6th grade Math At-Home Learning Packet

Notes & Assignments for 5.11-5.22

Lesson: 10.4 Interquartile Range and Box Plots
- ○ p. 561-566
- ○ Work through the lesson: complete Learn- Examples- Check- Apply

Assignment: Practice 10.4 <u>Friday 5.15</u>
- ● P.567-568
- ● You will need to submit this assignment to your teacher

Lesson 10.6 Dot Plots and Histograms
- ○ p. 575-580
- ○ Work through the lesson: complete Learn- Examples- Check- Apply

Assignment: Practice 10.6 <u>Due Friday 5.22</u>
- ● P.581-582
- ● You will need to submit this assignment to your teacher

Interquartile Range and Box Plots

I Can... understand how a measure of variation describes the variability of a data set with a single value, display a numerical data set in a box plot, and summarize the data.
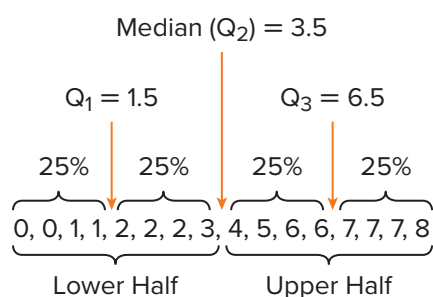
Learn Measures of Variation

Measures of variation are values that describe the variability, or spread, of a data set. They describe how the values of a data set vary with a single number.

One measure of variation is the **range**, which is the difference between the greatest and least data values in a data set. Consider the data set shown.

0, 0, 1, 1, 2, 2, 2, 3, 4, 5, 6, 6, 7, 7, 7, 8   The data values range from 0 to 8. The range is 8–0, or 8.

Another measure of variation is the *interquartile range*. Before you can find this measure, you first need to understand and find quartiles. **Quartiles** divide the data into four equal parts. The **first quartile**, Q_1, is the median of the data values less than the median. The **third quartile**, Q_3, is the median of the data values greater than the median. The median is also known as the **second quartile**, Q_2.

Median $(Q_2) = 3.5$

$Q_1 = 1.5$ $Q_3 = 6.5$

| 25% | 25% | 25% | 25% |

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

Lower Half   Upper Half

> The median divides the data into two halves. The quartiles divide the data into fourths. Each fourth represents 25% of the data.

The **interquartile range (IQR)** is the distance between the first and third quartiles of the data set. To find the IQR, subtract the first quartile from the third quartile.

The interquartile range represents the middle half, or middle 50%, of the data. The lower the IQR is for a data set, the closer the middle half of the data is to the median.

IQR $= 6.5 - 1.5$, or 5

$Q_1 = 1.5$ $Q_3 = 6.5$

| 25% | 25% | 25% | 25% |

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

50%
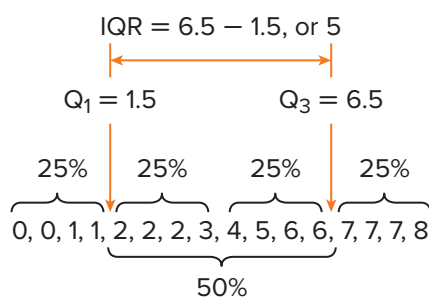
In the given data set, the IQR is $6.5 - 1.5$, or 5.

What Vocabulary Will You Learn?
box plot
first quartile
interquartile range
measures of variation
quartiles
range
second quartile
third quartile

💬 Talk About It!
How does knowing that the data is divided into four equal parts help you remember the vocabulary term quartile?

~~See students' responses.~~

💬 Talk About It!
If the median describes the center of a data set, what does the interquartile range describe?

~~Sample answer: The interquartile range describes how spread out the middle 50% of the values are around the median.~~

🌐 **Example 1** Find the Range and Interquartile Range

The table shows the approximate maximum speeds, in miles per hour, of different animals.

Use the range and interquartile range to describe how the data vary.

Animal	Speed (mph)
Housecat	30
Cheetah	70
Elephant	25
Lion	50
Mouse	8
Spider	1

Part A Describe the variation of the data using the range.

The greatest speed in the data set is 70 miles per hour. The least speed in the data set is 1 mile per hour.

The range is 70 − 1, or 69 miles per hour.

The speeds of animals vary by 69 miles per hour.

Part B Describe the variation of the data using the interquartile range.

Step 1 Find the median.

Write the speeds in order from least to greatest.

1	8	25	30	50	70

least   **greatest**

The median is __27.5__. Find the mean of the two middle numbers, 25 and 30.

Step 2 Find the first and third quartiles.

The first quartile is __8__. Find the median of the lower half of the data.

The third quartile is __50__. Find the median of the upper half of the data.

Step 3 Find the interquartile range.

Interquartile range = $Q_3 - Q_1$

$$= \boxed{50} - \boxed{8} \qquad Q_3 = 50; Q_1 = 8$$

$$= \boxed{42} \qquad \text{Subtract.}$$

So, the spread of the middle 50% of the data is __42__. This means that the middle half of the data values vary by __42__ miles per hour.

Check

The average wind speeds for several cities in Pennsylvania are given in the table. Use the range and interquartile range to describe how the data vary.

Wind Speed	
City	Speed (mph)
Allentown	8.9
Erie	11.0
Harrisburg	7.5
Middletown	7.7
Philadelphia	9.5
Pittsburgh	9.0
Williamsport	7.6

Part A

Describe the variation of the data using the range.

Show your work here — **The data vary by a range of 3.5 miles per hour.**

Part B

Describe the variation of the data using the interquartile range.

Show your work here — **The middle half of the data values vary by 1.9 miles per hour.**

Go Online You can complete an Extra Example online.

Learn Construct Box Plots

A **box plot**, or box-and-whisker plot, uses a number line to show the distribution of a data set by plotting the median, quartiles, and extreme values. The extreme values, or extremes, are the greatest and least values in the data set. The extremes, quartiles, and median are referred to as the *five-number summary*.

A box is drawn around the two quartile values. The whiskers extend from each quartile to the extreme data values, unless the extremes are very far apart from the rest of the data set. The median is marked with a vertical line, and separates the box into two boxes.

lower extreme Q_1 median Q_2 upper extreme
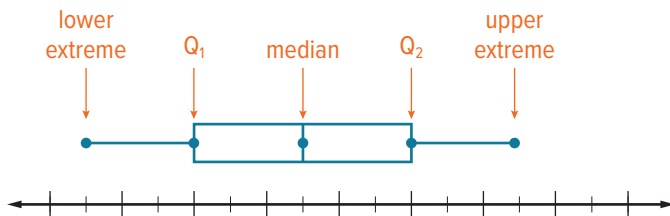
Box plots separate data into four sections. These sections are visual representations of quartiles. Even though the parts may differ in length, each contain 25% of the data. The two boxes represent the middle 50% of the data. A longer box or whisker indicates the data are more spread out in that section. A longer box or whisker does not mean there are more data values in that section. Each section contains the same number of values, 25% of the data.

Math History Minute

Florence Nightingale (1820–1910) used statistics to help improve the survival rates of hospital patients. She discovered that by improving sanitation, survival rates improved. She designed charts to display the data, as statistics had rarely been presented with charts before. She is known for inventing the *coxcomb graph,* which is a variation of the circle graph.

🌐 **Example 2** Interpret Box Plots

The box plot shows the annual snowfall totals, in inches, for a certain city over a period of 20 years.

Annual Snowfall (in.)

Describe the distribution of the data. What does it tell you about the snowfall in this city?

The annual snowfall ranges from about 110 inches to about 250 inches. The middle half of the data range from about 140 inches to about 195 inches. Because the boxes are shorter than the whiskers, there is less variation among the middle half of the data. Having less variation means there is a greater consistency among the middle 50% of the data than in either whisker.

Check

The average gas mileage, in miles per gallon, for various sedans is shown in the box plot. Describe the distribution of the data. What does it tell you about the average gas mileage for those sedans?

Average Gas Mileage (mpg)
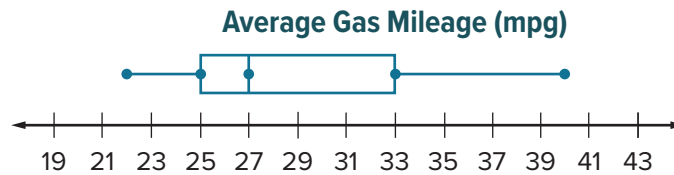
> Show your work here

Sample answer: The average gas mileage ranges from about 22 to 40 mpg. The middle half of the data range from 25 to 33 mpg. The median is 27 mpg; half of the sedans have a gas mileage above 27 mpg and half of them have a gas mileage below 27 mpg. Because the left whisker and left box are shorter than the right whisker and right box, there is less variation among the lower half of the data. Having less variation means there is a greater consistency among the gas mileages in the sedans that have a gas mileage under 27 mpg.

💬 **Think About It!**

What does the length of the box and whiskers tell you about the spread of the data in the box plot?
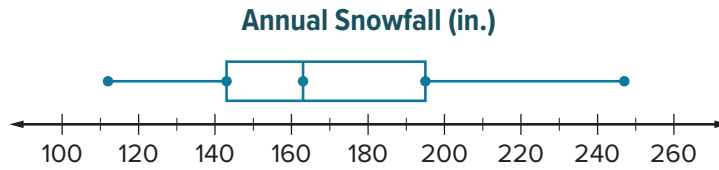
See students' responses.

💬 **Talk About It!**

What does the interquartile range describe in the context of the problem?

Sample answer: The middle 50% of the data is clustered between about 140 inches and 195 inches. So, in half of the years, the city received between 140 inches and 195 inches of snow.

🌐 **Go Online** You can complete an Extra Example online.

🌐 Example 3 Construct and Interpret Box Plots

The table shows the recorded speeds of cars traveling on a country road.

Car Speeds (mph)									
25	35	27	22	34	40	20	19	23	25

Construct a box plot to represent the data. Then describe the distribution of the data.

Part A Construct a box plot.

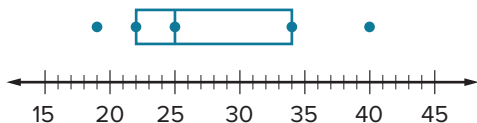**Step 1** Order the values from least to greatest.

In order from least to greatest, the speeds are 19, 20, 22, 23, 25, 25, 27, 30, 34, 35, and 40 miles per hour.

Step 2 Graph the values above a number line.

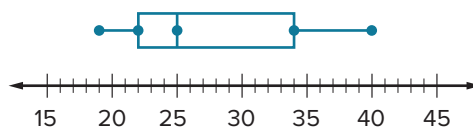Find the median, the extremes, and the first and third quartiles. Graph the values above a number line.

Median: 25

Q_1: 22 Q_3: 34

Lower Extreme: 19   Upper Extreme: 40

15 20 25 30 35 40 45
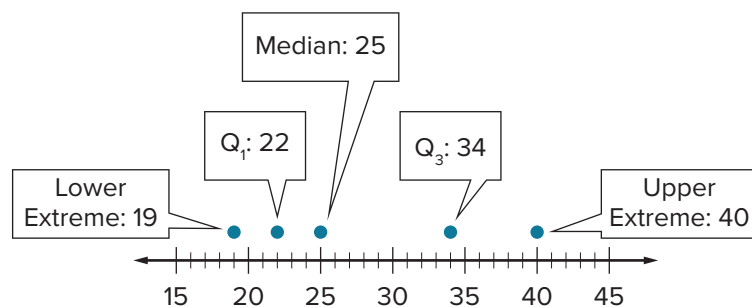
Step 3 Draw the box plot.

15 20 25 30 35 40 45

Draw a box around the first quartile and the third quartile. Draw a line through the median.

Car Speeds (mph)

15 20 25 30 35 40 45

Draw a line from the first quartile to the least value. Draw a line from the third quartile to the greatest value. Add a title.

Part B Describe the distribution of the data.

The recorded speeds range from 19 miles per hour to 40 miles per hour. The middle half of the data range from 22 miles per hour to 34 miles per hour. Because the boxes are longer than the whiskers, there is more variation among the middle half of the data. Having more variation means there is a lesser consistency among the middle 50% of the data than in either whisker.

Copyright © McGraw-Hill Education

🧠 Think About It!
What are the different measures of variation you need to find in order to construct a box plot?

quartiles, extremes, median

💬 Talk About It!
How does constructing a box plot to represent the data help you to understand the distribution of the data?

Sample answer: From the box plot, I can compare the sizes of the boxes and whiskers to make conclusions about the distribution of the data.
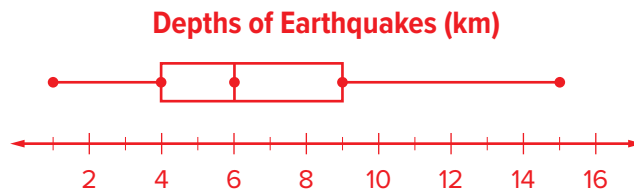
Check

Earthquakes occur at different depths below Earth's surface. Stronger earthquakes happen at depths that are closer to the surface. The table shows the depths of recent earthquakes, in kilometers.

Depth of Recent Earthquakes (km)						
5	15	1	11	2	7	3
9	5	4	9	10	5	7

Part A Construct a box plot to represent the data.

Show your work here

Depths of Earthquakes (km)

Part B Describe the distribution of the data.

Show your work here

Sample answer: The earthquake depths range from 1 km to 15 km. The middle half of the data range from 4 km to 9 km. The median is 6 km; half of the earthquakes occurred at a depth greater than 6 km and half of them occurred at a depth less than 6 km. Because the left whisker and left box are shorter than the right whisker and right box, there is less variation among the lower half of the data. Having less variation means there is a greater consistency among the earthquakes that have a depth of less than 6 km.

👆 **Go Online** You can complete an Extra Example online.

📖 **Foldables** It's time to update your Foldable, located in the Module Review, based on what you learned in this lesson. If you haven't already assembled your Foldable, you can find the instructions on page FL1.

Statistical Displays

dot plot	Best used to...
histogram	Best used to...
box plot	Best used to...

Copyright © McGraw-Hill Education

566 **Module 10** • Statistical Measures and Displays

THIS MATERIAL IS PROVIDED FOR INDIVIDUAL EDUCATIONAL PURPOSES ONLY AND MAY NOT BE DOWNLOADED, REPRODUCED, OR FURTHER DISTRIBUTED.

Practice 10.4 Interquartile Range and Box Plots Practice

1. Cameron surveyed her friends about the number of apps they use. The responses were 15, 16, 18, 9, 18, 4, 19, 20, 17, and 36 apps. Use the range and interquartile range to describe how the data vary. (Example 1)

2. The table shows the number of hours different animals spend sleeping per day. Use the range and interquartile range to describe how the data vary. (Example 1)

Time Animals Spend Sleeping (h)					
12	20	16	11	4	2

3. The box plot shows the ages of vice presidents when they took office. Describe the distribution of the data. What does it tell you about the ages of vice presidents? (Example 2)

Ages of US Vice Presidents
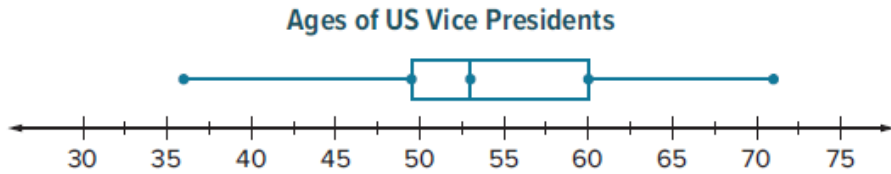
4. The ages of children taking a hip-hop dance class are 10, 9, 9, 7, 12, 14, 14, 9, and 16 years old. Construct a box plot of the data. Then describe the distribution of the data. (Example 3)

5. Open Response The cost of tents on sale at a sporting goods store are $66, $72, $78, $69, $64, $70, $67, $72, and $66. Use the range and interquartile range to describe how the data vary.

Apply

6. The table shows the number of points scored by the seventh and eighth grade girls basketball teams in each of their games this season. Construct a box plot to represent the data for each team. Then use the box plots to compare the data.

Points Scored per Game							
Seventh Grade Team				Eighth Grade Team			
39	36	40	27	34	36	47	40
35	29	36	29	39	38	45	43
31	38	30	34	42	41	45	42

7. Justify Conclusions Determine if the following statement is *true* or *false*. If *false*, justify your reasoning.

You can determine the mean of a data set from a box plot.

8. Create Provide a set of real-world data and then construct a box plot of the data.

9. Make an Argument A student said that, in a box plot, if the box to the right of the median is longer than the box to the left of the median, there are more data values represented by the longer box. Is the student's reasoning correct? Construct an argument to defend your solution.

10. Reason Inductively What can you conclude about a data set shown in a box plot where the whiskers and boxes are all the same length?

Outliers

I Can... understand how an outlier may affect a measure of center, and determine which measure of center is most appropriate to use when describing a data set that does or does not contain an outlier.

Learn Outliers

An **outlier** is a data value that is very far away from the other data values. It can be much greater in value or much less than the other values. Consider the data set shown.

225, 245, 295, 305, 360, 387, 388, 420, 470, 480, 625, 780

How do you know if either of the extreme values, 225 or 780, are considered outliers?

Most of the values are around 400. The value of 780 is an outlier because it is much bigger than most of the other values.

Below is a box-and-whisker of the data above. The outlier of 780 must be taken out of the data set when you create the box-and-whisker.

The box plot represents the data set. Outliers are indicated with an asterisk (*).
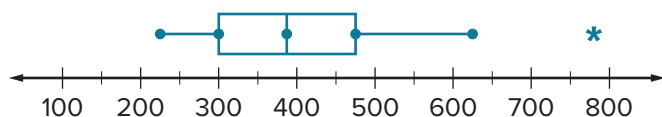
What Vocabulary Will You Learn?
outlier

💬 **Talk About It!**
If the outlier was removed from the data set, will the median still be 387.5? Why or why not?

no; Sample answer: There would now be an odd number of data values. The median would be 387.

Example 1 Identify Outliers

The ages, in years, of the candidates in an election are 55, 49, 48, 57, 23, 63, and 72.

Identify any outliers in the data set.

The value of 23 is much lower than the rest of the data. The outlier of the data is 23.

Check

The lengths, in feet, of various bridges are 354, 88, 251, 275, 727, and 1,121. Identify any outliers in the data.
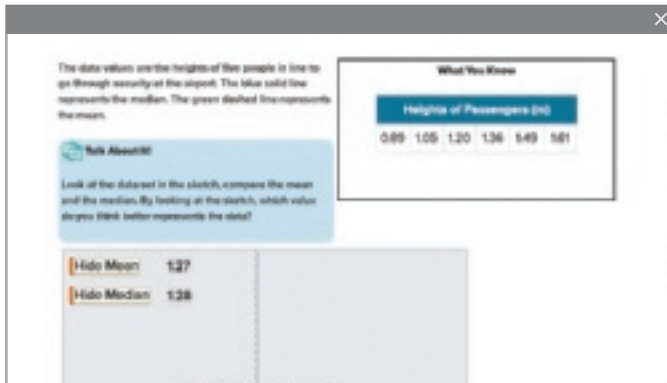
Show your work here

There are no outliers in the data set.

🌀 **Go Online** You can complete an Extra Example online.

Explore Mean, Median, and Outliers

Online Activity You will use Web Sketchpad to explore how outliers affect the mean and median.

Learn Describe the Effect of Outliers

If a data set contains an outlier, the outlier may affect the measures of center and/or variation.

Suppose you track the daily high temperatures for one week and the results are recorded in the table shown.

High Temperatures (°F)	
Sunday	72
Monday	68
Tuesday	71
Wednesday	74
Thursday	75
Friday	72

Suppose that the high temperature on Saturday is 42°F. This temperature is much lower than the other temperatures in the data set. It is also an outlier, because 42 is less than the lower limit for outliers.

To see how an outlier affects the measures of center and variation, calculate the measures both with and without the outlier.

Calculate the measures with the outlier.

Mean

$$\frac{42 + 68 + 71 + 72 + 72 + 74 + 75}{7} \approx 67.7$$

Median = 72

$Q_1 = 68$ $Q_3 = 74$

42, 68, 71, 72, 72, 74, 75

IQR = 74 − 68, or 6

Talk About It!

Suppose Saturday's temperature had been 59°F, which does not qualify as an outlier, but is cooler than the rest. How does this affect the mean? the median?
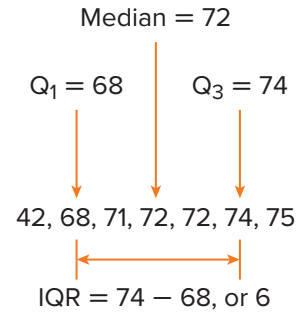
Sample answer: The mean and median without Saturday's temperature was 72, when Saturday's temperature was added, the mean decreased to 70, but the median was unchanged.
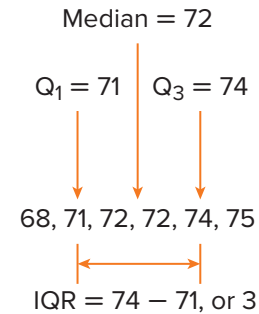
Median

The median is ___72___.

Interquartile Range (IQR)

The IQR is ___6___.

Calculate the measures without the outlier.

Mean

$$\frac{68 + 71 + 72 + 72 + 74 + 75}{6} = 72$$

Median = 72

$Q_1 = 71$ $Q_3 = 74$

68, 71, 72, 72, 74, 75

IQR = 74 − 71, or 3

Median

The median is ___72___.

Interquartile Range (IQR)

The IQR is ___3___.

The median was not affected by the inclusion of the outlier. Without the outlier, the mean,   , and IQR all increased in value. With the outlier, the mean is not the best representation of center, because most of the values are higher than 67.7.

Use either the mean or median when the data does not contain any outliers. Use only the median when the data contains an outlier. While the median might change a little when an outlier is included or removed, it does not change as much as the mean.

Use the corresponding measure of variation to describe the spread of the data.

- If you choose the median to describe the center, choose the IQR to describe the variation.

🌐 **Example 2** Describe the Effect of Outliers

The table shows the average lifespans of selected animals.

Calculate the mean and median with and without the outlier, 200. Then choose the measure that best describes the center.

Average Lifespan	
Animal	**Lifespan (years)**
Elephant	35
Dolphin	30
Chimpanzee	50
Tortoise	200
Gorilla	30
Gray Whale	70
Horse	20

Step 1 Calculate the mean and median with the outlier. Round to the nearest tenth, if necessary.

Mean

$$\frac{35 + 30 + 50 + 200 + 30 + 70 + 20}{7} \approx 62.1$$

The mean lifespan is about __**62**__ years.

Median

> Show your work here

The median lifespan is __**35**__ years.

Step 2 Calculate the mean and median without the outlier. Round to the nearest tenth, if necessary.

Mean

$$\frac{35 + 30 + 50 + 30 + 70 + 20}{6} \approx 39.2$$

The mean lifespan is about __**39**__ years.

Median

> Show your work here

The median lifespan is __**32.5**__ years.

Step 3 Choose the measure that best describes the center.

The __**mean**__ was most affected by the inclusion of the outlier.
The __**median**__ changed very little.

So, the __**median**__ best describes the center of the data.

💭 **Think About It!**
Will the outlier affect the mean or the median more? Explain your reasoning.

See students' responses.

💬 **Talk About It!**
Explain why it makes sense that the lifespan of the animals listed in the table are centered around 32.5 or 35 years, rather than around 39 or 62 years.

Sample answer: Most of the animals have lifespans that are 50 years or less.

Check

The table shows the cooking temperatures for different recipes. Calculate the mean and median with and without the outlier. Round to the nearest tenth, if necessary. Then choose the measure that best describes the center.

Cooking Temperature (°F)			
175	325	325	350
350	350	400	450

Show your work here

mean with outlier: ≈ 340.6
median with outlier: 350
mean without outlier: ≈ 364.3
median without outlier: 350
Sample answer: The mean was affected the most by the outlier. The median temperature did not change when the outlier was removed, so it would be the best measure of center to represent the data set.

Go Online You can complete an Extra Example online.

Pause and Reflect

Create a graphic organizer that will help you study the concepts you learned today in class.

Record your observations here

See students' observations.

Practice 10-6: Outliers Practice

1. Last week, Joakim spent 40, 25, 60, 30, 35, and 40 minutes practicing the piano. Identify any outliers in the data. (Example 1)

2. Last month, a basketball team scored 83, 84, 85, 87, 89, 88, 67, 79, and 81 points in their games. Identify any outliers in the data. (Example 1)

3. Abrianna sold 20, 23, 18, 4, 17, 21, 15, and 56 boxes of cookies after different football games. Identify any outliers in the data. (Example 1)

4. Last week a certain pet store had 52, 72, 96, 21, 58, 40, and 75 paying customers. Identify any outliers in the data. (Example 1)

5. The prices of trees that Sahana bought are $46, $39, $40, $45, $44, $68, and $51. Calculate the mean and median with and without the outlier. Round to the nearest tenth, if necessary. Choose the measure that best describes the center. (Example 2)

6. The prices of backpacks are $37, $43, $41, $36, $44, and $70. Calculate the mean and median with and without the outlier. Round to the nearest tenth, if necessary. Choose the measure that best describes the center. (Example 2)

7. The table shows the number of points scored by a football team. Calculate the mean and median with and without the outlier. Round to the nearest tenth, if necessary. Choose the measure that best describes the center. Explain. (Example 2)

Points Scored by a Football Team			
14	20	3	9
18	35	21	24
7	12	31	68

8. **Multiple Choice** The table shows the number of points scored by the players in a trivia game. Which measure of center best represents the data?

Points Scored in a Trivia Game			
12	9	5	11
6	0	14	7

   A mean
   B median
   C mode
   D range

Apply

9. **Create** Generate a set of real-world data that contains two outliers.

10. **Justify Conclusions** The ages, in years, of participants in a relay race are 12, 15, 14, 13, 15, 12, 22, 16, and 11. Identify any outliers in the data set. Justify your response.

11. **Construct an Argument** Explain how an outlier may or may not affect the mean and median.

12. **Justify Conclusions** Does an outlier affect the range of a data set? Explain.