LESSON 29

Dear Family,

This week your student is exploring how to display and describe data collected to answer **statistical questions**. The data collected to answer a statistical question are expected to vary, or show **variability**.

You can collect data to answer this statistical question.

How old are the students in the swim class?

You can use a table to show the **frequency** of each data value, or the number of times each value occurs. A frequency table displays the **distribution** of the data.

Age	Frequency
9	|||
10	𝍸 |
11	||

Your student will be displaying and describing distributions of data sets like the one below.

This list shows the number of pets owned by each of 11 students.

1, 0, 0, 1, 0, 1, 1, 3, 1, 1, 0

➤ **ONE WAY** to display the data distribution is in a frequency table.

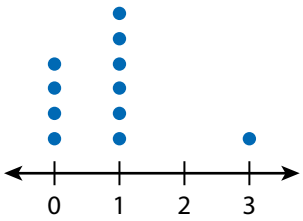
Number of Pets	Frequency
0	||||
1	𝍸 |
2	
3	|

You can describe the data as being spread out from 0 pets to 3 pets.

The data show that most people have 0 pets or 1 pet.

➤ **ANOTHER WAY** is to use a **dot plot**.

Each dot represents one data value.



The shape of the data distribution shows one large group of data points at 0 and 1, with a single point at 3.

Both displays show the data values in order from least to greatest and can help you describe the distribution of the data.

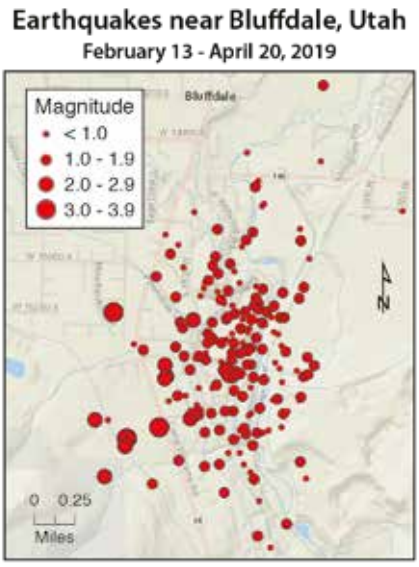
Use the next page to start a conversation about statistical questions.

Activity Thinking About Statistical Questions Around You

➤ **Do this activity together to investigate statistical questions in the real world.**

Did you know that scientists collect data to study earthquake patterns? They do this by using special equipment to measure the magnitude, or strength, of an earthquake each time one occurs.

Scientists use the data they collect to help predict when the next earthquakes may happen. Scientists ask and answer statistical questions such as *What magnitudes of earthquakes occur in this region?* and *How often do earthquakes of these magnitudes occur?*



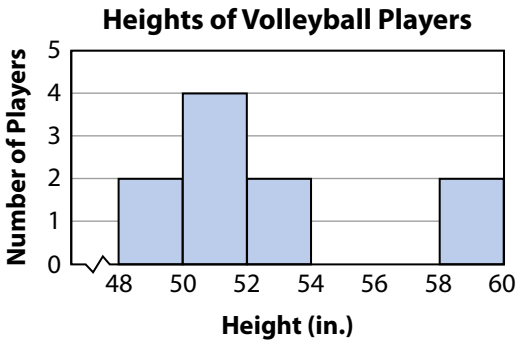
? Where else do you see statistical questions being used in the world around you?

LESSON 30

Dear Family,

This week your student is learning how to use dot plots and histograms to describe data distributions.

A **histogram** displays the frequency of data in equal-size intervals of the number line. This histogram shows 4 players whose heights are from 50 in. up to, but not including, 52 in. A player whose height is 52 in. would instead be counted in the next interval.



Your student will be learning how to solve problems like the one below.

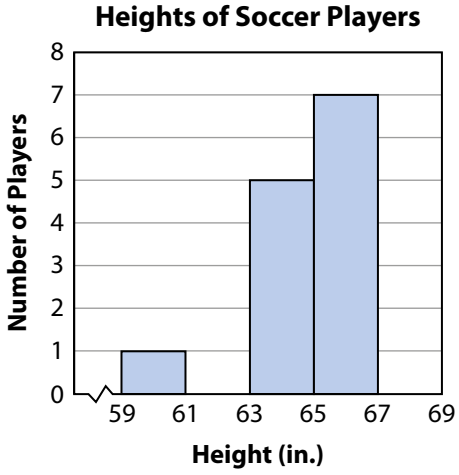
> The list shows the heights, in inches, of 13 soccer players. Display and describe the distribution of the data.
>
> 65, 63, 64, 59, 66, 65, 64, 66, 66, 64, 63, 65, 66

➤ **ONE WAY** to display a data distribution is with a frequency table.

Height (in.)	Frequency
59–60	|
61–62	
63–64	𝍸
65–66	𝍸 ||

No player is 61 in. or 62 in. tall. This is a **gap** in the data. Most of the data are in a **cluster** near the higher values. There is a **peak** at 65 in. to 66 in.

➤ **ANOTHER WAY** is to use a histogram.



When data are clustered near higher values, the distribution is **skewed left**.

Both displays can be used to describe the shape of the data distribution.

Use the next page to start a conversation about data displays.

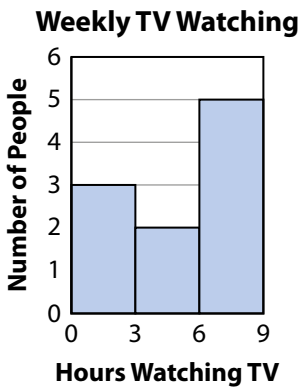
Use Dot Plots and Histograms to Describe Data Distributions

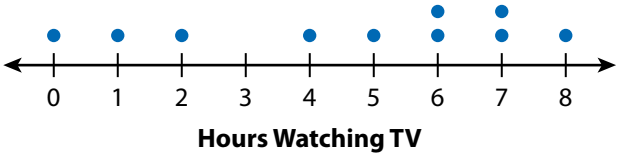
Activity Exploring Dot Plots and Histograms

➤ **Do this activity together to look for relationships between dot plots and histograms.**

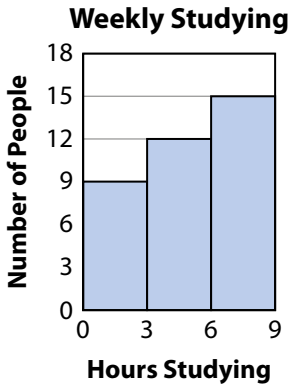
Dot plots and histograms can be used to display numerical data. There are two sets of dot plots and histograms below. The two graphs in each set represent the same data. What do you notice about each set?

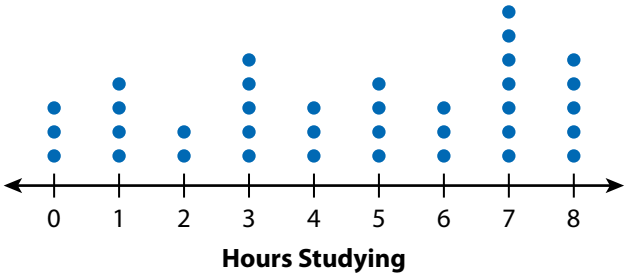
SET 1





SET 2





? How can you use each dot plot to make a histogram with a larger or smaller interval?

LESSON 31

Interpret Median and IQR in Box Plots

Dear Family,

This week your student is learning how to summarize a data set using a measure of center and a measure of variability. A **measure of center** is a single number that represents a typical value. A **measure of variability** describes how spread out the values in a data set are.

Your student will be learning to solve problems like the one below.

> The ages, in years, of 11 new members of a family fitness center are listed. How can you summarize the ages of the new members?
>
> 11, 5, 32, 7, 10, 41, 40, 15, 28, 80, 10

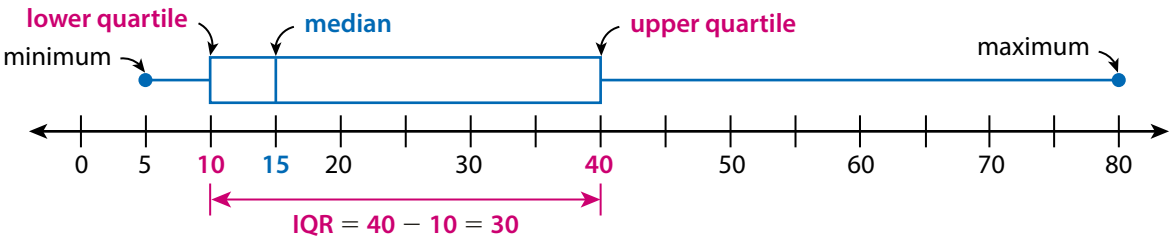
➤ **ONE WAY** to summarize a data set is with a measure of center.

The **median** of a data set is the middle number when the data values are listed from least to greatest.

5, 7, 10, 10, 11, (15), 28, 32, 40, 41, 80
median

➤ **ANOTHER WAY** to summarize a data set is with a measure of variability.

The **median** separates the data into two halves. The median of the lower half is the **lower quartile** and the median of the upper half is the **upper quartile**. You can display the data on a **box plot** to see how the data in each quarter are spread out. The box in the middle shows the **interquartile range (IQR)**.



You can use the median and IQR together to describe the distribution of the data. The median age of a new member is 15 years and 50% (or half) the ages are within the 30-year range of 10 years to 40 years.

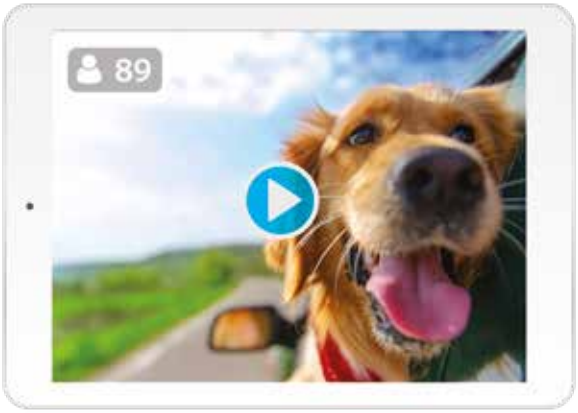
Use the next page to start a conversation about medians.

Activity Thinking About Medians Around You

➤ **Do this activity together to investigate medians in the real world.**

Online videos are a popular form of entertainment. Sometimes videos with a lot of views can earn money from showing advertisements. There are videos that have as many as one billion views!

Did you know that the median number of views for all videos uploaded in 2016 was only 89? This means half of these videos were viewed fewer than 89 times. That view count is much too low to earn any money!



? Describe a situation where you might want to know the median of a data set.

LESSON 32

Interpret Mean and Mean Absolute Deviation

Dear Family,

This week your student is learning how to find the mean and mean absolute deviation of a data set.

The **mean** of a data set is a measure of center that represents the average of the data values. The **mean absolute deviation (MAD)** is a measure of variability that represents the average distance of a data point from the mean. The MAD describes how spread out the values are from the mean value.

Your student will be learning to solve problems like the one below.

> The ages, in years, of 5 new members of a school architecture club are listed. How can you describe the ages of the new members?
>
> 12, 12, 16, 15, 15

➤ **ONE WAY** to describe, or summarize, a data set is with a measure of center.

To find the **mean** of a data set, divide the **sum of the data values** by the **number of data values**.

$$\frac{12 + 12 + 16 + 15 + 15}{5} = \frac{70}{5} = 14$$

➤ **ANOTHER WAY** to summarize a data set is with a measure of variabiity.

To find the average distance of a data value from the mean, you must first find each data value's distance from the **mean**, 14.

Data Value	12	12	16	15	15
Distance from 14	2	2	2	1	1

Now look at the data set in the second row of the table. Find the mean of this data set. Divide the **sum of the distances from the mean** by the **number of distances**. The result is the **mean absolute deviation (MAD)** of the *original* data set.

$$\frac{2 + 2 + 2 + 1 + 1}{5} = \frac{8}{5} = 1.6$$

You can use the mean and MAD together to describe the distribution of the data. The mean age of a new member is 14 years. On average, a new member's age is within 1.6 years of 14 years.

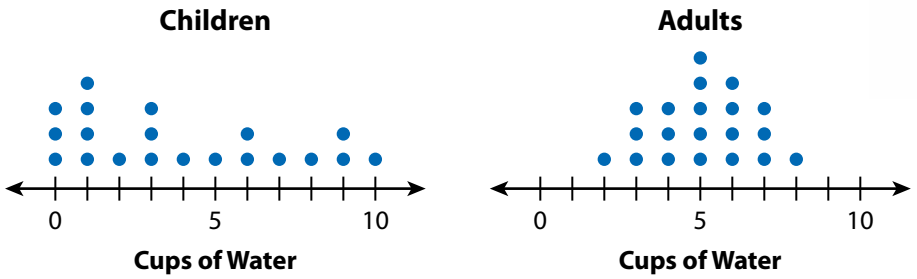
Use the next page to start a conversation about measures of variability.

Activity Thinking About Variability Around You

➤ **Do this activity together to investigate variability in the real world.**

Drinking water helps you to stay hydrated, but did you know that drinking enough water has other benefits? Water also helps to bring nutrients and oxygen to all the cells in your body. Drinking water even helps you digest your food!

Look at the two dot plots. One dot plot shows how many cups of water 20 children drank in one day. The other dot plot shows how many cups of water 20 adults drank in one day. There is a lot more variability in the number of cups of water children drink than there is in the number of cups of water adults drink.



? Why do you think there might be more variability in the amount of water children drink than adults drink?

LESSON 33

Dear Family,

This week your student is learning about choosing measures of center and variability to summarize and compare data sets.

Your student will be learning to solve problems like the one below.

> Alberto and Hannah are practicing throwing fastballs for a baseball game. The lists show the two pitchers' practice speeds in miles per hour.
>
> Alberto: 63, 65, 71, 62, 60, 62, 71, 65, 69, 67, 64
>
> Hannah: 57, 66, 73, 62, 56, 73, 71, 74, 55, 70, 68
>
> How can you compare Alberto's and Hannah's fastball pitching speeds?

➤ **ONE WAY** to compare the two data sets is to compare medians and IQRs.

Use the **median** to represent a typical fastball speed. Use the **lower** and **upper quartiles** to find the **interquartile range (IQR)**.

Alberto

60, 62, **62**, 63, 64, **65**, 65, 67, **69**, 71, 71 IQR = **69** − **62** = **7**

Hannah

55, 56, **57**, 62, 66, **68**, 70, 71, **73**, 73, 74 IQR = **73** − **57** = **16**

➤ **ANOTHER WAY** is to compare the means and MADs of the data sets.

Use the **mean**, or average, to represent a typical fastball speed. Use the distances of the speeds from the mean to find the average distance from the mean, or the **mean absolute deviation (MAD)**.

Alberto

Mean = $\frac{719}{11}$, or about **65.4** MAD = $\frac{33.2}{11}$, or about **3**

Hannah

Mean = $\frac{725}{11}$, or about **65.9** MAD = $\frac{67.3}{11}$, or about **6**

Using either pair of measures of center and variability, you can see that Hannah's typical fastball is faster than Alberto's and Alberto's fastball speeds show less variability. Hannah may often pitch faster than Alberto, but Alberto is the more consistent pitcher.

Use the next page to start a conversation about measures of center.

Activity Choosing Measures of Center

➤ **Do this activity together to examine choices between measures of center.**

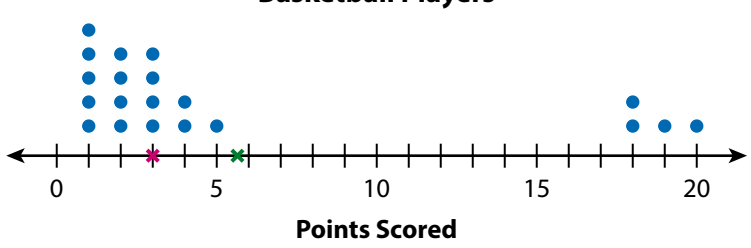
The shape of a distribution can help you choose either the mean, the median, or both as a measure of center that summarizes the data. For the data sets below, the circled measures of center are good choices for a typical value of the data set.

What do you notice about how the choice of measure of center relates to the shape of the distribution?

DATA SET 1

This dot plot shows how many points each player on a basketball team scored during the season.

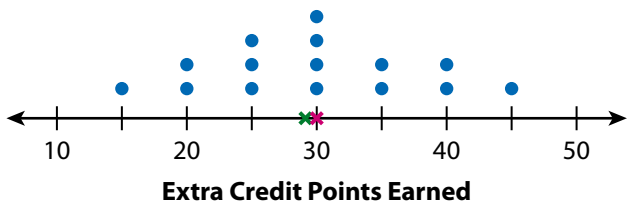
(Median = 3) Mean = 5.65



DATA SET 2

This dot plot shows how many extra credit points math students earned in one month.

(Median = 30) (Mean = $29\frac{2}{3}$)



? How would you describe the variability of the two data sets?