

## What We Are Learning

## Collecting and Describing Data

## Vocabulary

These are the math words we are learning:

**back-to-back stem-and-leaf plot** a means to compare two sets of data

**biased sample** a sample that is not based on a good representation of the population

**box-and-whisker plot** a way to show the distribution of the data through the use of quartiles and the median

**convenience sample** a sample whose members are easiest to reach

**line plot** a data display that uses a number line to show how often an value occurs

**mean** the sum of the values, divided by the total number of values

**median** if an odd number of values, the middle value; if an even number of values, the average of the two middle values

**mode** the value or values in a data set that occur most often

**outlier** an extreme value that may have an effect on the mean of the data.

**population** the entire group being studied

*Dear Family,*

When studying a group or a population, it is important to get true and accurate information. While interviewing everyone in a population would achieve the most precise results, this is often not feasible. Usually a **sample** of the **population** is used to represent the views of an entire group.

**Identify the population and the sample. Give a reason why the sample could be biased.**

*A local newspaper sends a survey to 500 local subscribers to find which mayoral candidate people prefer.*

Population: People in the local community

Sample: Up to 500 subscribers who take the survey.

Possible Bias: Not all people in the community may subscribe to the local paper.

Your child will learn to organize data so it can be evaluated. A **stem-and-leaf plot** is one method used to organize large amounts of data in a simple, yet precise, manner.

**Create a stem-and-leaf plot of the data values.**

85, 74, 91, 77, 86, 80, 71, 79, 82, 84, 99, 62

**Step 1** Find the least data value and the greatest data value. Since the data values range from 62 to 99, use the tens digits for the **stems** and the ones digits for the **leaves**.

**Step 2** List the stems from least to greatest on the plot.

**Step 3** List the leaves for each stem from least to greatest. For the number 62, the 6 is the stem and the 2 is the leaf.

Stem	Leaves
6	2
7	1 4 7 9
8	0 2 4 5 6
9	1 9

**quartile** the division of a set of data into four equal parts

**random sample** every member of the population has an equal chance of being chosen

**range** the largest data value minus the smallest data value in a set of data

**sample** part of the population being surveyed

**stem-and-leaf plot**

a graph used to organize and display data so that the frequencies can be compared

**stratified sample**

a sampling method that chooses from randomly chosen subgroups

**systematic sample**

a sampling method that follows a rule or formula

**variability** the description of how spread out a data set is

**Venn diagram** a diagram that shows relationships between sets

**voluntary-response**

**sample** a sample whose members chose to respond

Once the data is organized, your child will use the measures of central tendency to describe the data.

<b>Mean</b>	The sum of the values divided by the number of values. Otherwise known as the average.
<b>Median</b>	The median is the middle value if there is an odd number of values. If there is an even number of values, the median is the average of the two middle numbers.
<b>Mode</b>	The value or values that occur most often in a set of data. If no value occurs more than once, then there is no mode.

**Find the range, mean, median, and mode of the data set.**

26, 27, 28, 22, 3, 28, 26, 26, 22, 20

**Range**

$$28 - 3 = 25$$

The range is 25.

Subtract the least value from the greatest value.

**Mean**

$$26 + 27 + 28 + 22 + 3 + 28 + 26 + 26 + 22 + 20 = 228$$

$$228 \div 10 = 22.8$$

The mean is 22.8.

Add the values.

Divide the sum by the number of items.

**Median**

3, 20, 22, 22, 26, 26, 26, 27, 28, 28

The median is 26.

Arrange the values in order. Choose the number in the middle of the data set.

**Mode**

The mode is 26.

The value 26 occurs 3 times.

Your child will also learn to identify an outlier. An outlier is an extreme piece of data. For instance, in the data set above, 3 is an outlier. If you remove 3 from the data set the mean becomes 25. Ask your child to explain how the other measures of central tendency are affected when you remove the outlier.

Encourage your child to be aware of the everyday instances of data analysis and the importance it has in our lives.

**Sincerely,**

**CHAPTER**  
**9**

**Family Letter**

**Collecting and Describing Data**

**Identify the population and sample. Give a reason why the sample could be biased.**

1. A deli clerk surveys 20 customers to see what kind of cheese is purchased most frequently.

Population                      Sample                      Possible Bias

\_\_\_\_\_

2. A librarian asks the first four people who return the latest bestseller if they liked the book.

Population                      Sample                      Possible Bias

\_\_\_\_\_

**Identify the sampling method used.**

3. The names of all children who are 12 years old and want to play soccer are each written on a piece of paper and placed into a bowl. Each coach selects ten names, one at a time, to make a team.

\_\_\_\_\_

**List the data values in the stem-and-leaf plot.**

4.

1	1	2
2	3	4 6
3	2	8

\_\_\_\_\_

5.

4	0	1	2
5	1	5	8
6	0	1	1 5
7	0	3	5 5 8

\_\_\_\_\_

**Find the mean, median, and mode of each data set.**

6. 15, 40, 22, 46, 7, 20, 22, 12  
mean              median              mode

\_\_\_\_\_

7. 106, 46, 93, 98, 92, 95, 100  
mean              median              mode

\_\_\_\_\_

**Find the range and the first and third quartiles for each data set.**

8. 14, 8, 25, 26, 11, 16, 23, 15, 40  
range              first quartile              third quartile

\_\_\_\_\_

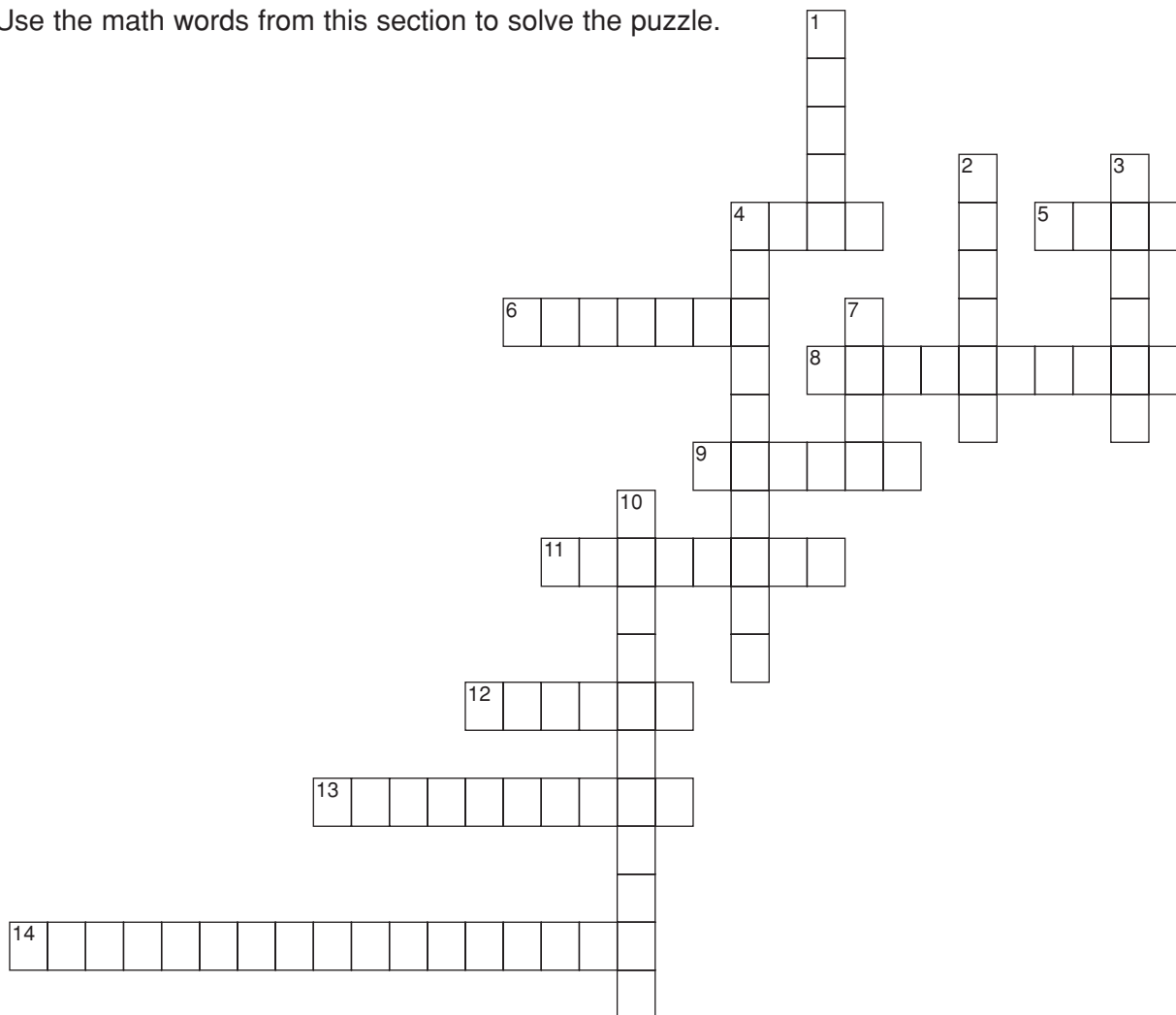
9. 94, 85, 76, 98, 82, 87, 78  
range              first quartile              third quartile

\_\_\_\_\_

**Answers:** 1. Population: all store customers; Sample: 20 customers; Possible Bias: Not all customers buy or like cheese 2. Population: All people who checked out books; Sample: 4 people who returned bestseller first; Possible Bias: Not all readers may have finished the book; person bringing back the book did not read it. 3. random 4. 11, 12, 23, 24, 26, 32, 38 5. 40, 41, 42, 51, 55, 58, 60, 61, 65, 70, 73, 75, 75, 78 6. mean: 23; median: 21; mode 22 7. mean: 90; median: 95; mode: no mode 8. range: 32; first quartile: 12.5; third quartile: 25.5 9. range: 22; first quartile: 78; third quartile: 94

**Directions**

Use the math words from this section to solve the puzzle.



**Across**

- 4. The \_\_\_\_ is the first value in this type of plot.
- 5. Average
- 6. An extreme value that affects the mean.
- 8. The group being studied.
- 9. This type of sample does not represent the whole population
- 11. The division of data into four equal parts.
- 12. The middle number in a set of data
- 13. Sampling method example - Every third child will get a letter home.
- 14. Shows the distribution of data using the median and quartiles

**Down**

- 1. The difference between the largest value and the smallest value.
- 2. The part of the population being studied.
- 3. A sampling method where anyone in the population can be chosen.
- 4. A type of sample that chooses from subgroups.
- 7. In the set 3, 7, 3, 4, 6, 8, 5, 3, 8, the \_\_\_\_ is 3.
- 10. Description of a data spread.

**Answers:** Across: 4. stem 5. mean 6. outlier 8. population 9. biased 11. quartile 12. median 13. systematic 14. box and whisker plot Down: 1. range 2. sample 3. random 4. stratified 7. mode 10. variability

## What We Are Learning

## Displaying Data

## Vocabulary

These are the math words we are learning:

**correlation** describes the type of relationship between two data sets

**double-bar graph** a way to display related sets of data that are grouped into categories

**double-line graph** a way to display related sets of data that change over time

**frequency table** a way to organize data in terms of the number of times each value occurs

**histogram** a type of bar graph where the data is grouped in intervals

**line of best fit** the line that comes closest to all the points on a scatter plot

**scatter plot** shows the relationship between two sets of data

*Dear Family,*

In this section your child will be interpreting and constructing many different types of graphs used to display data.

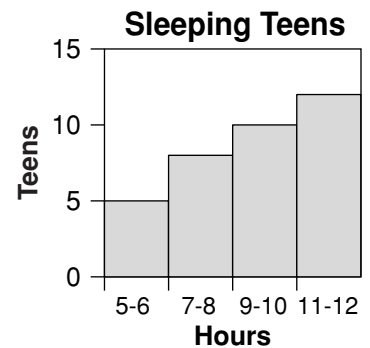
**Bar graphs** are helpful when you want to display data that can be put into categories.

**Frequency tables** help you organize data so you can tell how many items fall into a particular category.

A **histogram** is a type of bar graph that shows the frequency of data within equal intervals.

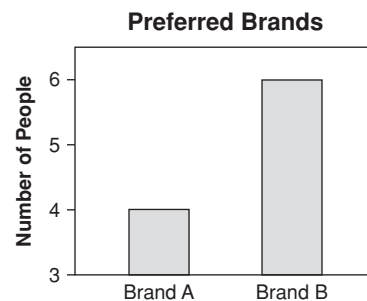
**The frequency table shows the average number of hours per day in the summer that teenagers spend sleeping. Use this data to make a histogram.**

Hours	Frequency
5–6	5
7–8	8
9–10	10
11–12	12



- Step 1** Choose an appropriate scale and interval for the vertical axis. The greatest value for this axis should be at least as great as the greatest frequency.
- Step 2** Draw a bar for each interval. The height of the bar is the frequency. Bars must touch, not overlap.
- Step 3** Label both axes. Give the graph a title.

As your child learns to create different types of graphs, it is critical that she or he learns to distinguish an accurate portrayal of data versus a misleading one.

**Explain why the graph is misleading.**

At first glance, it appears that twice as many people prefer Brand B than Brand A. If you look closely, the scale does not start at zero, but at 3. By reading the graph you can learn that 4 people prefer Brand A and 6 people prefer Brand B.

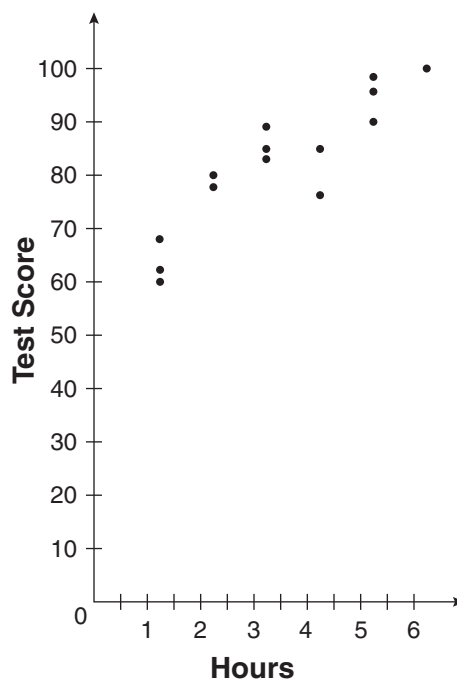
The graph is misleading because the difference between brands is not as large as the graph portrays.

Another way to graph data is to make a **scatter plot**. A scatter plot is made by plotting individual points for each data entry. By studying the grouping of the points, you can tell whether or not the two sets of data are related.

**Make a scatter plot and tell whether or not time spent studying is related to test grades.**

**Study Time and Test Scores in Mr. Baez's Class**

Hours	Score
5	98
1	63
2	80
3	89
4	77
5	96
0	52
1	67
6	100
2	78
3	84
1	60
3	85
4	85
5	90



There is a relationship between study time and test scores. The more time spent studying, the higher the test score.

Encourage your child to be aware of the day-to-day examples of how data is presented. Discuss these displays with your child to help him or her gain understanding of the concepts presented in this section.

**Sincerely,**

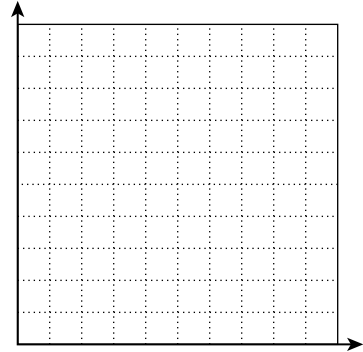
**CHAPTER**  
**9**

**Family Letter**  
**Displaying Data**

1. Organize the data into a frequency table and make a bar graph.

25 24 22 26 23 26 29  
22 26 26 23 28 25 27

Interval	Frequency
22–23	
24–25	
26–27	
28–29	

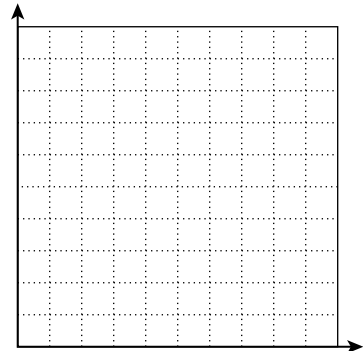


Explain why the statistic is misleading.

2. A school reporter asked 100 students what their favorite subject was. Of the 30 that responded, 14 said math, 8 said science, 5 said history, and 3 said physical education. The reporter wrote, "Half of all the students pick math as their favorite subject."

3. A group of students had babysitting jobs over the weekend. Use the given data to make a scatter plot.

Name	Total Hours Worked	Amount Earned
Sheri	5	\$22.50
Jordan	8	\$36.00
Lydia	11	\$49.50
Alexis	6	\$27.00



4. Use the data from the scatter plot to predict the number of hours Sam would have to work to make \$30.

Do the sets of data have a positive, negative, or no correlation?

5. The weight of a baby and the month that it is born.

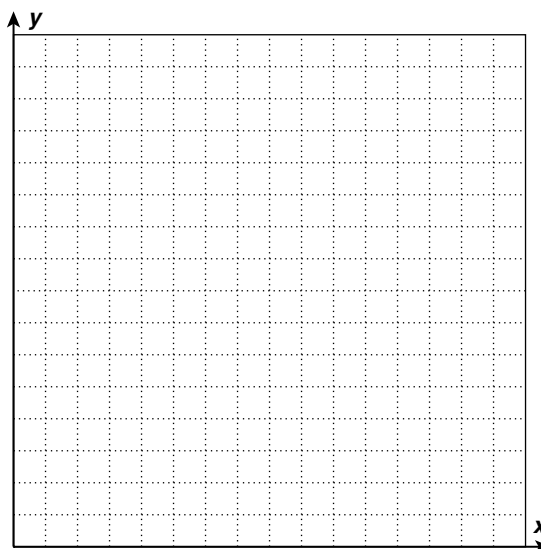
6. The amount of free time you have and the number of sports that you play.

**Answers:** 1. Table: 4 marks, 5 marks, 3 marks, 2 marks; Bars on the graph should measure accordingly. 2. The sample response was too small. Only 30% responded and less than 50% stated math was their favorite subject. 3. Points should be plotted at: (5, 22.50), (8, 36), (11, 49.50) and (6, 27); x-axis is hours worked, y-axis is amount earned 4. about 7 hours 5. no correlation 6. negative correlation

**CHAPTER 9** **Family Fun**  
**Graphing for Greatness**

It is said that there is a statistic for just about everything in the game of baseball. Choose a player and look up his statistics, like homeruns, strikeouts, RBI's, or stolen bases. Make a line graph on the graph below using your player's statistics over a given period of time. Be sure to label the axes and title your graph. Then answer the questions below.

For the non-baseball fan, research the statistics of another athlete, such as a tennis player, golfer, swimmer, or gymnast.



1. Who is your athlete? What team was he involved with, if any?  
 \_\_\_\_\_
2. Which statistic did you investigate? \_\_\_\_\_
3. Which seasons/years were the most successful for your athlete?  
 \_\_\_\_\_
4. Which seasons/years did your athlete seem to struggle?  
 \_\_\_\_\_
5. How long did your player play his/her sport? \_\_\_\_\_
6. What, if anything, affected your player's performance? (i.e. age, illness, experience)  
 \_\_\_\_\_
7. How do you think this information helped the opposing team compete against your player?  
 \_\_\_\_\_