Investigation 1

Looking at Data

The problems in this investigation involve people's names. Family traditions are often involved when a child is named. A person's name may reveal information about his or her ancestors.

Many people have interesting stories about how they were named. Here is one student's story: "I'm a twin, and my mom and dad didn't know they were going to have twins. My sister was born first. She was named Sukey. I was a surprise. My mom named me after the woman in the next hospital bed. Her name was Takara."

- Do you know anything interesting about how you were named or about the history behind your family's name?

Did You Know?

Rhoshandiatellyneshiaunneveshenk Koyaanisquatsiuth Williams is the longest name on a birth certificate.

Shortly after Rhoshandiatellyneshiaunneveshenk was born, her father lengthened her first name to 1,019 letters and her middle name to 36 letters. What is a good nickname for her?

1.1 Organizing and Interpreting Data

Most parents do not worry about the number of letters in their children's names. Sometimes though, name length does matter. For example, only a limited number of letters may fit on a bracelet or a library card.

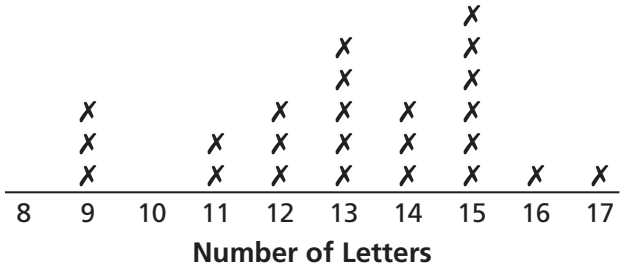
Getting Ready for Problem 1.1

What do you think is the typical number of letters in the full names (first and last names) of your classmates?

- What data do you need to collect and how would you collect it?
- How would you organize and represent your data?
- If a new student joined your class today, how might you use your results to predict the length of that student's name?

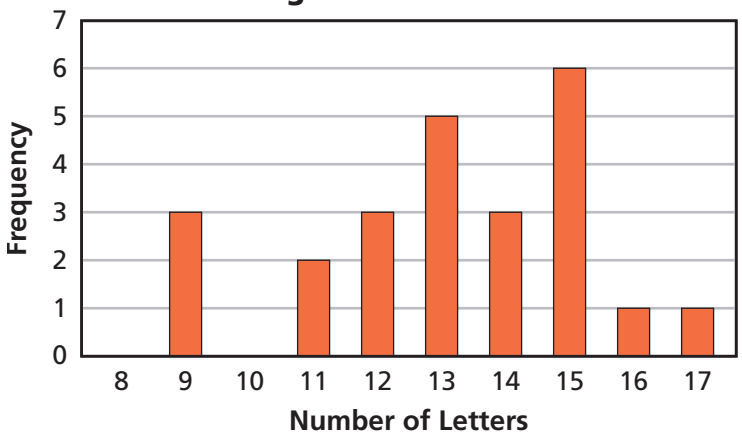
The students in Ms. Jee's class made a **line plot** to display the distribution of their class's data.

Name Lengths of Ms. Jee's Students



Another group displayed the same data using a **bar graph.**

Name Lengths of Ms. Jee's Students



To describe how the data are distributed, you might look at where the data values cluster, how much they vary, and the high and low values.

Problem 1.1 Organizing and Interpreting Data

Examine the line plot and the bar graph.

A. Describe the distribution of the data. Do you see any patterns?

B. How are the two graphs alike? How are they different?

C. How can you use each graph to determine the total number of letters in all the names?

D. Fahimeh Ghomizadeh said, "My name has the most letters, but the bar that shows my name length is one of the shortest on the graph. Why?" How would you answer this question?

E. Collect the data for your class's name lengths. Represent the data distribution using a line plot or a bar graph.

F. What are some similarities and differences between the data distribution from Ms. Jee's class and the data distribution from your class?

ACE Homework starts on page 21.

Did You Know?

In Africa, a child's name is often very meaningful. Names such as Sekelaga, which means "rejoice," and Tusajigwe, which means "we are blessed," reflect the happiness the family felt at the child's birth. Names such as Mvula, meaning "rain," reflect events that happened at the time the child was born.

Go Online PHSchool.com **For:** Information about African names **Web Code:** ame-9031

1.2 Useful Statistics

In the data for Ms. Jee's class, the name length of 15 letters occurs most often. Notice that 15 has the highest stack of X's in the line plot and the tallest bar in the bar graph. We call the most frequent value the **mode** of the data set.

The least value and the greatest value are important values in a data set. They give a sense of the variability in the data. In Ms. Jee's class, the data vary from 9 letters to 17 letters. The difference between the least value and the greatest value is called the **range** of the data. The range of Ms. Jee's class data is 17 – 9, or 8 letters.

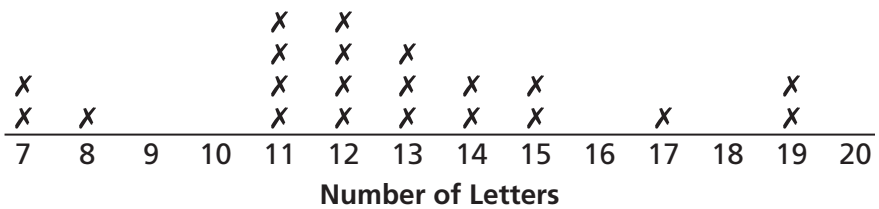
Still another important statistic is the **median,** or the midpoint, of the data set.

The table and line plot below show the distribution of the name-length data for Mr. Gray's class. Notice that these data have two modes, 11 letters and 12 letters. We say the distribution is *bimodal*. The data vary from 7 letters to 19 letters. The range of the data is 19 – 7, or 12 letters.

Name Lengths of Mr. Gray's Students

Name	Number of Letters
Jeffrey Piersonjones	19
Thomas Petes	11
Clarence Jenkins	15
Michelle Hughes	14
Shoshana White	13
Deborah Black	12
Terry Van Bourgondien	19
Maxi Swanson	11
Tonya Stewart	12
Jorge Bastante	13
Richard Mudd	11
Joachim Caruso	13
Robert Northcott	15
Tony Tung	8
Joshua Klein	11
Jan Wong	7
Bob King	7
Veronica Rodriguez	17
Charlene Greene	14
Peter Juliano	12
Linora Haynes	12

Name Lengths of Mr. Gray's Students



Problem 1.2 Useful Statistics

Here is a way to help you think about how to identify the median. Cut a strip of 21 squares from a sheet of grid paper. Each square is for the length of a student's name in Mr. Gray's class. Write the name lengths of Mr. Gray's students in order from least to greatest on the grid paper as shown.

7	7	8	11	11	11	11	12	12	12	12	13	13	13	14	14	15	15	17	19	19

A. Fold the strip in half.

1. On what number is the crease caused by the fold?

2. How many numbers occur to the left of this number?

3. How many numbers occur to the right of this number?

4. The median is the number that is the midpoint of a set of data. The same number of data values occurs before and after the median. What is the median for these data?

B. Suppose a new student, Suzanne Mannerstrale, joins Mr. Gray's class. The class now has 22 students. On a strip of 22 squares, list the name lengths, including Suzanne's, in order from least to greatest. Fold the strip in half.

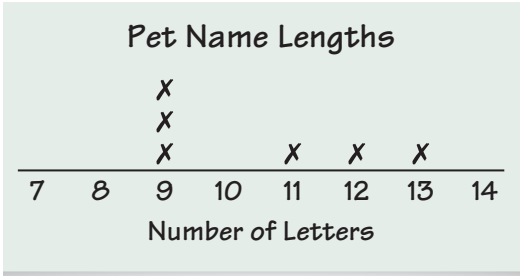
1. On what number is the crease caused by the fold?

2. How many numbers occur to the left of the crease?

3. How many numbers occur to the right of the crease?

4. What is the median for these data?

C. Suzanne has six pets. She made the line plot shown of her pets' name lengths. Find the median length of her pets' names. Find the mode for the data set.



D. There are 15 students in a class. Use the information about the class's name lengths below.

- Mode: 12 letters
- Median: 12 letters
- The data vary from 8 letters to 16 letters

1. Find a possible set of name lengths for the class.

2. Make a line plot to display your data distribution.

3. Compare your graph with the graphs of your classmates. How are the graphs alike? How are they different?

ACE Homework starts on page 21.

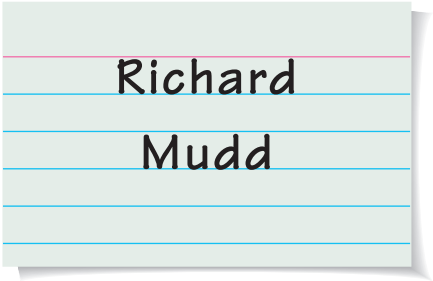
1.3 Experimenting With the Median

You can use the median and the mode of a set of data to describe what is typical about the distribution. They are sometimes called *measures of center.*

Use the following ten names. Write each name on an index card. On the back of each card, write the number of letters in the name. A sample index card is shown below.

Student Name Lengths

Name	Number of Letters
Thomas Petes	11
Michelle Hughes	14
Shoshana White	13
Deborah Black	12
Tonya Stewart	12
Richard Mudd	11
Tony Tung	8
Janice Wong	10
Bobby King	9
Charlene Greene	14



front

back

Order the cards from shortest name length to longest name length, and identify the median of the data.

Problem 1.3 Experimenting With the Median

Use your cards to complete each task below. Keep a record of your work.

A. Remove two names from the original data set so that

1. the median stays the same.

2. the median increases.

3. the median decreases.

B. Add two new names to the original data set so that

1. the median stays the same.

2. the median increases.

3. the median decreases.

C. How does the median of the original data set change if you add a name

1. with 16 letters?

2. with 1,019 letters?

ACE Homework starts on page 21.

Did You Know?

Names from many parts of the world have special origins. European family names (last names) often came from the father's first name. For example, Ian Robertson was the son of Robert, Janos Ivanovich was the son (vich) of Ivan, and John Peters was the son of Peter.

Family names also came from words that described a person's hometown or job. This resulted in such names as William Hill and Gilbert Baker.

Family names in China and Vietnam are almost always one-syllable words that are related to names of ruling families. Chang is one such example.

You can read more about names in books such as *Names from Africa* by Ogonna Chuks-Orji and *Do People Grow on Family Trees?* by Ira Wolfman.

Go Online PHSchool.com **For:** Information about names **Web Code:** ame-9031

1.4 Using Different Data Types

When you are interested in learning more about something, you ask questions about it. Some questions have answers that are words or categories. For example, what is your favorite sport? Other questions have answers that are numbers. For example, how many inches tall are you?

Categorical data are data that are specific labels or names for categories. They are usually not numbers. Suppose you ask people in which month they were born or what kinds of pets they have. Their answers would be categorical data.

Numerical data are data that are counts or measures. Suppose you ask people how tall they are or how many pets they have. Their responses would be numerical data.

Getting Ready for Problem 1.4

Read each of the questions below. Which questions have words or categories as answers? Which questions have numbers as answers?

- In which month were you born?
- What is your favorite kind of pet?
- How many pets do you have?
- Who is your favorite author?
- How much time do you spend watching television in a day?
- What's your highest score in the game?
- How many movies have you watched in the past week?

The kinds of pets people have often depend on where they live. People who live in cities often have small pets. People who live on farms often have large pets. People who live in apartments sometimes cannot have pets at all.

One middle-school class gathered data about their pets by tallying students' responses to these questions:

- What is your favorite kind of pet?
- How many pets do you have?

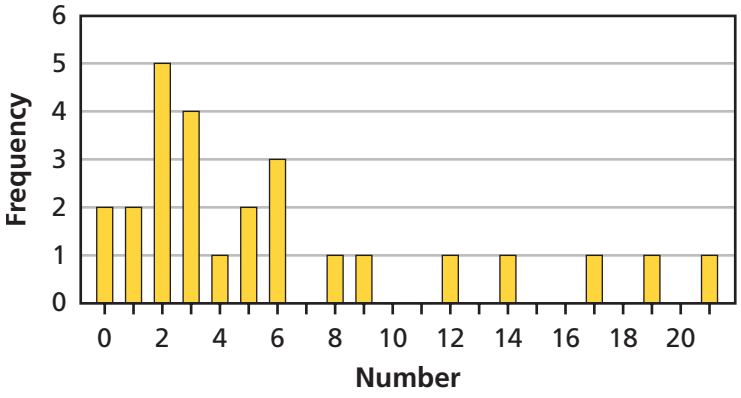
The students made tables to show the tallies or frequencies. Then they made bar graphs to display the data distributions.

Do you think the students surveyed live in a city, the suburbs, or the country? Explain.

Number of Pets

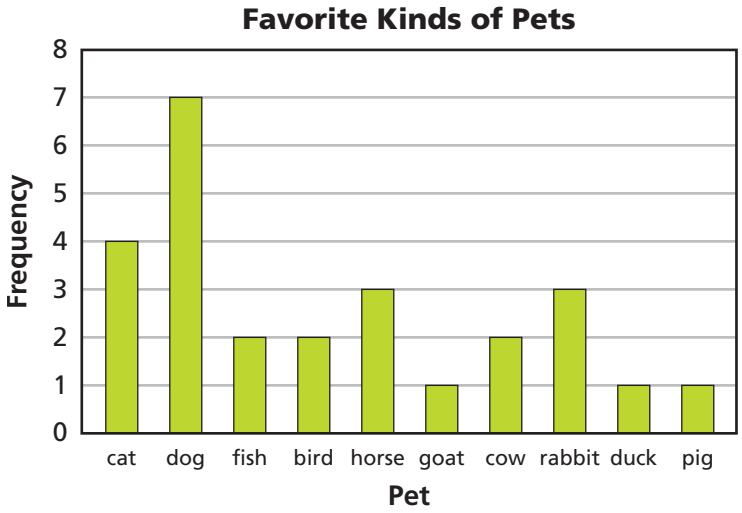
Number	Frequency
0	2
1	2
2	5
3	4
4	1
5	2
6	3
7	0
8	1
9	1
10	0
11	0
12	1
13	0
14	1
15	0
16	0
17	1
18	0
19	1
20	0
21	1

Number of Pets



Favorite Kinds of Pets

Pet	Frequency
cat	4
dog	7
fish	2
bird	2
horse	3
goat	1
cow	2
rabbit	3
duck	1
pig	1



Problem 1.4 Using Different Data Types

Decide whether each question can be answered by using data from the graphs and tables the students created. If so, give the answer and explain how you got it. If not, explain why not and tell what additional information you would need to answer the question.

A. Which graph shows categorical data?

B. Which graph shows numerical data?

C. What is the total number of pets the students have?

D. What is the greatest number of pets a student has?

E. How many students are in the class?

F. How many students chose cats as their favorite kind of pet?

G. How many cats do students have as pets?

H. What is the mode for the favorite kind of pet?

I. What is the median number of pets students have?

J. What is the range of the numbers of pets students have?

K. Tomas is a student in this class. How many pets does he have?

L. Do the girls have more pets than the boys?

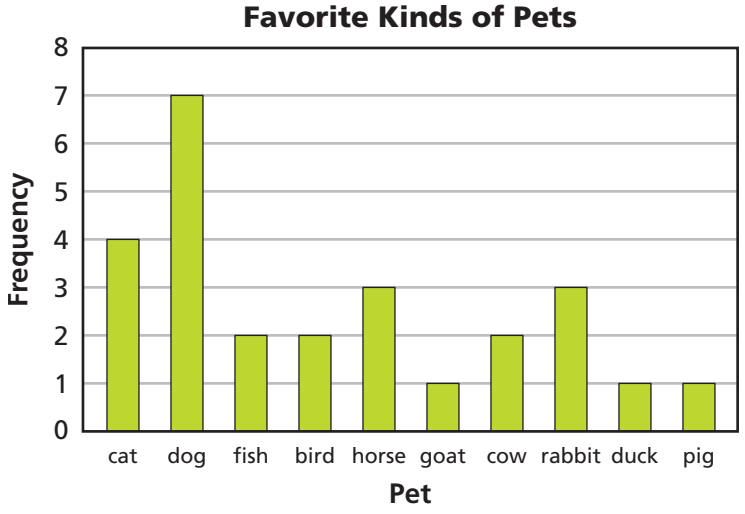
ACE **Homework starts on page 21.**

1.5 Vertical Bar Graphs and Horizontal Bar Graphs

You have used bar graphs to display distributions of data. *Vertical bar graphs* display data on the horizontal axis with vertical bars. On vertical bar graphs, the heights can be compared to the vertical frequency axis.

Look at the vertical bar graph below.

- What information does the horizontal axis show?
- What information does the vertical axis show?



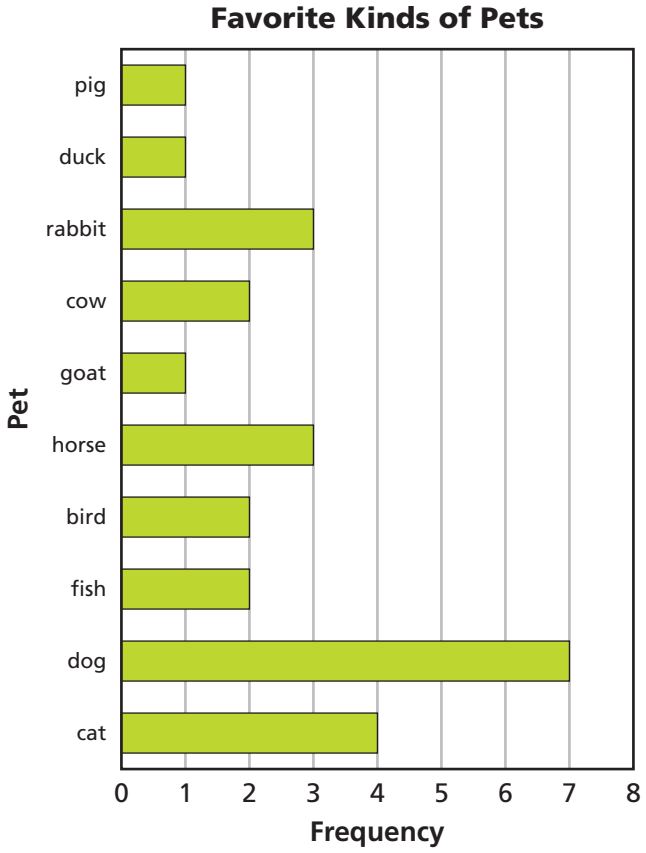
- How do you find out how many people chose "dog" as their favorite kind of pet using the vertical bar graph?

Suppose five more students are surveyed. Three identify birds as their favorite kind of pet. Two identify cats as their favorite kind of pet.

- What changes would you make in the vertical bar graph to show the new distribution?

Below is the distribution of the original pet data shown on a *horizontal bar graph*.

- Compare the vertical bar graph to the horizontal bar graph. How are they alike? How are they different?



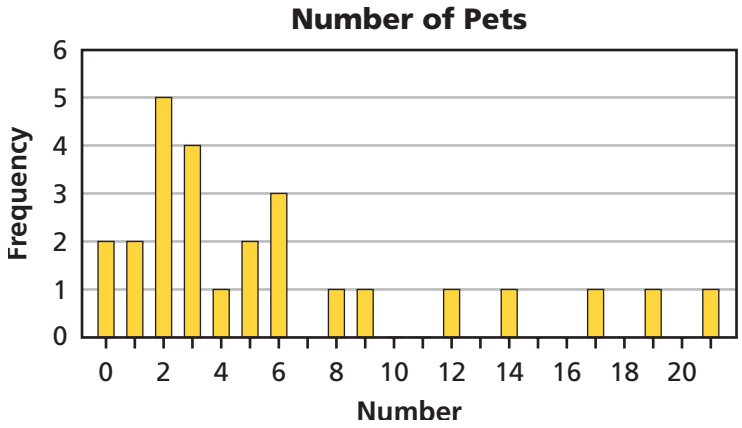
- How do you find out how many people chose "dog" as their favorite kind of pet using the horizontal bar graph?

Suppose five more students were surveyed. Three identify birds as their favorite kind of pet. Two identify cats as their favorite kind of pet.

- What changes would you make in the horizontal bar graph to show the new distribution?

Problem 1.5 Vertical Bar Graphs and Horizontal Bar Graphs

Below is a vertical bar graph showing the distribution of the number of pets students have.



A. Make a horizontal bar graph to show this distribution of data.

For each question below, explain:

- how you can find the answer to the question using the vertical bar graph
- how you can find the answer to the question using the horizontal bar graph

B. How many students in the class have more than five pets?

C. What is the least number of pets that any student in the class has?

D. What is the greatest number of pets that any student in the class has?

E. What is the median number of pets?

F. Three students were absent when these data were collected. Malcolm has 7 pets, Makana has 1 pet, and Jake has 3 pets. Add their data to each graph. What is the median number of pets now?

ACE Homework starts on page 21.