Section 3
Analyzing Your Data

Key Concept Scientists analyze data in order to answer questions, understand results, and make predictions.

What You Will Learn
- Mathematics is an important tool for understanding and summarizing large quantities of information.
- The accuracy and reproducibility of data affect the results and conclusions of scientific studies.
- The mean, median, and mode are terms used to describe and analyze an entire set of data.
- Slope is the degree of slant, or steepness, of a straight line. The slope of a linear graph represents a constant that can be used to help analyze a set of data.

Why It Matters
Knowing how to analyze data can help you understand the results and predictions made in scientific studies.

One way to analyze data is to use mathematics. Mathematical models in the form of computer simulations can answer questions about how rockets, such as the one in **Figure 1,** will fly and react to different conditions before the rockets even leave the ground.

Figure 1 Mathematics was used to design the White Knight carrier plane and the SpaceShipOne rocket at right.

Why Mathematics?

Just like making observations, conducting experiments, and organizing data, mathematics is used to answer questions. Mathematics helps determine important properties of substances, such as area, volume, and density. Mathematics also allows scientists to understand and sum up a lot of information. As a result, scientists can make predictions. For example, a meteorologist gathers data on hurricane movement. Mathematics helps her to see patterns in the data. Then, she uses these patterns to predict where future hurricanes might hit land.

There are scientists in every country around the world. They speak many different languages. Mathematics is often called the *language of science* because it allows scientists to easily share their findings with each other in a language that everyone understands: numbers!

Accuracy of Data

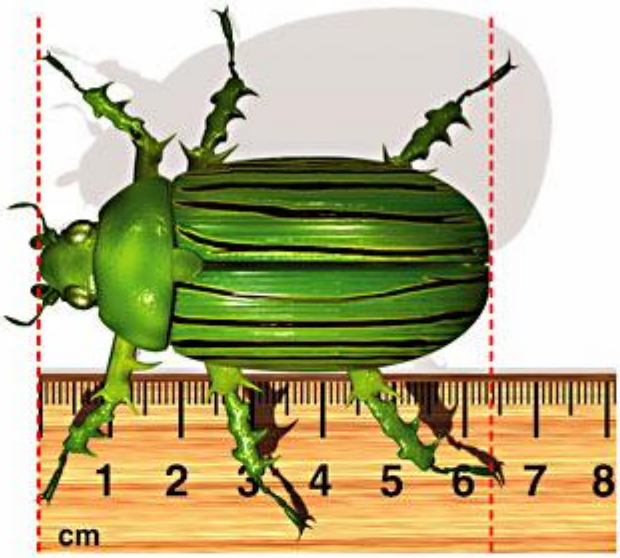
When scientists conduct experiments, they want to collect accurate data. In other words, they want the values to be correct. If you place 525 g on a balance and get a mass reading of 450 g, your reading is not accurate. Some reasons why a value is inaccurate include using broken equipment, using the wrong tool, or using a tool incorrectly.

Choosing Tools and Using Them Correctly

You have probably made measurements using the tools shown in **Figure 2.** Using a graduated cylinder to measure volume can provide a more accurate measurement than using a measuring cup found in a kitchen. But your data will still be inaccurate if you do not read the volume at the bottom of the meniscus at eye level. Likewise, to get an accurate reading using a ruler, you should look straight down on the end of the thing you are measuring. If you move your head to either side, you will get a slightly different measurement.

Figure 2 Using Tools Correctly

The most accurate reading of the beetle's length is from directly overhead. ***If you make your measurement when you are slightly to the right of the beetle rather than directly overhead, how would your measurement compare with the actual length of the beetle?***



Reproducibility of Data

Imagine that you and a friend do an experiment using the same procedure and equipment. You'd expect to get data that are very similar. When scientists conduct investigations, they want their results to be able to be repeated, or reproduced, by other scientists. If the data are not reproducible, then there is no way for the results of

the experiment to be supported and accepted by other scientists. Data must be reproducible in order for other people to agree with your conclusions.

Standards Check Explain what it means to say that the data from a scientific investigation are reproducible.

Describing the Entire Set of Data

When scientists analyze data, they often find it helpful to use a single number to describe the entire set of data. Three terms that are used for this purpose are *mean*, *median*, and *mode*. **Figure 3** shows how to find each one. The **mean,** or average, is found by adding all the data points together, then dividing the sum by the total number of data points. The **median** is the value of the data point in the middle when the data are placed in order from smallest to largest. The median is especially helpful when one data point is much smaller or larger than the rest of the data points, as in Week 9 below. The **mode** is the number that appears most often in a data set.

Figure 3 Analyzing the Entire Set of Data

A class of students recorded the total number of hours they exercised during Weeks 6–10.

Week	Number of hours exercised
6	40 hr
7	46 hr
8	43 hr
9	96 hr
10	40 hr

Finding the mean

Add your data together.
Divide the sum by the number of observations in your data set.

40
46
43
96
40
sum = 265

Mean = 265 ÷ 5

Mean = 53 hours exercised

Finding the mode

Place your data in order from smallest to largest. The number that appears most frequently is the mode.

40, 40, 43, 46, 96

Mode = 40 hours exercised

Finding the median

Place your data in order from smallest to largest. The number in the middle is the median.

40, 40, 43, 46, 96

Median = 43 hours exercised

[redacted]

Slope of a Line

What do you think of when you hear the word *slope*? You might think of how something slants uphill or downhill. **Figure 4** shows a skier moving down a slanted mountainside. Imagine drawing a straight line under the skier that runs along the mountainside. The steepness of the line representing the slant of the mountainside is called the slope of the line. In math, **slope** is defined as the degree of slant of a line.

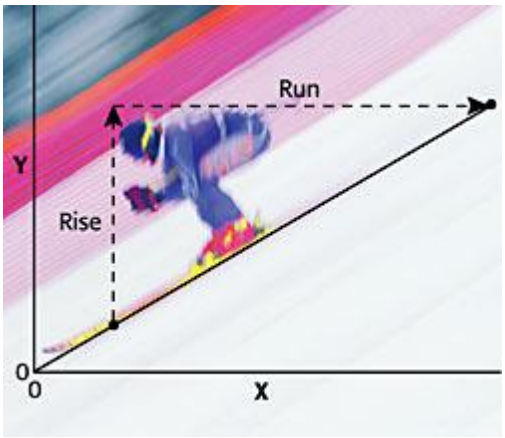


Figure 4 In skiing, the term slope refers to a slanted mountainside. The steeper a slope is, the higher its difficulty rating will be for a skier. In math, slope is calculated by dividing the rise by the run.

Calculating Slope

To calculate the slope of a line, it is helpful to use the terms *rise* and *run*, as shown in **Figure 4.** The rise represents a vertical change. The run represents a horizontal change. For a line on the coordinate plane, the change in y, or the rise, indicates the number of units moved up or down. The change in x, or the run, is the number of units moved to the right or left. Slope is calculated by dividing the vertical change (the change in y) by the horizontal change (the change in x). In other words, the slope of a straight line is found by dividing the rise by the run, often described as rise over run.

Using Slope to Analyze Data

The slope of a line graphed from data can help you analyze the data. Look at **Figure 5.** Three sets of data result in three different lines, each with a different slope. For each line, the value of the slope between any two points on that line will be a constant number. Here, the slope (rise over run) represents the speed (meters over seconds) of an object moving at a constant rate. **Figure 5** shows that slope can have a positive, negative, or zero value. The slope of all horizontal lines is zero.

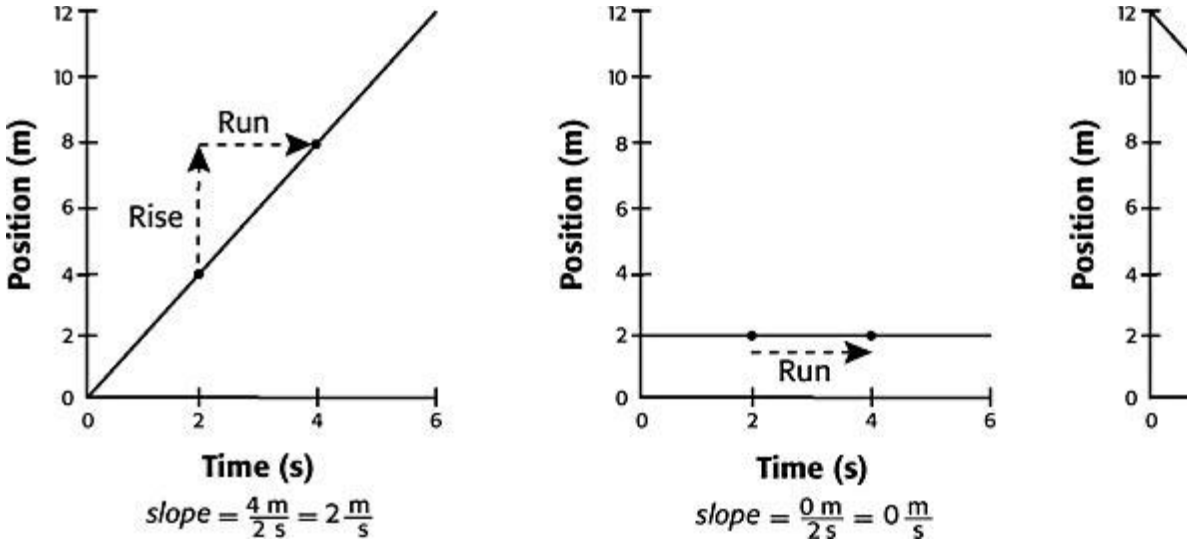
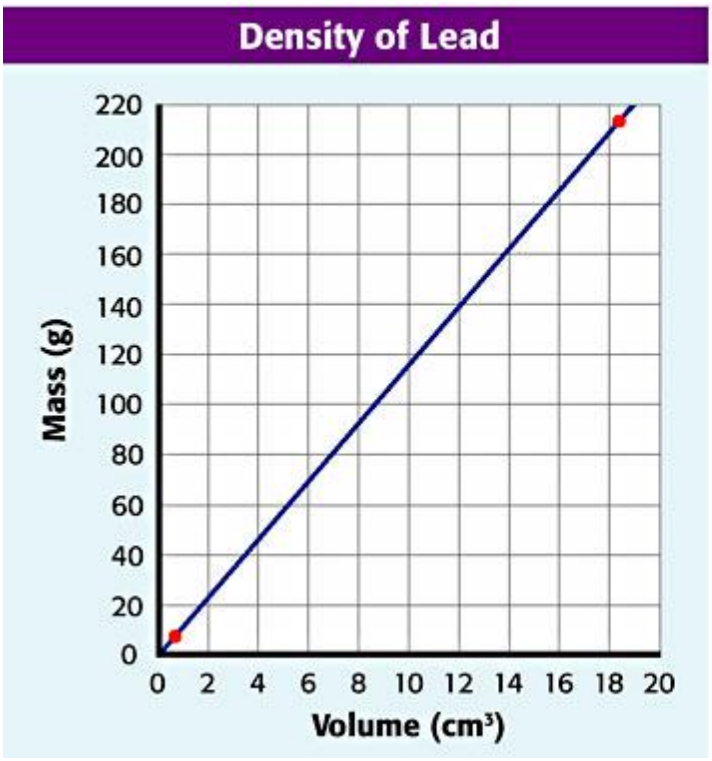


Figure 5 The results from three sets of data can be graphed to analyze the data. Here, the slope of a line on a position versus time graph represents the speed of the object.

Standards Check Define the slope of a straight line using the terms *rise* and *run.*

Slope as the Constant in *y = kx*

Scientists sometimes find that their data form a straight line that includes the point that has *x* and *y* values of zero. They can then use the equation $y = kx$ to represent their data. **Figure 6** is a graph representing the density of lead. Here, *y* represents measurements of mass. And *x* represents measurements of volume. Then, *k* represents the constant term, which is the slope of the line. The slope of this line (rise divided by run) represents the density (mass divided by volume) of lead. If you rearrange the formula for density into the form $y = kx$, the equation is $m = DV$. So, density is the constant, *k,* which is the slope of the line.



Lead fishing weight
Mass = 211.0 g
Volume = 18.59 cm^3

Figure 6 The line in the graph at left represents the constant density of lead. When plotting the mass and volume measurements for the two lead fishing weights above, the data points fall on the line shown in the graph.

Comparing Linear and Nonlinear Graphs

The lines on a graph can help you draw conclusions about your data. The slope of a straight line shows how much one factor (y) always changes in relation to another factor (x). A straight line shows that a constant linear relationship exists between the factors you are studying. In other words, y always changes the same amount in relation to x. But what

if your data do not form a straight line? Many relationships that scientists study are not linear. In a nonlinear graph, each unit change in one factor (y) does not always bring about the same change in another factor (x). The graph of this relationship will be a curve instead of a straight line. So, there is a nonlinear relationship between the factors you are studying.

Section Summary

- Mathematics is an important tool for understanding and summarizing data.
- The accuracy and reproducibility of data used in scientific investigations affect the results.
- Mean, median, and mode summarize an entire set of data.
- Slope is the degree of slant of a straight line.
- The slope of a straight line represents a constant that can be used to understand and analyze data.
- Linear and nonlinear graphs result from different relationships in the data.

Chapter Summary

The Big Idea

Scientists use tools to collect, organize, and analyze data while conducting investigations.

Section 1

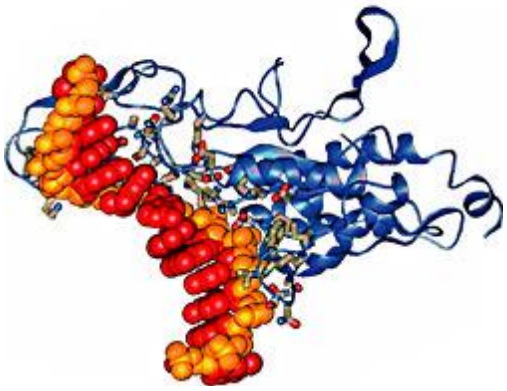
Tools and Models in Science

Key Concept Scientists use tools and models to increase their ability to investigate the natural

world.

- Tools are used to make accurate measurements while collecting data.
- The International System of Units (SI) is a system of measurement used by most scientists.
- A model uses familiar things to describe unfamiliar things. Physical, conceptual, and mathematical models are commonly used in science.
- Models help scientific progress through their use in theories and laws.

Scientists use models to help understand and explain data.



Section 2
Organizing Your Data

Key Concept Scientists organize data to make quantitative statements about the relationships between the variables in an investigation.

- Scientists use data tables and graphs to organize information.
- The independent variable is the factor or parameter that the investigator can change.
- The dependent variable is the factor or parameter that the investigator measures.
- Graphs help show patterns in data. Linear and nonlinear graphs show different relationships between the variables.

Organizing your data helps you see trends and draw conclusions.

Section 3
Analyzing Your Data

Key Concept Scientists analyze data in order to answer questions, understand results, and make predictions.

- Mathematics is an important tool for understanding and summarizing large quantities of information.
- The accuracy and reproducibility of data affect the results and conclusions of scientific studies.
- The mean, median, and mode are terms used to describe and analyze an entire set of data.
- Slope is the degree of slant, or steepness, of a straight line. The slope of a linear graph represents a constant that can be used to help analyze a set of data.

The accuracy of measurements affects your data.