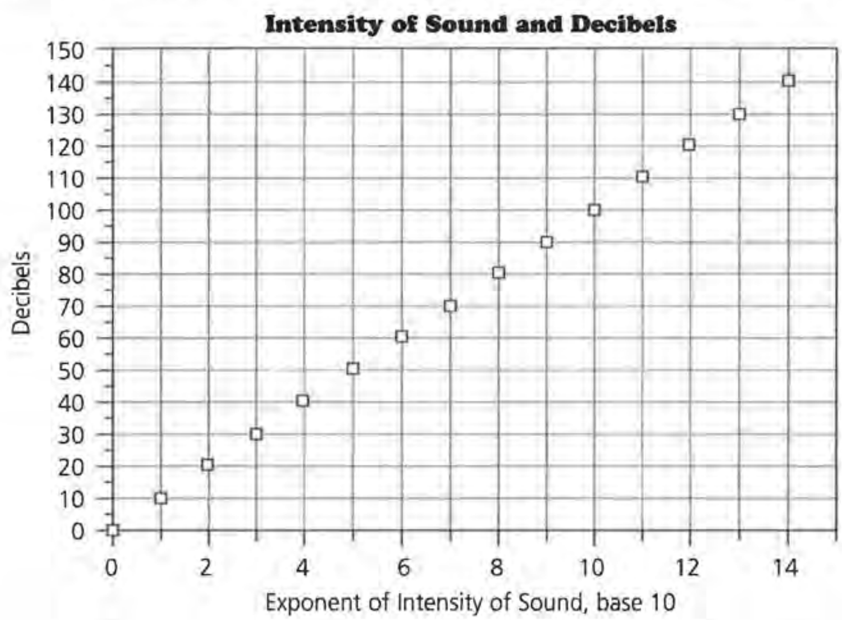


It is reasonable to consider a change that would compress the largest intervals in the x-range without also compressing the smallest intervals. One way to preserve the values of the points plotted and create a scale in which the horizontal values are the same distance apart is to use a scale containing the numbers represented in exponential form. Consider the following scale:



This type of scale change, in which the distances between points appear equal but actually represent different values, is an example of a transformation.

When the x-axis is transformed, the graph of (exponent of intensity of sound, decibels) changes as shown in this graph.



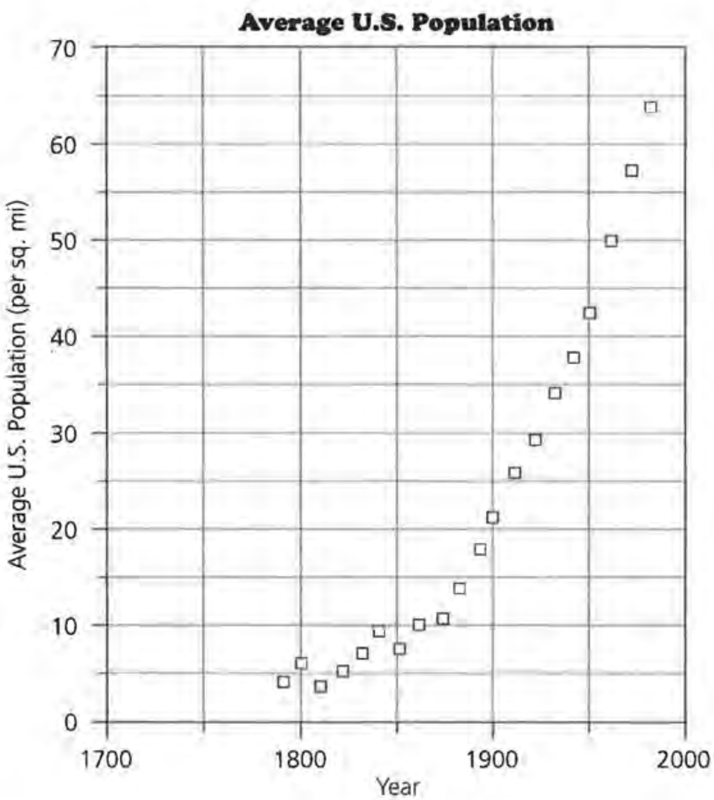
To enable us to graph the (exponent of intensity of sound, decibels) ordered pairs and compensate for the rapid increase in horizontal values, each intensity-of-sound value has been mathematically transformed.

6. Write a paragraph describing the changes you observe.

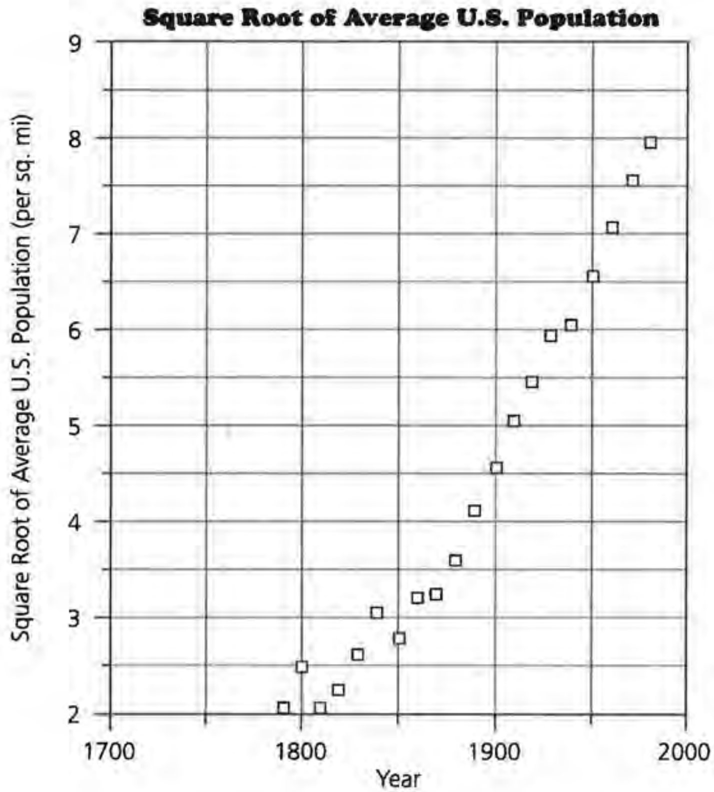
7. Write an equation to describe decibels in terms of the exponent, base 10, of the intensity of sound.

Straightening a graph through transformation is often an advantage to us, because it is much simpler to create the equation of a straight line than that of a curve. When we are able to ***linearize*** a data set by performing a transformation on the units, we can fit a straight line and create the equation of that line. The line or its equation can be used to describe patterns and make predictions about the data set.

Here is an example of how the transformation of units can linearize a graph.



Transform the data on the vertical axis above by taking the square root of each unit. The results are plotted using the new values on the vertical axis in the graph on page 41.



You will notice a straightening effect, or linearization, of the curve created by the transformation.

8. On the grid on *Activity Sheet 1,* draw a straight line that seems to be a good fit for the previous graph that models the relationship between the square root of the average population per square mile and the year.

9. Write an equation for the line of best fit for the square root of average population per square mile in terms of year.

Summary

In this lesson, you investigated the transformation of the units on one of the axes. A transformation of one axis is accomplished by applying a function to the x- or y-values. The results of a transformation are:

(1) The distances between the data points are changed. For example, in the (exponent of intensity of sound, decibels) graph in this lesson, the horizontal distances between the data points increase as x increases.

(2) A transformation alters the appearance of the graph of the data, often changing its shape.

We will now investigate specific methods of effecting changes in graphs.

Practice and Applications

10. Use the data table at the beginning of this lesson.

- **a.** Make a scatter plot of (circumference, area).
- **b.** Add a new column to the data table for circumference squared.
- **c.** Make a scatter plot of (circumference squared, area).
- **d.** Describe how the graph in part c is different from the graph in part a.

11. Write an equation for area in terms of the circumference squared.

LESSON 6

Exploring Changes on Graphs

What effect does changing the scale on an axis have on the graph of the data?

Is it true that a transformation using a function's inverse will linearize that function's graph?

INVESTIGATE

Soccer balls come in different sizes, which are indicated by numbers.

How do volume and circumference compare?

Size	Circumference (cm)	Volume (cm^3)	Age of Player
3	59	3468.2	under 8
4	62.8	4182.4	8–11 years
5	67	5078.9	12 years and older

Other sports also have balls with standardized circumferences.

Ball	Circumference (cm)	Volume (cm^3)
Soccer, size 3	59	3468.2
Soccer, size 4	62.8	4182.4
Soccer, size 5	67	5078.9
Softball	33	606.9
Basketball	82.5	9482.2
Golf Ball	13.5	41.5
Playground Ball	69	5547.5
Racquetball	18	98.5
Tennis Ball	20.2	139.2
Baseball	23.3	213.6

OBJECTIVE

Recognize and understand how the shape of a graph changes when a variable plotted in the graph is transformed.

Discussion and Practice

Use the data on page 43 to make five scatter plots: (circumference, volume), (circumference squared, volume), (circumference, square root of volume), (circumference cubed, volume), and (circumference, cube root of volume) on the grids provided on *Activity Sheets 2–4*. Use the graphs to answer the following questions.

1. Compare and then explain how transformations change the appearance of the graph.
2. Which graph(s) are easier to describe algebraically? Explain.
3. Compare the first graph to the fourth graph. How did the transformation of circumference into circumference cubed change that graph? How did the transformation of volume into cube root of volume in the last graph change the first graph?
4. Make a conjecture about your findings in Question 3.
5. Why do you think cube and cube-root transformations might be preferred over the square and square-root transformations in this example?
6. Classify each of the above graphs as linear or nonlinear and increasing or decreasing.
7. Measure the circumference of a beach ball in centimeters and use one of the graphs to predict the volume.
   a. Which graph did you select for this purpose? Why?
   b. Compute the volume of the beach ball and determine how close to the actual volume your prediction came.
8. Write an equation to represent each of following relations.
   a. The cube root of the volume of a sphere in terms of its circumference
   b. The volume of a sphere in terms of its circumference cubed

Summary

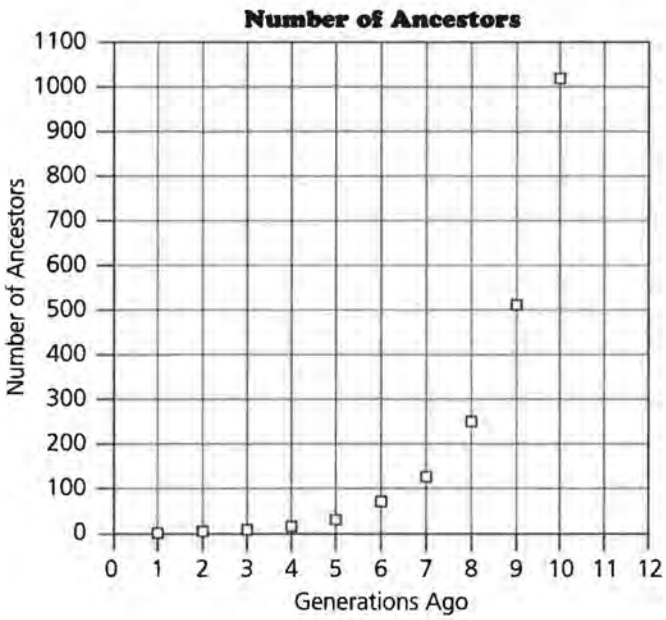
The association between a graph's shape and the scale on either axis is another important relationship in the process of modeling. In this unit, you investigated that relationship and became aware of the choice of the inverse of a function to effect the straightening of the curve.

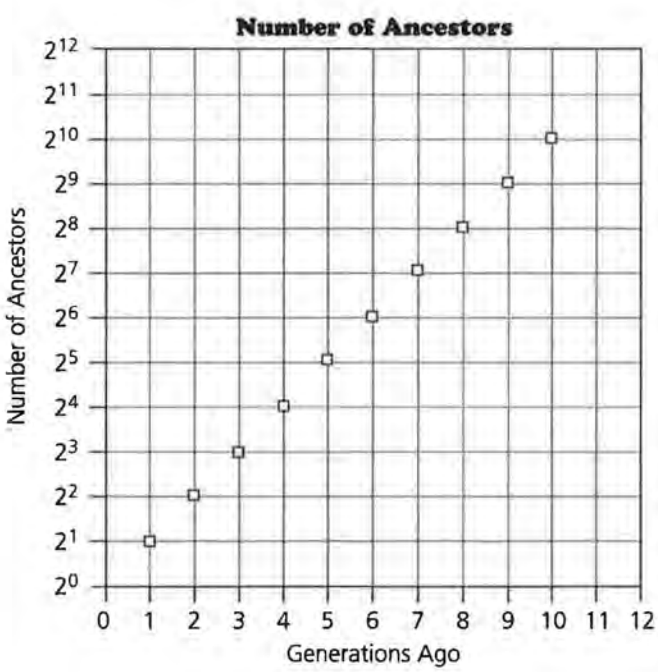
Practice and Applications

Ancestors Problem

In the ancestors problem of Lesson 1, you made a scatter plot of (number of ancestors, generations ago).

9. Study the following two graphs of the ancestors problem.
   a. Make a list of the ordered pairs graphed in each one.
   b. Explain how the graphs are alike and how they are different.





10. Write an equation representing the number of ancestors in terms of the exponent of the generations ago.

Notice that although the physical distances between the vertical tick marks, 2^1 and 2^2, and 2^2 and 2^3, and so on, appear equal, the numerical distances between them are not equal.

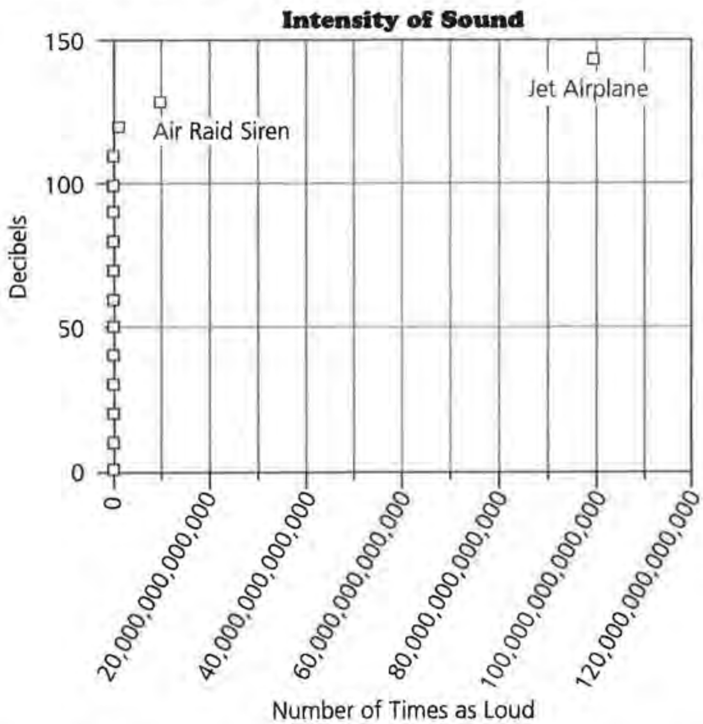
11. Were the x-values or the y-values transformed to create the second graph? Describe the effect of this transformation on the graph.

The decibel data from Lesson 5, repeated here, has been used to create the two graphs on the next page.

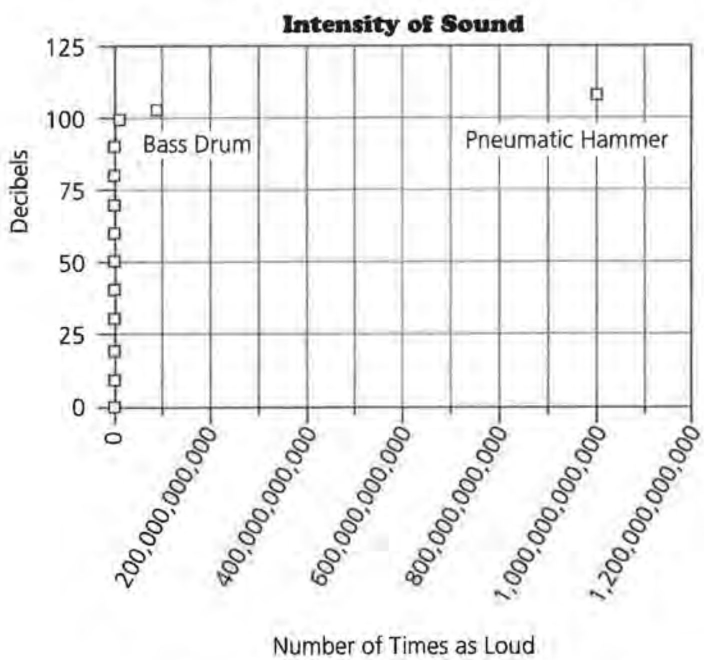
Sample Sound	Decibels	Number of Times as Loud as Softest Sound
Jet airplane	140 decibels	100,000,000,000,000
Air raid siren	130 decibels	10,000,000,000,000
Pneumatic hammer	120 decibels	1,000,000,000,000
Bass drum	110 decibels	100,000,000,000
Thunder clap	100 decibels	10,000,000,000
Niagara Falls	90 decibels	1,000,000,000
Loud radio	80 decibels	100,000,000
Busy street	70 decibels	10,000,000
Hotel lobby	60 decibels	1,000,000
Quiet automobile	50 decibels	100,000
Average residence	40 decibels	10,000
Average whisper	30 decibels	1,000
Faint whisper	20 decibels	100
Rustling leaves or paper	10 decibels	10
Softest sound heard	0 decibels	1

Source: *Definitions of Integrated Circuits, Logic, and Microelectronics Terms*

Study this graph of the decibel data.



12. The plot above looks linear except for two points. Consider the following plot with the data having a vertical scale from 0 to 125 decibels to eliminate the two highest data points.



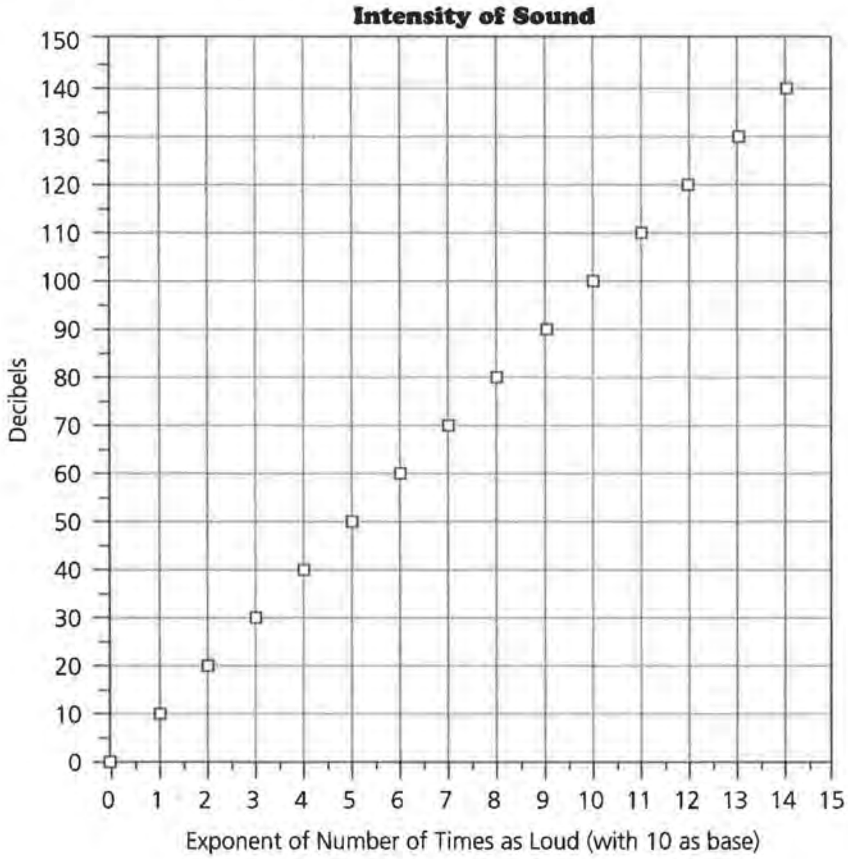
a. How has the horizontal scale changed?

b. How does the appearance of the graph change?

c. Is the graph easier to read?

13. Describe what happens when you draw the graph with a vertical scale from 0 to 100 and eliminate two additional points.

14. Does the following graph accurately describe the data in the decibels table? Justify your answer.



15. Use the graph above. For a decibel level of 80, what is the intensity of sound?

Stopping Distances

Driving students are usually taught to allow one car length, or about 15 feet, between their car and the next car for every ten miles of speed under normal driving conditions and a greater distance in adverse weather or road conditions. The faster a car is traveling, the longer it takes the driver to stop the car. The stopping distance depends on the driver-reaction distance and the braking distance. The total stopping distance is equal to the sum of the distance the car travels in the time it takes the driver to react and the distance the car travels after the brakes are applied.

Speed (mph)	Average Total Stopping Distance (ft)
20	42
25	56
30	73.5
35	91.5
40	116
45	142.5
50	173
55	209.5
60	248
65	292.5
70	343
75	401
80	464

Source: U.S. Bureau of Public Roads

1. Graph (speed, average total stopping distance) using the data from the table. Tell which function family you think your graph belongs to. Justify your answer.
2. Graph each of the following and then write an equation for the graph.
   a. (speed squared, average total stopping distance)
   b. (speed cubed, average total stopping distance)
   c. (speed, the square root of average total stopping distance)

The number of motor vehicles registered every 5 years in the United States has increased since 1945.

Year	Motor Vehicles Registered in U.S. (millions)
1940	32.4
1945	31.0
1950	49.2
1955	62.7
1960	73.9
1965	90.4
1970	108.4
1975	132.9
1980	155.8
1985	170.2

Source: *Moore & McCabe*

3. Plot (year, motor vehicles registered in the U.S.).
   a. Draw the straight line of best fit through the data plot. Then find an equation for this line.
   b. The number of vehicles registered in 1945 does not fit the pattern. Explain why.
4. Identify a transformation that would straighten the curve. Create the table and plot the graph.
   a. Draw a straight line on the graph and find its equation.
   b. Which line do you think is better? Why?
5. Use the line you chose to predict how many motor vehicles will be registered in the year 2000.

Unit III

Mathematical Models from Data

LESSON 7

Transforming Data Using Logarithms

What is a logarithm?

How does a logarithmic transformation change the appearance of a graph?

OBJECTIVE

Recognize how transforming either scale of a graph with the logarithmic function changes the shape of the graph.

Units of measure such as meter, centimeter, foot, and inch increase consistently by 1s as they get larger. Therefore, a board that is 10 feet long is twice as long as a board that is 5 feet long. A meter stick is 10 times as long as a 10-cm ruler.

However, the number of ancestors in each generation does not increase steadily by 1s as you consider previous generations. Rather, the number increases by multiples of 2. Two generations ago, you had 4, or 2^2, grandparents. Three generations ago, you had 8, or 2^3, great-grandparents. Four generations ago, you had 16, or 2^4, great-great-grandparents.

INVESTIGATE

When numbers increase or decrease rapidly, it may be convenient to look for patterns in their exponents, or ***logarithms.***

$\log_{10}10 = 1$	The log of 10 equals 1 because $10^1 = 10$ and the log is the exponent.
$\log_{10}100 = 2$	The log of 100 equals 2 because $10^2 = 100$ and the log is the exponent.
$\log_{10}1{,}000 = 3$	The log of 1,000 equals 3 because $10^3 = 1{,}000$ and the log is the exponent.

Discussion and Practice

1. What is $\log_{10}1{,}000{,}000$?
2. What is $\log_{10}1$? Why?
3. Complete the table.

Number	Power of 10	Log_{10} (Number)
100,000,000,000,000	10^{14}	______
10,000,000,000,000	______	______
1,000,000,000,000	______	______
100,000,000,000	______	______
10,000,000,000	______	______
1,000,000,000	______	______
100,000,000	______	______
10,000,000	______	______
1,000,000	______	______
100,000	______	______
10,000	______	______
1,000	______	______
100	______	______
10	______	______
1	______	______

4. What is $\log_{10}0.1$? Explain how you got your answer.
5. Use the table. Between what two numbers is $\log_{10}50$?
6. Use the $\boxed{\text{LOG}}$ button on your calculator to find $\log_{10}50$.
7. Between what two numbers is $\log_{10}5$? Use $\log_{10}50$ to find $\log_{10}5$.
8. Between what two numbers is $\log_{10}500$? Use $\log_{10}50$ to find $\log_{10}500$.
9. Is it possible for 10^n to be a negative number? Is it possible to find the $\log_{10}x$ if $x < 0$? Justify your answer.

Other bases may be used for logarithms. For example, in the ancestors problem, it is helpful to use base 2.

$\log_2 2 = 1$	The log base 2 of 2 equals 1.
$\log_2 4 = 2$	The log base 2 of 4 equals 2.
$\log_2 8 = 3$	The log base 2 of 8 equals 3.

10. What is $\log_2 1$?

11. What is $\log_2 0.5$?

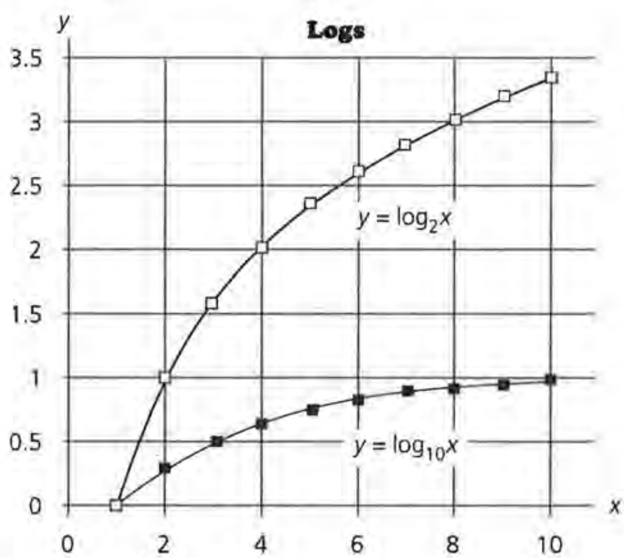
12. What is $\log_2 64$?

13. Use the graph below to find

a. $\log_{10} 1$ and $\log_2 1$.

b. $\log_{10} 20$ and $\log_2 20$.

c. $\log_{10} 100$ and $\log_2 100$.



Summary

The *logarithm* of a number is simply an exponent. The equations $b^y = x$ and $y = \log_b x$ are equivalent. In these equations, y is the logarithm, and b is the base. Numbers may be compared by looking at their logarithms. Logarithms are often considered useful when numbers have magnitudes that are very great or very small.

Practice and Applications

14. Complete the rows on this table for each year. The logarithm of the debt can be abbreviated log (debt) when the base 10 logarithm is used.

Year	Federal Debt	Federal Debt	Log (Debt)
1980	$909,100,000,000	9.091×10^{11}	11.9586
1981	$994,900,000,000	______	______
1982	$1,137,000,000,000	______	______
1983	$1,372,000,000,000	______	______
1984	$1,565,000,000,000	______	______
1985	$1,818,000,000,000	______	______
1986	$2,121,000,000,000	______	______
1987	$2,346,000,000,000	______	______
1988	$2,601,000,000,000	______	______
1989	$2,868,000,000,000	______	______
1990	$3,207,000,000,000	______	______
1991	$3,598,000,000,000	______	______
1992	$4,002,000,000,000	______	______
1993	$4,351,000,000,000	______	______
1994	$4,644,000,000,000	______	______
1995	$4,921,000,000,000	______	______

The following table shows the cumulative number of different kinds of U.S. postage stamps issued, by 10-year intervals. This number includes only regular and commemorative issues and excludes such items as airmail stamps, special-delivery stamps, and postal cards.

Year	Cumulative Number of Kinds of U.S. Stamps Issued	Log (Number of Stamps Issued)
1868	88	______
1878	181	______
1888	218	______
1898	293	______
1908	341	______
1918	529	______
1928	647	______
1938	838	______
1948	980	______
1958	1123	______
1968	1364	______
1978	1769	______
1988	2400	______

Source: *Scotts Standard Postage Stamp Catalog*, 1989

15. Find the log of the cumulative number of kinds of stamps for each year indicated.

16. Graph (year, cumulative number of kinds of stamps issued). Identify whether a linear or non-linear model best describes these data.

17. Graph (year, log(number of stamps issued)). Identify whether a linear or nonlinear model best describes these data.

18. Describe the change that occurs on the graph when each y-value is transformed to log y.

19. Draw a line to model the (year, log(number of stamps issued)) graph, and write the equation of your line.

Finding an Equation for Nonlinear Data

How can the application of an inverse function be used to find the model of the original data set?

What physical phenomenon can be modeled using an exponential function?

OBJECTIVE

Find the equation of a nonlinear data set using transformations.

Half-life is the amount of time it takes half of a substance to decay. Radioactive materials and other substances are characterized by their rates of decay, or decrease, and are rated in terms of their half-lives. The half-life of Carbon-14 allows scientists to date fossils. The half-life of a radioactive prescription substance is used to determine how frequently doses should be given.

INVESTIGATE

Half-life can be simulated by the following experiment.

Equipment Needed: cup to shake coins; 100–200 pennies or other coins

Procedure: Put the coins into the cup. Shake the coins and pour them out. Remove all coins that land heads up. Record the number of coins removed and the number remaining on a chart like the one on page 59. Repeat this procedure until the number of coins remaining is 3, 2, or 1.

Shake Number	Number of Coins Removed	Number of Coins Remaining
0	0	original number
1	______	______
2	______	______
3	______	______
4	______	______
5	______	______
6	______	______
7	______	______
8	______	______

Discussion and Practice

1. The half-life of Carbon-14 is 5730 years. How long would it take for one gram of Carbon-14 to decay to $\frac{1}{16}$ of a gram?
2. The volume of a mothball, a small ball of napthalene used as a moth repellent, decreases by about 20% each week. What is the half-life of a mothball?
3. Perform the half-life simulation and record your results.
4. Graph the (shakes, number of coins remaining) ordered pairs from your experiment.
   a. Describe the pattern on the graph.
   b. Does the graph cross the x-axis? Where? What is the meaning of that point?
5. Approximately what percent of the remaining pennies were removed on each shake? Why? How many shakes represent a half-life for the pennies?

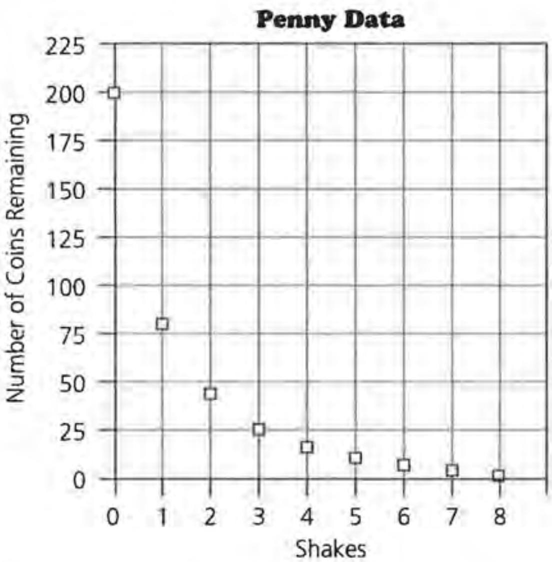
In order to decide what transformation best linearizes a curve, you may have to try more than one transformation.

6. Graph (shakes, square root of number of coins remaining). Draw the straight line that seems to best fit the graph.

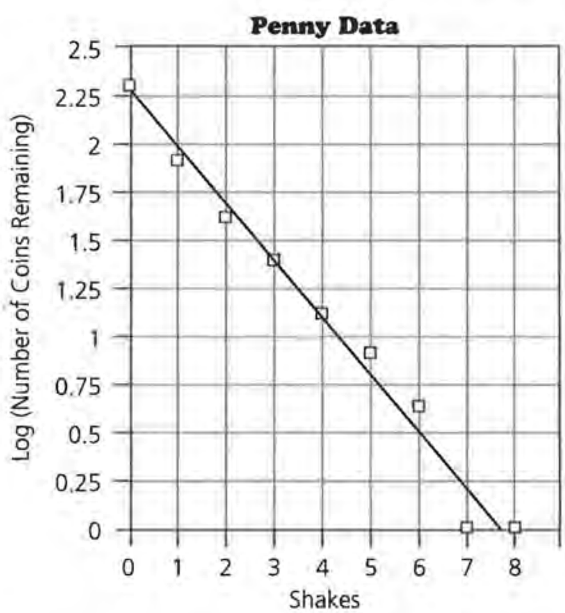
7. Graph (shakes, log(number of coins remaining)). Draw a straight line that seems to fit the graph best.

8. Compare the lines you drew for Questions 6 and 7. Which one do you think is better? Why?

The half-life simulation was performed by two students. Their data are graphed below.



The data were then transformed and a regression line was drawn as shown on the following graph.



The equation of the regression line in $y = ax + b$ form was $y = -0.298x + 2.275$.

The following example shows how this linear equation can be used to find the equation of the curve formed by the original data set. Study the example and make sure you understand what is happening.

Example: Using the Equation of a Regression Line to Find the Equation of the Original Data Set

Let the x- and y-variables in the ordered pairs (x, y) represent points on the curve that contain the original data points. The equation of the regression line of the transformed data in the form $y = ax + b$ is written

$$\log_{10} y = ax + b,$$

since the transformation used $\log_{10} y$.

In this case, $a = -0.298$ and $b = 2.275$. For this regression line, a is the slope and b is the y-intercept.

Therefore, $\log y = -0.298x + 2.275$.

Applying the definition that $\log_b x = y$ implies $b^y = x$, it follows that:

$$y = 10^{(-0.298x + 2.275)}$$
$$y = (10^{-0.298x})(10^{2.275})$$
$$y = (10^{-0.298})^x\,(10^{2.275})$$
$$y = (0.504^x)(188.36)$$
$$y = 188.36(0.504^x)$$

9. This is a mathematical model for the data set shown in the first graph of Penny Data on page 60.

   a. What does 188.36 represent?

   b. What does 0.504 represent?

10. Find the equation of a mathematical model for the data you collected in your experiment.

Summary

Mathematical models are often used to answer questions or study trends. The process used in finding the equation may be summarized in the following steps.

a. Collect and graph the data.

b. Observe patterns in the data.

c. If necessary, transform the data to straighten it. You may need to try more than one transformation.

d. Plot the transformed data and draw a linear regression line.

e. Use the equation of the regression line to find an equation of the original data set.

When a logarithmic transformation straightens a function, the function is an exponential function.

Practice and Applications

In Questions 11 and 12, transform the data, if necessary, to find a linear model. Then use the linear model and inverse functions to find an equation for the original data.

11.

Time (seconds)	Height of Bounce for Ball Dropped from 2.09 m (meters)
0	2.09
1.2	1.52
2.2	1.2
3	0.94
3.6	0.76
4.2	0.63
4.7	0.57
5.2	0.52
5.5	0.49
5.9	0.45
6.2	0.44
6.5	0.42

Source: Physics Class, Nicolet High School

12.

Distance for Metal Ball on Small Ramp at a Given Angle (cm)	Time (seconds)
15.0	0.61
30.0	0.95
45.0	1.24
60.0	1.35
75.0	1.65
90.0	1.77
105.0	1.91
120.0	2.02

Source: Physics Class, Mahivah High School

Residuals

What are residuals?

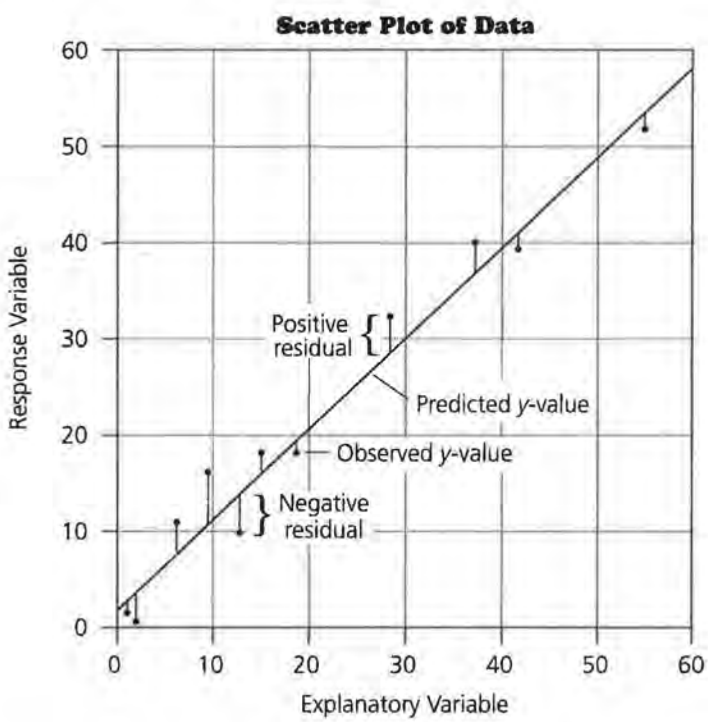
What do residuals reveal about the model?

OBJECTIVE

Use plots of residuals to help assess how well a mathematical model fits the data.

In many problems, you have been fitting a straight line to data in a scatter plot and evaluating the results to see if the line fits well. You have judged whether or not a fit was adequate mainly by the "eyeball" method: looking at the scatter plot, looking at the straight line in relation to the data points, and checking if the straight line makes sense as a model for the data. This method is important and is a good first step.

A numeric tool available to help in making this evaluation is a *residual.* The corresponding graphic tool is the *residual plot.*



The residual may be calculated for each point in a data set. It is the difference between the actual or observed y-value and the predicted y-value found by using the mathematical model:

residual = observed y-value – predicted y-value from the model

Using symbols, $r_i = y_i - (a + bx_i)$.

INVESTIGATE

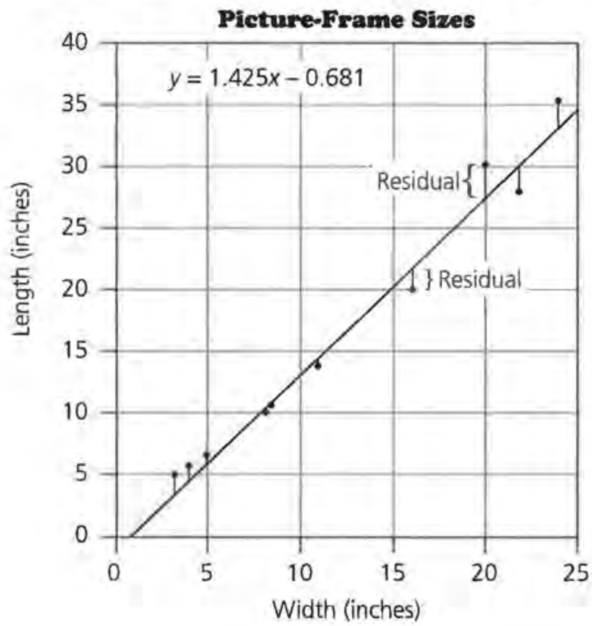
A fit, or predicted value, in the form $a + bx_i$ may be described by the symbol $\hat{y}_i$, read "y-hat sub i"; $\hat{y}_i = a + bx_i$:

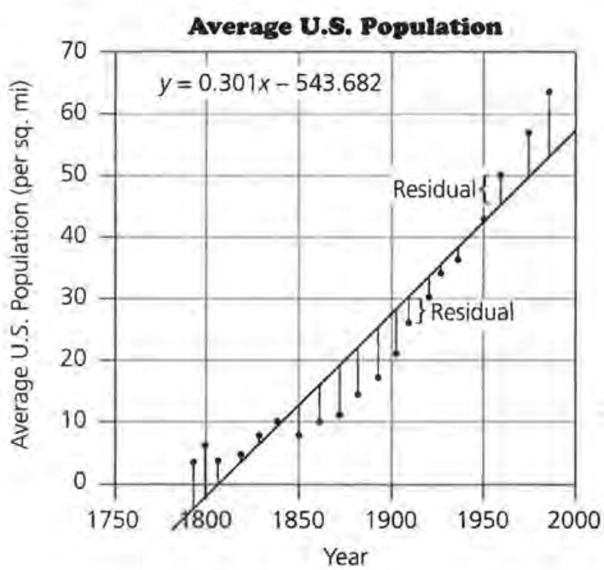
$$\text{residual}_i = \text{res}_i = y_i - \hat{y}_i$$

Also, observed data = fit + residual:

$$y_i = \hat{y}_i + (y_i - \hat{y}_i)$$

Consider the following scatter plots and residuals.

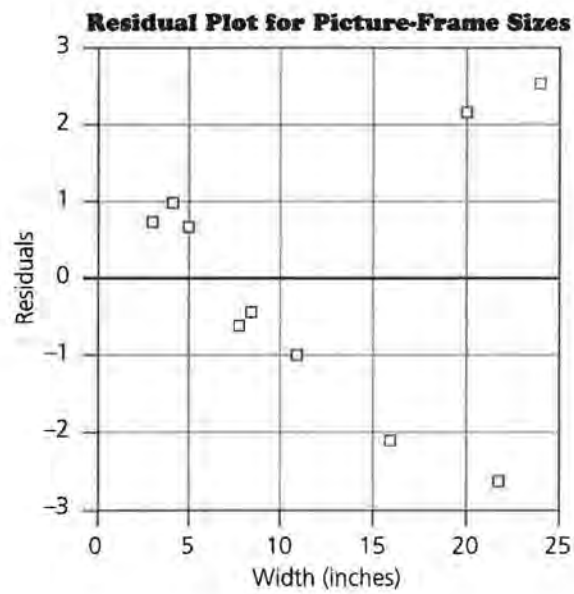


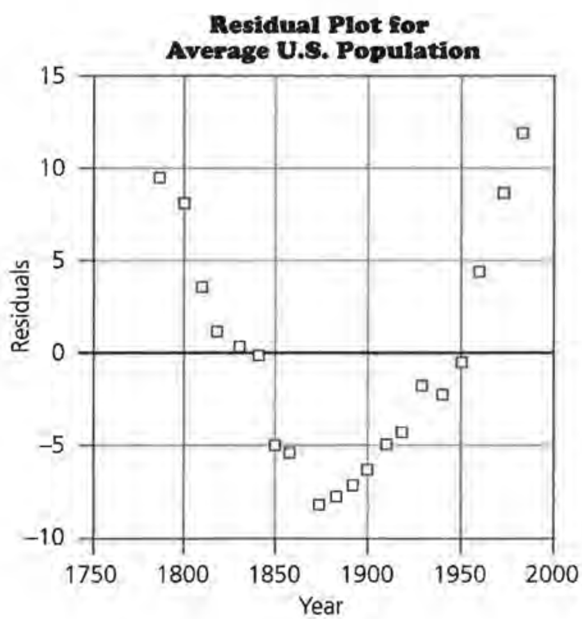


When actual data are collected, a mathematical model generally does not fit them perfectly. It is not expected that the residuals' sum will be exactly zero. The residuals should, however, represent random variation of the data from the fitted model, some above and some below the horizontal line.

While residuals can be observed in an original scatter plot, it is sometimes helpful to make a separate plot of them to see if they form a pattern. You can do this by graphing $(x_i, y_i - \hat{y}_i) = (x_i, \text{res}_i)$, which is called a *residual plot* or plot of residuals against the explanatory variable x. In a good model, the residuals should exhibit a random pattern.

Consider the residual plots for the previous two graphs.





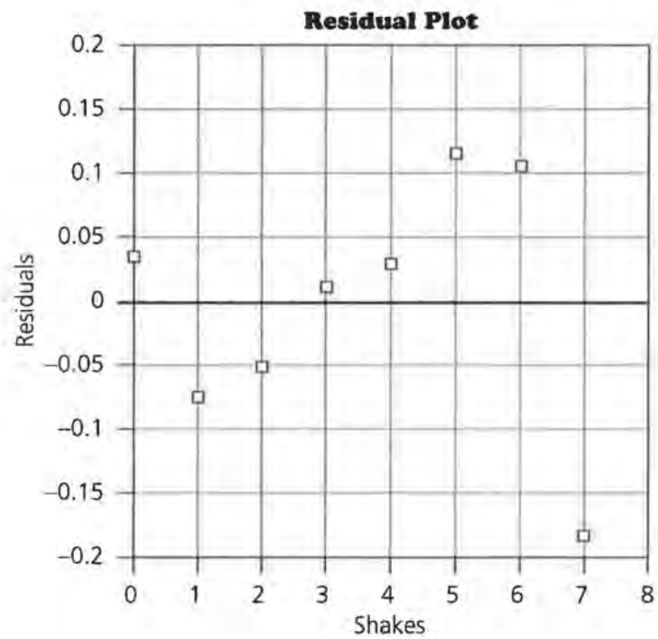
In general, information from the residuals may indicate that the mathematical model is a good fit. This occurs when

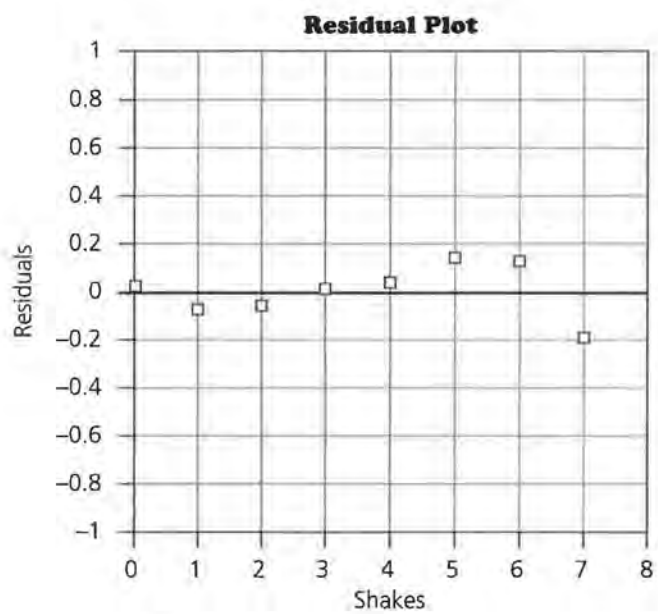
- there are no obvious patterns in the relationship of residuals to the explanatory variable, x;
- there is uniform variability in the relationship of residuals to x_i; and
- there are no individual outlying points.

The residual plot should form an approximately uniform, horizontal band going across the page. This indicates a random pattern to the residuals with no special relationship to the explanatory variable x. The following graph shows an approximate horizontal band of data points somewhat uniform.

If the residuals from a mathematical model show a pattern with an obvious structure, the structure should be incorporated into the "fit" by changing the mathematical model, if possible. This can be done by transforming the data and fitting a new line. Then the residuals for the new model can be calculated, plotted, and observed. This process may continue several times for a data set.

It is also important to consider the scale on the y-axis of the residual plot. The scale reveals the magnitudes of the residuals. The following plots of the same set of residuals for the Penny Data in Lesson 8 look different because the scales on the y-axes are different.





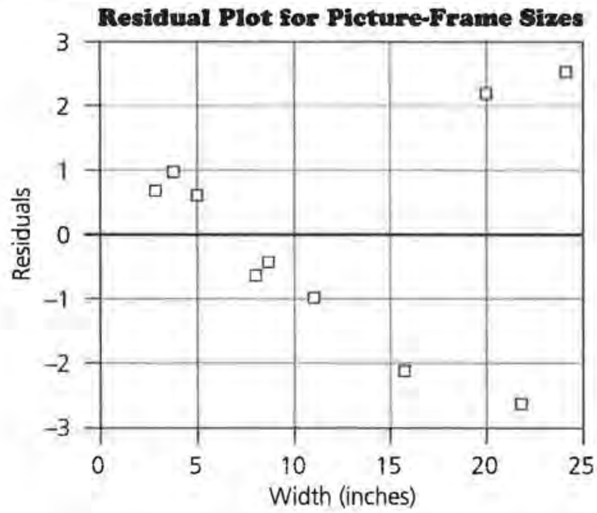
Discussion and Practice

1. Describe the vertical position on the graph of the actual or observed data point with respect to the value predicted by the mathematical model if the residual at the point is

a. positive. **b.** negative. **c.** zero.

2. Describe at least two mathematical ways you could summarize the residuals for a mathematical model on a data set.

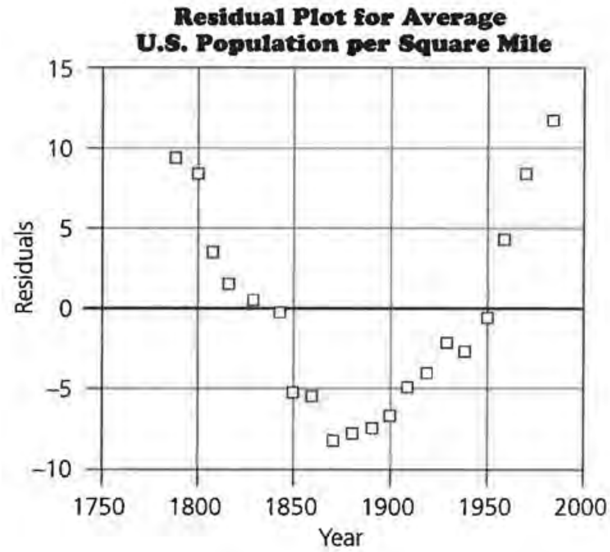
3. What patterns do you observe on the residual plot for Picture-Frame Sizes?

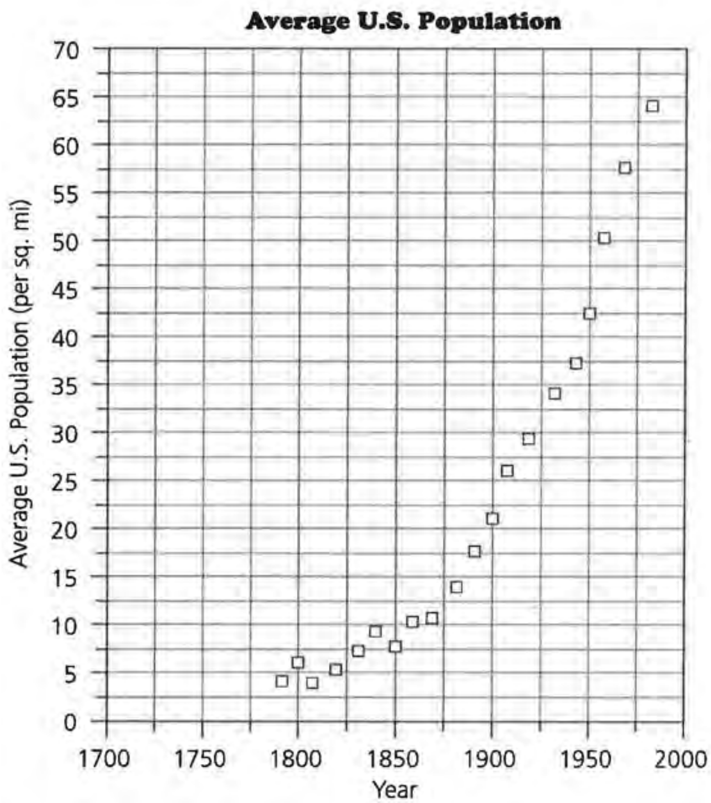


4. You have seen several graphs and the residual plot for the Average U.S. Population. The data set for these plots is given below.

Year	Average U.S. Population (per sq. mi)
1790	4.5
1800	6.1
1810	4.3
1820	5.5
1830	7.4
1840	9.8
1850	7.9
1860	10.6
1870	10.9
1880	14.2
1890	17.8
1900	21.5
1910	26.0
1920	29.9
1930	34.7
1940	37.2
1950	42.6
1960	50.6
1970	57.5
1980	64.0

Source: *World Almanac and Book of Facts,* 1988





a. The residual plot for average U.S. population shows a pattern that may be incorporated into a new model. Apply a transformation to one variable in the data set.

b. Find a mathematical model for the transformed data and make a residual plot.

c. Compare the previously given residual plot with the residual plot created in part b above. Decide which appears to have variation that is more random. Explain.

Summary

A mathematical model can be used to describe a data set. Residuals describe the deviations of the data from the model.

$$\text{residual} = \text{observed value} - \text{predicted value}$$
$$\text{data} = \text{fit} + \text{residual}$$

Residuals may be plotted and studied for patterns. These patterns may help in the search for a good mathematical model because the deviations from the model are emphasized in the residual plot, (x_i, res_i).

Use residuals in the modeling process using the following steps.

- Make a scatter plot of the data.
- Fit a line to the data.
- Calculate the residuals.
- Make a residual plot.
- Study the residual plot to see whether or not it exhibits random variation. If not, there may be a better mathematical model for the data.

Practice and Applications

5. Use the data in the table below to make scatter plots and residual plots for each of the following.

a. Median income of men over time

b. Median income of women over time

c. Ratio of median income of women to median income of men over time

Median Income of Men and Women ($1,000)

Year	Men	Women
1980	19.2	11.6
1981	20.7	12.4
1982	21.6	13.7
1983	22.5	14.5
1984	24.0	15.4
1985	25.0	16.2
1986	25.9	16.8
1987	26.7	17.5
1988	27.3	18.5
1989	28.4	19.6

6. What patterns do you notice in the scatter plots and the residual plots?

7. What observations can you make about median incomes of men and women?

8. Create a residual plot for the (year, log(number of stamps issued)) graph you created for Question 17 in Lesson 7.

Year	Cumulative Number of Kinds of U.S. Stamps Issued	Log (Number of Stamps Issued)
1868	88	______
1878	181	______
1888	218	______
1898	293	______
1908	341	______
1918	529	______
1928	647	______
1938	838	______
1948	980	______
1958	1123	______
1968	1364	______
1978	1769	______
1988	2400	______

Source: *Scotts Standard Postage Stamp Catalog*, 1989

LESSON 10

Correlation: r and r^2

What is correlation?

What additional information do r and r^2 provide regarding the fit of a model?

A numerical component in the analysis of how models fit data is *correlation*. ***Correlation*** is a measure of the strength of a linear relationship, or how tightly the points are packed around a straight line. The ***correlation coefficient*** is represented by the symbol r. The ***square of the correlation coefficient***, represented by r^2, also has a useful interpretation. The correlation coefficient r or its square, r^2, is often included in a statistical analysis with a least-squares line for a set of data. The formula for finding r is generally programmed into a graphing calculator, and a calculator can be used to quickly and easily find r for any pair of variables. On the TI-83, for instance, r can be made to appear on the screen when the least-squares linear regression line is calculated.

The correlation coefficient may be useful in assessing how well a mathematical model fits a set of data. It is important to understand what the correlation coefficient measures, what it does not measure, how it can be interpreted, its properties, and its limitations. These topics will be investigated in this lesson.

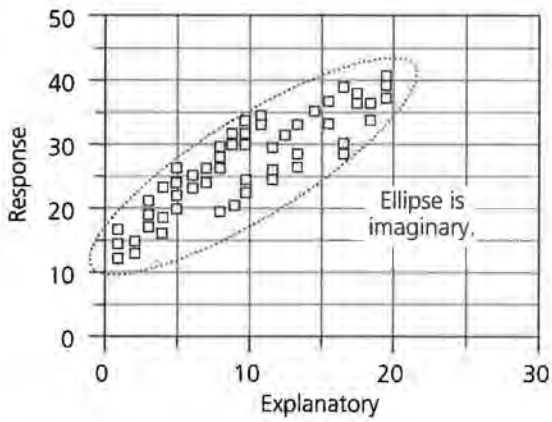
Taken together, residual plots and the correlation coefficient can help assess how well a linear mathematical model fits a data set. These two mathematical tools also help in the selection of an appropriate model for a given data set.

OBJECTIVE

Use the correlation coefficient and the square of the correlation coefficient along with residual plots to help assess how well a mathematical model fits a data set.

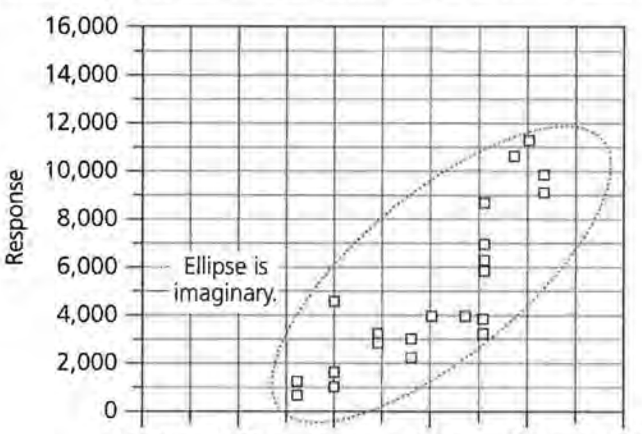
INVESTIGATE

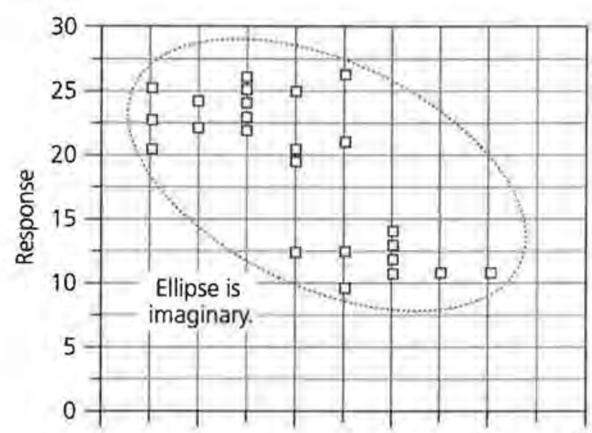
Consider the following scatter plot. If the data points are close to having a uniform spread throughout an imaginary ellipse enclosing all the data points, as in the plot shown below, it is reasonable to use correlation for measuring the association between these variables.



Imagine that you try to fit a line to the data set in the plot above. Do you see how tightly the points would be packed around the line you imagined?

If, however, the graph appears to have large gaps, empty areas, or a noticeable curved shape as in the next two plots, then correlation is not as useful for a measure of association.





Imagine lines through the data sets in these two plots. Do you see how tightly (or loosely) the points would be packed around the lines you imagined?

It is often difficult to estimate the strength of the relationship between the explanatory and response variables from a plot. It is also difficult to meaningfully compare the degree of association in two different plots. A numerical measure of association is therefore useful. The correlation statistic is based on comparing how well y can be predicted when x is known to how well y can be predicted when x is not known. Some general properties of the correlation coefficient are listed below.

Size

- The value of r always falls between -1 and 1. Positive r indicates a positive association between the variables; that is, as x increases, y increases. Negative r represents a negative association; that is, as x increases, y decreases.
- If $r = 0$, there is no linear relationship between the variables.
- The extreme values $r = -1$ and $r = 1$ occur only in the case of perfect linear association, when the points in the scatter plot lie exactly along a straight line.

Units

- The value of r is not changed when the unit of measurement of x, y, or both x and y changes.
- The correlation r has no unit of measurement; it is a dimensionless number.

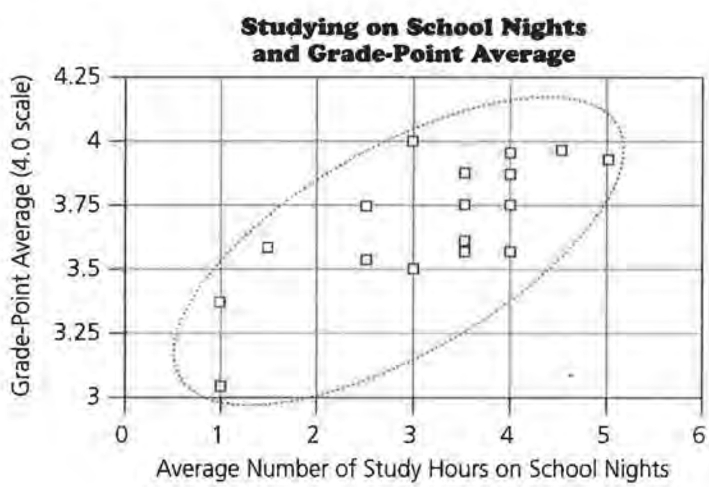
Linear Relation

- Correlation measures only the strength of linear association between two variables.
- Curved relationships between variables, no matter how strong, are not reflected in the correlation.
- The square of the correlation coefficient, r^2, is the proportion of the variation in y that can be explained by the variation in the value of x.

Consider the relationship between grade-point average and the number of hours students study. Grade-point averages may vary from 0.0 to 4.0 on a 4-point scale. The correlation, r, of this data set is about 0.7, so the association is fairly strong.

Number of Study Hours on School Nights	Grade-Point Average
5	3.96
1	3.38
2.5	3.55
1.5	3.6
3.5	3.62
4	3.75
3	3.5
3.5	3.75
4	3.6
4.5	3.99
3	4.0
2.5	3.75
4	3.95
4	3.83
3.5	3.85
1	3.05
3.5	3.6

Source: Precalculus H Class, Nicolet High School



Since $(0.7)^2 = 0.49$, $r^2 = 0.49$. This means that 49% of the variability in grade-point averages can be explained by a linear relationship with how much the students study. However, 51% of the variability is unexplained or is due to other factors such as difficulty of classes, amount and quality of homework, and so on.

Discussion and Practice

1. In general, what are the possible values for r?

   a. If r^2 is 0.81, what are the possible values for r?

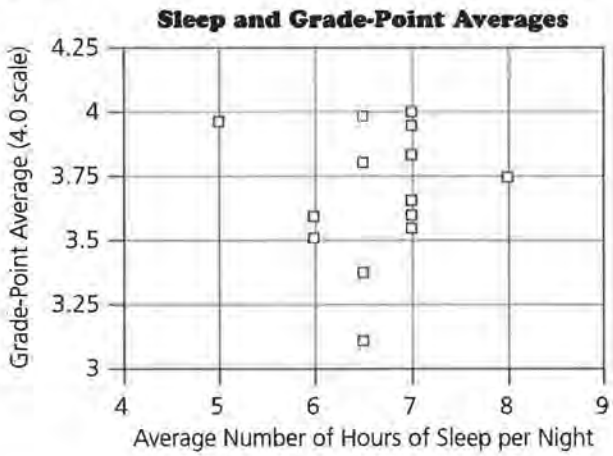
   b. Sketch a scatter plot and a line corresponding to a positive value of r. Make a similar sketch for a negative value of r.

   c. Sketch a plot showing a correlation close to 1.

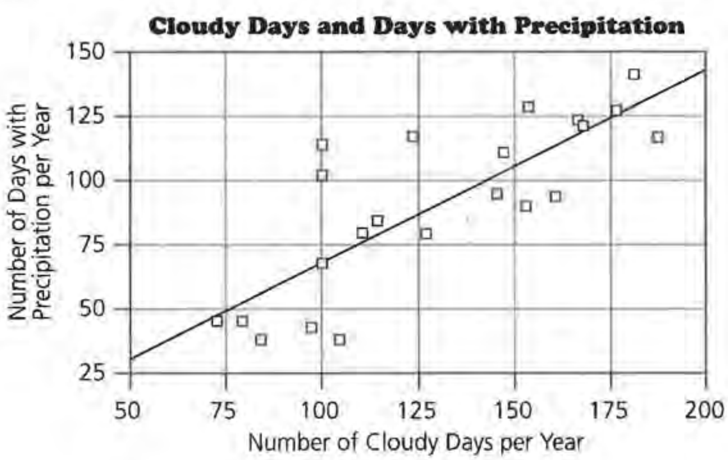
   d. Sketch a plot showing a correlation close to zero.

2. Describe the correlation you would expect from looking at the following plots.

   a. The hours of sleep on school nights versus the grade-point average



   b. The number of cloudy days per year in a set of cities versus number of days with precipitation per year



Correlation and Cause and Effect

People often confuse correlation with cause and effect. Just because two variables are correlated does not mean that one causes the other.

- They could both be a function of some other cause,
- one could cause the other, or
- the relationship could be purely coincidental.

Consider the relation between overall grade-point averages and grades in English. The association is probably strong, but English grades alone do not cause high grade-point averages; other courses contribute also. The association between grade-point average and hours of study is high, and it is reasonable to assume that the time spent studying is a primary cause of grade-point averages. The correlation between grade-point averages and SAT scores is strong, but neither variable causes the other. A good SAT score does not cause high grade-point averages.

Sometimes the relationship occurs purely by chance. It could happen that the correlation between grade-point averages and the distances students live from school is strong. It seems unlikely that all the good students live the same distance from school. Much more reasonable is the assumption that the connection is coincidental, and there is no real link between distance from school and grade-point average.

There are several different kinds of correlation and different procedures for finding correlation between variables. The correlation coefficient described here is called Pearson's r, and it is the most commonly used type of correlation.

Summary

The linear association between two variables can be measured by a number r called the correlation coefficient. If there is a perfect positive correlation, $r = 1$; if there is a perfect negative correlation, $r = -1$. A positive correlation indicates that as one variable increases, the other also tends to increase; while a negative correlation indicates that as one variable increases, the other tends to decrease.

If r is close to zero, then there is no good linear prediction of one variable from the other; that is, knowing the value of one does not help you predict the other using a linear model.

The correlation coefficient squared, r^2, indicates the proportion of error that can be explained by using the least-squares regression line. The closer r^2 is to 1, the more accurately x can be used to predict y.

The correlation coefficient measures only linear association rather than association in general. There may be a clear pattern in a set of data, but if it is not linear, the correlation may be close to zero.

Correlation is a number without any units attached. Therefore, correlation does not depend on the units chosen for either variable.

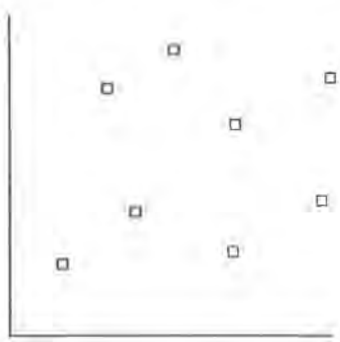
Many software packages calculate r automatically when they find the coefficients of the regression line. It is important, however, to look at the scatter plot to determine whether the relation is actually linear.

Correlation, the square of the correlation, and residual plots may be used to help assess a mathematical model or help select an appropriate mathematical model for data.

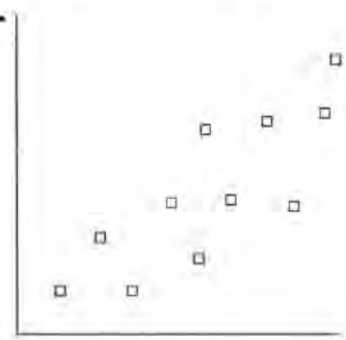
Practice and Applications

3. Match each correlation r and r^2 with the appropriate graph.

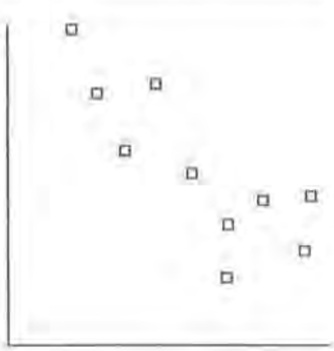
a.



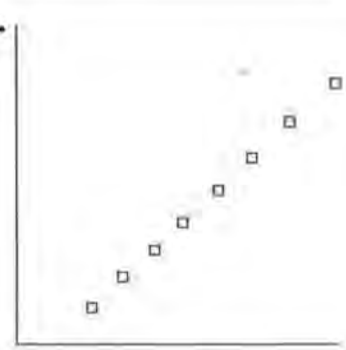
b.



c.



d.



$r = -0.8$ $r = 0.8$ $r = 0.99$ $r = 0.15$

$r^2 = 0.9801$ $r^2 = 0.0225$ $r^2 = 0.64$

4. The following quote appeared in a suburban Milwaukee newspaper article with the title *Spending More on Police Doesn't Reduce Crime.*

A CNI study of crime statistics and police department budgets over the last four years reveals there really is no correlation between what a community spends on law enforcement and its crime rate.

a. Sketch what you think a plot of the data would look like based on the quote from the article.

b. Use the data below about the suburban crime rate and the per-capita spending on police. Plot the data and find the correlation coefficient.

Community	Suburban Crime Rate per 1,000 Residents	Per-Capita Spending on Police
Glendale	74.39	$222.25
West Allis	50.43	$164.47
Greendale	50.25	$123.43
Greenfield	48.68	$143.59
Wauwatosa	48.52	$150.34
South Milwaukee	43.20	$110.64
Brookfield	42.16	$131.42
Cudahy	41.47	$137.12
St. Francis	41.32	$144.84
Shorewood	40.84	$156.39
Oak Creek	40.84	$160.41
Brown Deer	37.09	$150.57
Germantown	31.16	$125.86
Menomonee Falls	29.73	$159.28
Hales Corners	27.04	$155.40
New Berlin	25.45	$125.33
Franklin	23.09	$ 94.92
Elm Grove	21.86	$191.64
Whitefish Bay	21.28	$120.34
Muskego	17.00	$105.35
Fox Point	15.07	$132.85
Mequon	11.31	$136.39

Source: *Hub*, November 4, 1993

c. Does the correlation coefficient support the conclusions in the paragraph?

d. What does r^2 indicate about the relation between spending and the crime rate?

Developing a Mathematical Model

What is a mathematical model?

What mathematical concepts can be used to determine the best model for a given data set?

The Tree Growers Association has collected data about the ages of chestnut oak trees and their respective trunk sizes as measured by diameter. The Association would like to know if there is an optimum time to harvest the trees.

OBJECTIVE

Use the knowledge of transformations, logarithms, residuals, and correlation to develop a mathematical model.

INVESTIGATE

Because of your ability to develop mathematical models, you have been selected to prepare a report to be delivered at the next monthly meeting of the Tree Growers Association. Your task in this lesson includes two assignments.

- Find a mathematical model for the relationship between age and the size of the diameter of chestnut oak tree trunks using the data on page 81.
- Prepare a report that explains how you developed the model.

Your model will be used to help determine an optimum time to harvest the trees.

Chestnut Oak Trees

Age (years)	Trunk Diameter (inches)
4	0.8
5	0.8
8	1
8	2
8	3
10	2
10	3.5
12	4.9
13	3.5
14	2.5
16	4.5
18	4.6
20	5.5
22	5.8
23	4.7
25	6.5
28	6
29	4.5
30	6
30	7
33	8
34	6.5
35	7
38	5
38	7
40	7.5
42	7.5

Source: *Elements of Forest Mensuration,* Chapman and Demerritt

Discussion and Practice

1. Identify any patterns you observe when looking at the two columns of data individually. Consider the amount of increase within a column and the range of values.
2. Does the increase in diameter appear to be constant as age increases?
3. What patterns do you notice in the dependence between variables?

4. The Tree Growers Association reported they had recorded when the oak trees were planted; this fact made the data on their ages available. Determine the explanatory variable and the response variable for the data set. This identification will assist in determining the order in the ordered pairs.

Plot the Data

5. Enter the data into a graphing utility and draw a scatter plot. Write a short paragraph to identify the characteristics of the scatter plot.
6. Does a linear model fit the data?
   a. Draw a straight line that fits the scatter plot.
   b. Observe whether the line captures the characteristics you identified. Support your claim with a written argument.

Transform the Data to Straighten the Scatter Plot

To help you decide what transformations may straighten the data,

- use your knowledge of graph patterns for different functions, and
- consider what you know about the relationship between the explanatory and response variables.

7. Identify at least two transformations that you think will straighten the scatter plot of the original data set. Write a short paragraph explaining why you chose those transformations and to which variables you would apply them.
8. Make a scatter plot of each set of transformed data you think appears linear when graphed. Label the axes properly.
9. Use appropriate technology to draw the linear regression line on each graph. Then find and record the equation for each regression line and the correlation coefficient.

10. Use the lines you chose to predict the diameter of a tree for these specific ages.

a. 33 years

b. 21 years

c. 60 years

Compare One Transformation to Another

In order to determine which transformation is better, it is helpful to consider two different statistical tools: residuals and correlation.

Following are some questions to consider when looking at the plots of residuals.

- Do the residuals reveal a pattern that can be used to predict the error? If so, the line may not be considered a good fit for the data. It is better to have a random distribution of points.
- Do one or more of the data points have more influence that they should on the regression line? If so, examine the data points again.
- Do the residuals have a narrow vertical spread at one end of the plot and a wider spread at the other end? It is better to have a constant variation in the spread across the values of the explanatory variable (domain).

11. Use appropriate technology to draw residual plots for two of the scatter plots and lines you drew in Questions 8 and 9. Then discuss the three questions above for your residual plots.

12. Compare the correlation coefficients and the squares of the correlation coefficients for the lines you drew in Questions 8 and 9. Explain what they tell you about the data and the mathematical model.

13. Use what you learned by looking at residuals and correlation to choose a mathematical model for the oak trees data set.

Find a Model for the Original Data Set

14. Use the model you selected in Question 13. Rewrite each equation replacing the explanatory and response variables with the transformed variables. Remember that this equation is linear.

15. Use your knowledge of algebra and inverse functions to find a mathematical model for the original data set.

16. Graph the chestnut oak trees data and your model on the same graph. Describe your observations.

Interpret Results and Write a Summary

17. Write a summary report. Include the following.

- **a.** A summary of how you determined your model.
- **b.** A model (function) for the original data set.
- **c.** A graph that contains the original data set and the model.
- **d.** Your recommendation for the Tree Growers Association.

Summary

Process for Finding a Mathematical Model

- Study the data and identify patterns.
- Make a scatter plot of the data and examine the plot for any patterns.
- Look for functional relationships and try one or more transformations to straighten the scatter plot. Find linear models for the transformed data.
- Use residuals and correlation to assist in determining the best transformation for linearizing the data.
- Use your knowledge of transformations and functions to generate a model of the original data set.
- Interpret the results and write a summary of your findings.

It must be understood that the procedure used in this lesson is the development of a mathematical model. The concept of modeling is to use mathematical and statistical tools and techniques to create an equation or model to better understand a more complex process. If more sophisticated tools or additional data become available, the model can be changed to incorporate the new information. The interaction between mathematical models and data continues as long as man increases his knowledge.

Practice and Applications

18. In 1981, Boeing Aircraft Company did a study of people and their fear of flying. Data were obtained from a simple survey question, "Are you afraid of flying?" with responses of "yes" or "no" and the person's age.

Median Age	Percent of Population Sampled Afraid of Flying
20	1.840
25	2.670
30	3.690
35	4.890
40	6.280
45	7.850
50	9.610
55	11.550
60	13.680
65	15.990
70	18.490

a. Create a model to describe the relationship between median age and the percent afraid of flying.

b. Prepare an argument defending your model.

PROJECT

Alligators' Lengths and Weights

The following table of data was created by the Florida Game and Freshwater Fish Commission. Your assignment is to determine what function could be used to model the relationship between the length and weight of an alligator, or whether such a relationship even exists. Prepare a formal presentation (charts, graphs, and numeric and symbolic arguments), utilizing all the processes you have studied in this module, in defense of your position.

Length (in.)	Weight (lb)	Length (in.)	Weight (lb)
94	130	86	83
74	51	88	70
147	640	72	61
58	28	74	54
86	80	61	44
94	110	90	106
63	33	89	84
86	90	68	39
69	36	76	42
72	38	114	197
128	366	90	102
85	84	78	57
82	80		

Your report must include

- scatter plots of original and transformed data with the patterns in those graphs identified in writing,
- residual plots used and an analysis of each plot,
- correlation coefficients and related conclusions, and
- equations of lines, including an equation that can be used with the original data as a predictor equation.

ASSESSMENT

The Growth of Bluegills

The following table of data concerning the length of bluegills was created by the Florida Game and Freshwater Fish Commission.

Initial Length (mm)	Total Length After 1 Year (mm)	Initial Length (mm)	Total Length After 1 Year (mm)
48	69	138	160
52	71	138	157
51	69	130	156
53	75	140	161
69	101	160	173
71	107	157	168
69	100	156	172
75	104	161	178
101	138	173	176
107	138	168	174
100	130	172	173
104	140	178	178

OBJECTIVES

Find and interpret slope as a rate of change.

Find the rate of change from data.

1. Make a scatter plot of these data.
2. Identify in a short paragraph the characteristics you determine from the numeric and graphic displays of these data.
3. Draw a straight line through the scatter plot and determine whether or not the line captures any or all of the characteristics you identified in Question 2.
4. If the data set's scatter plot does not appear linear, perform a series of transformations on one or both of the variables to attempt to linearize the plot.
5. Select the plot or plots that appear the most linear, find the linear model, and check r and r^2 to determine how well the model fits the data.

6. Plot the residuals against x and consider the scale to determine a best line.

7. Use inverses of the transformations used for the equation you chose in Question 5 to create an equation that will be the best predictor equation of the original data set.

Dale Seymour Publications® is a leading publisher of K–12 educational materials in mathematics, thinking skills, science, language arts, and art education.

DALE
SEYMOUR
PUBLICATIONS®

90000

9 781572 322509

ISBN 1-57232-250-0
21185