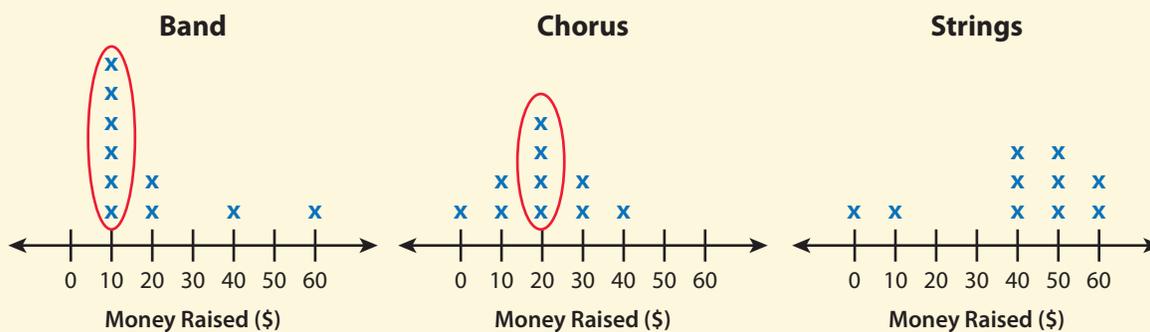


Measures of Center and Variability

Use What You Know

In the previous lesson, you learned that statistical data varies. In this lesson, you'll learn to explain variability in data. Take a look at this problem.

Students in band, chorus, and strings classes collected donations for a fundraiser concert to raise money for the music program. These line plots show the results of asking 10 students from each class how much money they collected.



Use the shape of each line plot to explain how the data vary.

Use the math you already know to solve this problem.

a. Between what numbers are the most common responses for each group of students?

Band _____ Chorus _____ Strings _____

b. Based on the line plots, which class likely raised the most money? Explain.

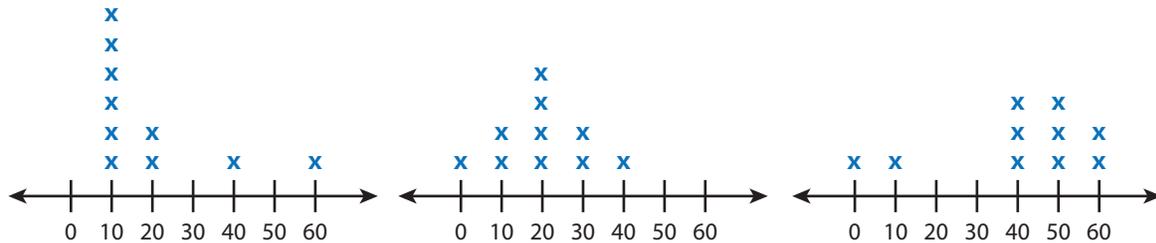
c. Based on the line plots, which class likely raised the least money? Explain.

d. Do you think these graphs would look different if 10 different students were surveyed from each class? Explain.

Find Out More

On the previous page, you looked at the shape of each line plot to think about what the data mean. You observed that the data for band were clustered around lesser values, and the data for strings were clustered around greater values. A **cluster** of data is a group of data points that crowd near each other.

Here are some other terms you can use to describe the shape of data points on a graph:



- A graph is **skewed left** when most of the data points are clustered near greater values.
- A graph is **skewed right** when most of the data points are clustered near lesser values.
- Symmetrical** graphs have the same shape on either side of a middle point.
- A **peak** forms when many data points are at one value.
- An **outlier** is a data point far away from the other data points; it doesn't quite fit with the rest of the data points.

Reflect

- Can an outlier value be part of a cluster of values? Explain.

Connect It Now you will solve the problem and explain how the mean describes the center of a data set.

2 Explain how to find the mean using the first model.

3 Look at the second *Model It*. What does the 200 represent? The 10? Why do you divide 200 by 10?

4 Are there any outliers in the data? What are they? How do you know?

5 Calculate the mean without outliers.

6 How do outliers affect the mean? Explain.

Try It Use what you just learned about mean to solve these problems. Show your work on a separate sheet of paper.

7 Here are the chorus data: {0, 10, 10, 20, 20, 20, 20, 30, 30, 40} What is the mean?

8 Here are the strings data: {0, 10, 40, 40, 40, 50, 50, 50, 60, 60} What is the mean?

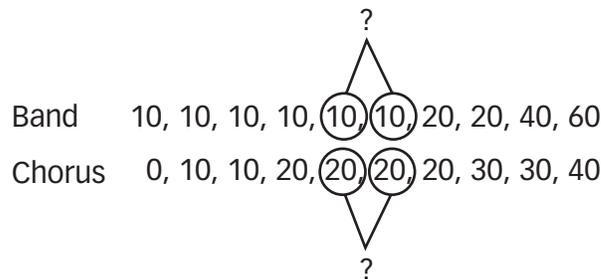
Learn About ▶ Median and Mode

Read the problem below. Then explore how to describe the center of data sets using median and mode.

You found that the band and chorus data both have a mean of \$20. Find the **median**, the middle number, and the **mode**, the most common number, to compare the center of each data set in a different way.

Model It You can use the median to describe the center of a data set.

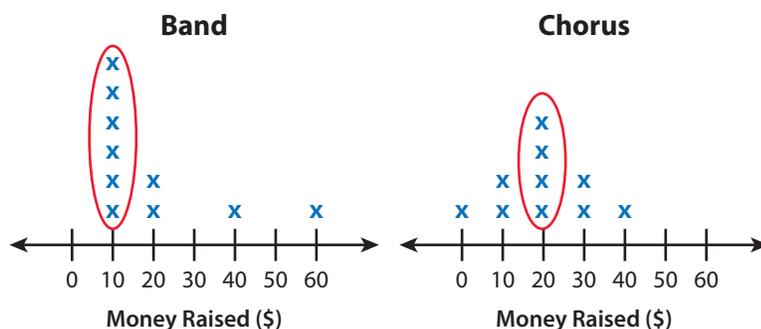
To find the **median**, or middle number, you can order the data points from least to greatest. Then, you can circle the middle number. Half of the values in a data set are above the median and half are below the median.



When there are two middle numbers, the median is the mean, or average, of the two numbers.

Model It You can use the mode to describe the center of a data set.

To find the **mode**, or most common number, you can look at the line plots to see which number in each data set appears the most often.



There can be more than one value that appears the most often.

Connect It Now you will solve the problem using the models and explain how median and mode describe the center of a data set.

9 Look at the first *Model It*. What is the median for the band data? the chorus data? Explain how you know.

10 Why does it make sense that the mean for the band data is greater than the median?

11 Why do you put the numbers in order from least to greatest to find the median?

12 Look at the second *Model It*. What is the mode for the band data? the chorus data?

13 Explain the meaning of both the median and the mode within this problem situation.

Try It Use what you just learned about the median and mode as a measure of center to solve this problem. Show your work on a separate sheet of paper.

14 Look at the strings data. {0, 10, 40, 40, 40, 50, 50, 50, 60, 60} Find the median and mode. Is the mean greater than or less than the median? Explain.

Learn About Variability

Read the problem below. Then explore how to describe the spread of data sets.

Jess knows the range of any data set is the difference between the greatest and least values. The range of a data set is a measure of the data set's variability, also called its spread. Is there another way to measure the spread of a data set? If so, how is it done?

Model It You can find the mean absolute deviation (MAD) to describe the spread of data points from the mean.

Range and MAD describe the variability of a data set in different ways.

- **Range** is the difference between the greatest and least values.
- **MAD** is the average of the distances of each data point from the mean.

To find the MAD:

1. Find the deviation, or distance, of each data value from the mean.
2. Find the absolute value of the deviation of each data value from the mean.
3. Find the average of these absolute deviations.

Band		
Data Value	Deviation (distance) from Mean	Absolute Deviation
10	-10	10
10	-10	10
10	-10	10
10	-10	10
10	-10	10
10	-10	10
20	0	0
20	0	0
40	20	20
60	40	40
MAD:		$\frac{120}{10} = 12$

The MAD of 12 means that, on average, every data point is \$12 from the mean.

Strings		
Data Value	Deviation (distance) from Mean	Absolute Deviation
0	-40	40
10	-30	30
40	0	0
40	0	0
40	0	0
50	10	10
50	10	10
50	10	10
60	20	20
60	20	20
MAD:		$\frac{140}{10} = 14$

On average, every data point is \$14 from the mean.

Connect It Now you will solve the problem using the model and describe the spread of data sets.

15 Look at the tables in *Model It*.

What does a negative deviation mean? _____

What does a positive deviation mean? _____

What does 0 deviation mean? _____

16 Why do you take the absolute value of the deviation?

17 Compare the band and strings MAD values. Which data set had slightly less variability in data points from the mean? Explain how you know.

18 Which indicates the greater degree of variability in a data set: A high MAD value or a low MAD value? Explain your answer.

Try It Use what you just learned about describing the spread of data to solve this problem. Show your work on a separate sheet of paper.

19 Look at the chorus data: {0, 10, 10, 20, 20, 20, 20, 30, 30, 40}

A. Find the range. What does the range mean?

B. The mean is \$20. Find the MAD. What does the MAD value mean?

Practice  **Finding Measures of Center and Variability**

Study the example below. Then solve problems 20–22.

Example

The table below shows the 5 highest temperatures recorded in the U.S. What is the mean of these data points?

Death Valley, California	134 °F
Lake Havasu City, Arizona	128 °F
Laughlin, Nevada	125 °F
Lakewood, New Mexico	122 °F
Alton, Kansas	121 °F

Look at how you can show your work with calculations.

Add the data values.

$$134 + 128 + 125 + 122 + 121 = 630$$

Divide by the total number of data values.

$$630 \div 5 = 126$$

Solution The mean is 126 °F



To find the mean, this student added the data values and divided by the total number of data values.

**Pair/Share**

Can the mean ever be equal to the greatest value in a data set?

- 20** 15 sixth graders and 15 seventh graders were asked: How many extracurricular activities do you participate in? Here are the survey results:

6th grade: {0, 0, 1, 1, 1, 1, 2, 2, 2, 2, 2, 3, 3, 3, 4, 4}

7th grade: {1, 1, 1, 1, 1, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4}

Find the mean for each data set. Based on these survey results, in which grade are students more likely to participate in activities?

Show your work.



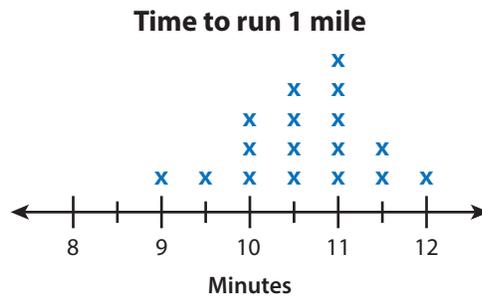
Look at the numbers in each data set. Can you predict which grade will have a greater mean?

**Pair/Share**

What is the median for each grade? How does it compare to the mean?

Solution _____

- 21 The line plot below shows how long it took students in a P.E. class to run 1 mile.



Describe the shape of the graph.

Solution _____

- 22 Look at the data in the line plot in problem 21. Which of the statements below is FALSE? Circle the letter of the correct answer.

- A The median is 10.5
- B The mean is 10.5.
- C The mode is 11.
- D The range is 3.

Lisa chose **A** as the correct answer. Why might Lisa have chosen A?



What is the difference between a symmetric and a skewed graph?

Pair/Share

Are there any outliers?
 How do outliers change the center and spread of a graph?



How can you look at the line plots to find the median, mode, and range?

Pair/Share

How are peak and mode related when looking at a line plot?

- 4 In a marketing study, two different groups of 12 people previewed a new movie. They rated the movie from 10, the best, to 1, the worst. The data is shown below.

Group A: 8, 7, 1, 6, 8, 5, 5, 8, 8, 1, 7, 8

Group B: 8, 7, 1, 6, 5, 5, 7, 2, 8, 1, 7, 6

Which statement **must** be true? Circle all that apply.

- A The mode of Group A exceeds the mode of Group B by 1.
 - B The mean of Group A exceeds the mean of Group B by 1.
 - C The median of Group A is equal to the median of Group B.
 - D The range of Group A is equal to the range of Group B.
- 5 Ten sixth graders were asked two questions. Below are the questions and survey results.

Question 1: How many hours per day do you spend playing outside?

{0, 0, 0, 0, 1, 1, 1, 1, 2, 3}

Question 2: How many hours per day do you spend using an electronic device?

{0, 2, 4, 4, 5, 6, 6, 7, 8, 8}

Part A Draw a line plot for each set of data. Then describe the shape of each line plot.

Part B Find the mean of the data sets.

Part C On average, how much more time do sixth graders spend playing with electronic devices than they do playing outside?

 **Self Check** Go back and see what you can check off on the Self Check on page 273.