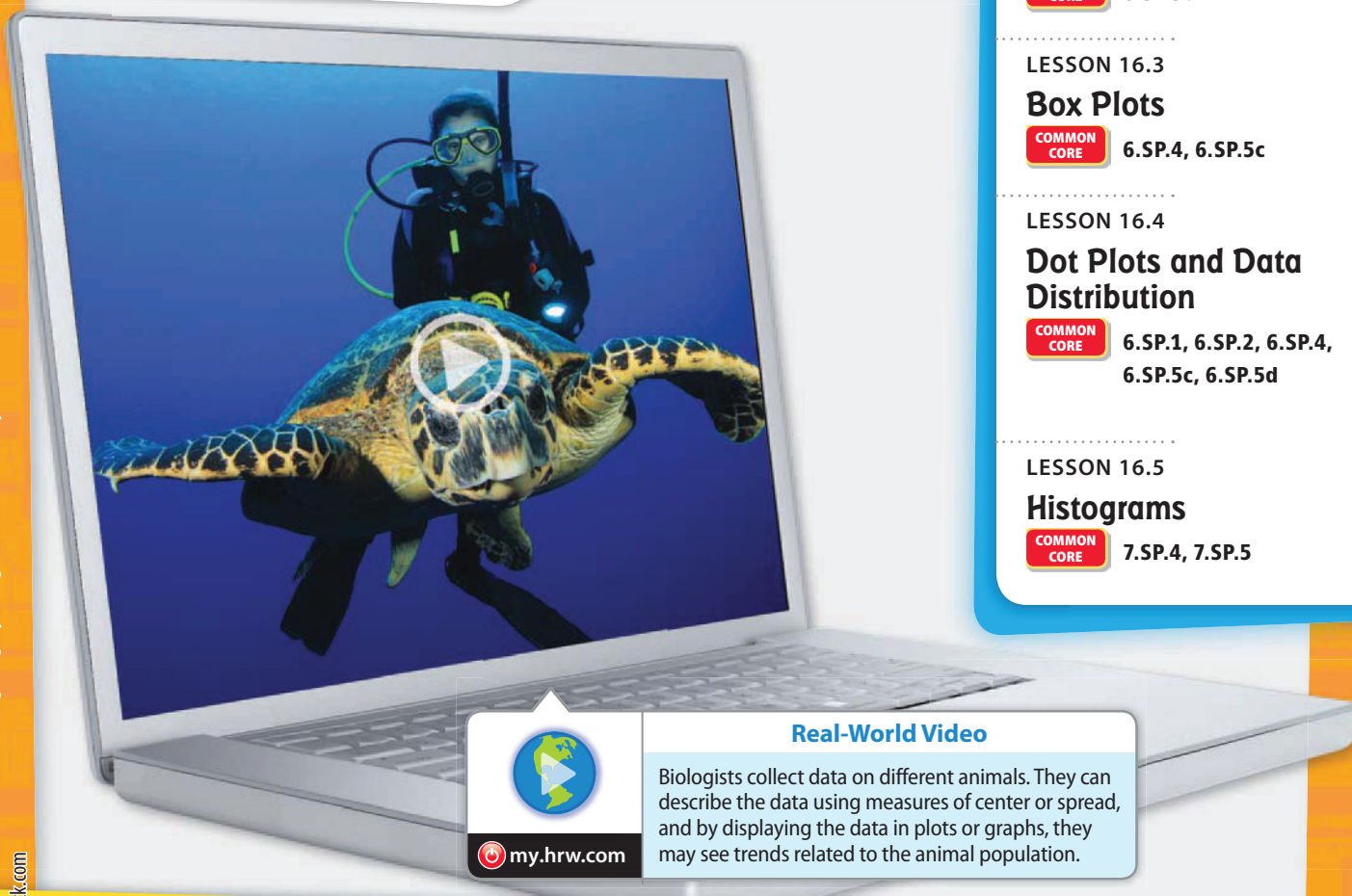


Displaying, Analyzing, and Summarizing Data



ESSENTIAL QUESTION

How can you solve real-world problems by displaying, analyzing, and summarizing data?



MODULE



16

LESSON 16.1

Measures of Center



6.SP.3, 6.SP.5,
6.SP.5a, 6.SP.5b,
6.SP.5c, 6.SP.5d

LESSON 16.2

Mean Absolute Deviation



6.SP.5c

LESSON 16.3

Box Plots



6.SP.4, 6.SP.5c

LESSON 16.4

Dot Plots and Data Distribution



6.SP.1, 6.SP.2, 6.SP.4,
6.SP.5c, 6.SP.5d

LESSON 16.5

Histograms



7.SP.4, 7.SP.5



my.hrw.com

Real-World Video

Biologists collect data on different animals. They can describe the data using measures of center or spread, and by displaying the data in plots or graphs, they may see trends related to the animal population.

© Houghton Mifflin Harcourt Publishing Company • Image Credits: © Rich Carey/Shutterstock.com

GO
DIGITAL
my.hrw.com



my.hrw.com

Go digital with your write-in student edition, accessible on any device.



Math On the Spot

Scan with your smart phone to jump directly to the online edition, video tutor, and more.



Animated Math

Interactively explore key concepts to see how math works.



Personal Math Trainer

Get immediate feedback and help as you work through practice sets.

Are YOU Ready?

Complete these exercises to review skills you will need for this module.



Personal Math Trainer

Online Assessment and Intervention

my.hrw.com

Remainders

EXAMPLE

$$\begin{array}{r} 7.25 \\ 12 \overline{)87.00} \\ \underline{84} \\ 30 \\ \underline{-24} \\ 60 \\ \underline{-60} \\ 0 \end{array}$$

Write a decimal point and a zero in the dividend.

Place a decimal point in the quotient.

Add more zeros to the dividend if necessary.

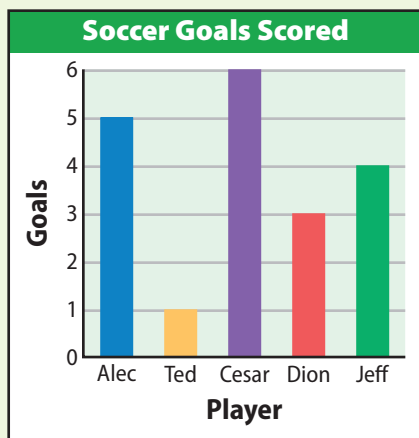
Find the quotient. Write the remainder as a decimal.

1. $15 \overline{)42}$ _____ 2. $75 \overline{)93}$ _____ 3. $52 \overline{)91}$ _____ 4. $24 \overline{)57}$ _____

Read Bar Graphs

EXAMPLE

How many goals did Alec score?



The first bar shows how many goals Alec scored.

The bar extends to a height of 5.

Alec scored 5 goals.

5. How many goals did Dion score? _____
6. Which two players together scored the same number of goals as Jeff? _____
7. How many fewer goals than Cesar did Alec score? _____

Reading Start-Up

Visualize Vocabulary

Use the review words to complete the chart.

Introduction to Statistics		
Definition	Example	Review word
A group of facts	The grades of all of the students in a school	
A tool used to gather information from individuals	A questionnaire given to all students to find the number of hours each student spends studying in 1 week	
A value that summarizes a set of unequal values, found by addition and division	Results of the survey show that students typically spend 5 hours a week studying	

Understand Vocabulary

Complete the sentences using the preview words.

1. The average of a data set is the _____.
2. The _____ is the middle value of a data set.
3. The number or category that occurs most frequently in a data set is the _____.

Vocabulary

Review Words

average (*promedio*)
data (*datos*)
survey (*encuesta*)

Preview Words

box plot (*diagrama de caja*)
categorical data (*datos categóricos*)
dot plot (*diagrama de puntos*)
histogram (*histograma*)
interquartile range (*rango entre cuartiles*)
lower quartile (*cuartil inferior*)
✓ mean (*media*)
mean absolute deviation (MAD) (*desviación absoluta media, (DAM)*)
✓ median (*mediana*)
measure of center (*medida central*)
measure of spread (*medida de dispersión*)
✓ mode (*moda*)
range (*rango*)
statistical question (*pregunta estadística*)
upper quartile (*cuartil superior*)

Active Reading

Layered Book Before beginning the module, create a layered book to help you learn the concepts in this module. Label each flap with lesson titles from this module. As you study each lesson, write important ideas, such as vocabulary and formulas under the appropriate flap. Refer to your finished layered book as you work on exercises from this module.





Unpacking the Standards

Understanding the standards and the vocabulary terms in the standards will help you know exactly what you are expected to learn in this module.

COMMON CORE 6.SP.5C

Summarize numerical data sets in relation to their context, such as by giving quantitative measures of center (median and/or mean) and variability (interquartile range and/or mean absolute deviation), as well as describing any overall pattern and any striking deviations from the overall pattern with reference to the context in which the data were gathered.

What It Means to You

You will use measures of center to describe a data set.

UNPACKING EXAMPLE 6.SP.5C

Several students' scores on a history test are shown. Find the mean score and the median score. Which measure better describes the typical score for these students? Explain.

History Test Scores							
73	48	88	90	90	81	83	

$$\text{Mean: } \frac{73 + 48 + 88 + 90 + 90 + 81 + 83}{7} = \frac{553}{7} = 79$$

To find the median, write the data values in order from least to greatest and find the middle value.

Median: 48 73 81 **83** 88 90 90

The median better describes the typical score.

The mean is affected by the low score of 48.

COMMON CORE 6.SP.4

Display numerical data in plots on a number line, including dot plots, histograms, and box plots.

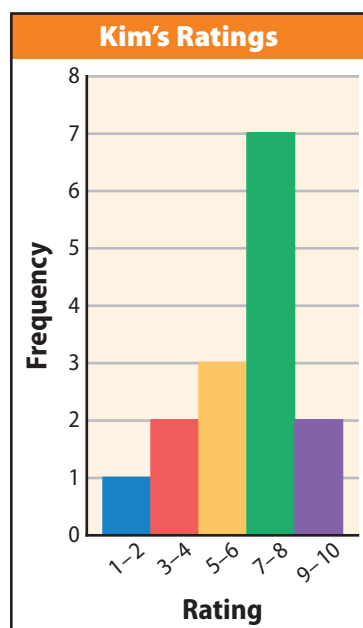
What It Means to You

You will interpret the data from a dot plot, histogram, or box plot.

UNPACKING EXAMPLE 6.SP.4

Kim has started rating each movie she sees using a scale of 1 to 10 on an online site. She made a histogram that shows how she rated the movies. What does the shape of the distribution tell you about the movies Kim has rated?

Of the 15 movies that Kim rated, she rated almost half a 7 or an 8 and did not generally give extreme ratings.



Visit my.hrw.com to see all the **Common Core Standards** unpacked.

LESSON 16.1 Measures of Center

COMMON
CORE

6.SP.5

Summarize numerical data sets in relation to their context, ... Also 6.SP.3, 6.SP.5a, 6.SP.5b, 6.SP.5c, 6.SP.5d



ESSENTIAL QUESTION

How can you use measures of center to describe a data set?

EXPLORE ACTIVITY 1



COMMON
CORE

6.SP.5c, 6.SP.3, 6.SP.5a

Finding the Mean

A **measure of center** is a single number used to describe a set of numeric data. A measure of center describes a typical value from the data set.

One measure of center is the *mean*. The **mean**, or average, of a data set is the sum of the data values divided by the number of data values in the set.

Tami surveyed five of her friends to find out how many brothers and sisters they have. Her results are shown in the table.

Number of Siblings				
Amy	Ben	Cal	Don	Eva
2	3	1	1	3

- A** Model each person's response as a group of counters.

Amy	Ben	Cal	Don	Eva
●	●	●	●	●
●	●			●
	●			●

- B** Now rearrange the counters so that each group has the same number of counters.

--	--	--	--	--

Each group now has _____ counter(s). This value is the mean. This model demonstrates how the mean "evens out" the data values.

- C** Use numbers to calculate the mean.

The sum of the data values is $2 + 3 + \square + \square + \square = \square$.

How many data values are in the set? _____

$$\text{Mean} = \frac{\text{sum of data values}}{\text{number of data values}} = \frac{\boxed{}}{\boxed{}} = \boxed{}$$

Math Talk**Mathematical Practices**

Suppose you have a data set in which all of the values are 2. What is the mean?

Reflect

1. Can the mean be greater than the greatest value in a data set? Why or why not?



Math On the Spot

my.hrw.com

Finding the Median

Another measure of center is the *median*. The **median** represents the middle value of an ordered data set.

EXAMPLE 1COMMON
CORE

6.SP.5c, 6.SP.5b

- A** A coach records the distances that some cross-country team members ran last week. Find the median.

Write the data values in order from least to greatest.

3 4 5 5 6 7 7 8 10 10 10

This value is the median.

Equal number of values on either side of the median

The median is 7.

- B** Find the median of these test scores: 87, 90, 77, 83, 99, 94, 93, 90, 85, 83.

Write the data values in order from least to greatest.

77 83 83 85 87 90 90 93 94 99

This data set has two middle values: 87 and 90.

The median is the average of these two values:

$$\text{Median} = \frac{87 + 90}{2} = 88.5$$

The median is 88.5.

**Math Talk****Mathematical Practices**

Why does the data set in **A** have one middle value while the data set in **B** has two middle values?

Reflect

2. **What If?** Which units are used for the data in **A**? If the coach had recorded some distances in kilometers and some in miles, can you still find the median of the data? Explain.

YOUR TURN

3. Charlotte recorded the number of minutes she spent exercising in the past ten days: 12, 4, 5, 6, 8, 7, 9, 8, 2, 1. Find the median of the data.



**Personal
Math Trainer**

Online Assessment
and Intervention

 my.hrw.com

EXPLORE ACTIVITY 2



COMMON
CORE

6.SP.5d, 6.SP.5c

Comparing the Mean and the Median

The mean and median of a data set may be equal, very close to each other, or very different from each other. For data sets where the mean and median differ greatly, one likely describes the data set better than the other.

The monthly earnings of several teenagers are \$200, \$320, \$275, \$250, \$750, \$350, and \$310.



- A** Find the mean.

$$\frac{\boxed{} + \boxed{} + \boxed{} + \boxed{} + \boxed{} + \boxed{} + \boxed{}}{\boxed{}} = \frac{\boxed{}}{\boxed{}} \approx \boxed{}$$

- B** Write the data values in order from least to greatest and find the median.

- C** The mean and the median differ by about \$_____. Why?

- D** Which measure of center better describes the typical monthly earnings for this group of teenagers—the mean or the median? Explain.

EXPLORE ACTIVITY 2 (cont'd)

Reflect

4. **Communicate Mathematical Ideas** Luka's final exam scores for this semester are 70, 72, 99, 72, and 69. Find the mean and median. Which is a better description of Luka's typical exam score? Explain your thinking.

Guided Practice

1. Spencer surveyed five of his friends to find out how many pets they have. His results are shown in the table. What is the mean number of pets? ([Explore Activity 1](#))

Number of Pets				
Lara	Cody	Sam	Ella	Maria
3	5	2	4	1

$$\text{Mean} = \frac{\text{sum of data values}}{\text{number of data values}} = \frac{\boxed{}}{\boxed{}} = \boxed{}$$

The mean number of pets is _____

2. The following are the weights, in pounds, of some dogs at a kennel: 36, 45, 29, 39, 51, 49. ([Example 1](#))
- a. Find the median. _____
- b. Suppose one of the weights were given in kilograms. Can you still find the median? Explain.

3. a. Find the mean and the median of this data set: 9, 6, 5, 3, 28, 6, 4, 7. ([Explore Activity 2](#))

- b. Which better describes the data set, the mean or the median? Explain.



ESSENTIAL QUESTION CHECK-IN

4. How can you use measures of center to describe a data set?

16.1 Independent Practice



6.SP.3, 6.SP.5, 6.SP.5a, 6.SP.5b, 6.SP.5c, 6.SP.5d

Personal Math Trainer

Online Assessment and Intervention

my.hrw.com

Several students in Ashton's class were randomly selected and asked how many text messages they sent yesterday. Their answers were 1, 0, 10, 7, 13, 2, 9, 15, 0, 3.

5. How many students were asked? How do you know?

6. Find the mean and the median for these data.

Mean = _____ Median = _____

The points scored by a basketball team in its last 6 games are shown. Use these data for 7 and 8.

Points Scored					
73	77	85	84	37	115

7. Find the mean score and the median score.

Mean = _____ Median = _____

8. Which measure better describes the typical number of points scored? Explain.

Some people were asked how long it takes them to commute to work. Use the data for 9–11.

9. What units are used for the data? What should you do before finding the mean and median number of minutes?

10. Find the mean and median number of minutes.

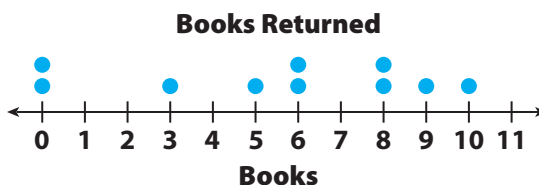
Mean = _____ Median = _____

11. Which measure do you think is more typical of the data?

16 min	5 min
7 min	8 min
14 min	12 min
0.5 hr	1 hr



- 12. Critique Reasoning** For two weeks, the school librarian recorded the number of library books returned each morning. The data are shown in the dot plot. The librarian found the mean number of books returned each morning.



$$\frac{8 + 6 + 10 + 5 + 9 + 8 + 3 + 6}{8} = \frac{55}{8} \approx 6.9$$

Is this the correct mean of this data set? If not, explain and correct the answer.

- 13. Critical Thinking** Lauren's scores on her math tests are 93, 91, 98, 100, 95, 92, and 96. What score could Lauren get on her next math test so that the mean and median remain the same? Explain your answer.

- 14. Persevere in Problem Solving** Yuko wants to take a job selling cars. Since she will get a commission for every car she sells, she finds out the sale price of the last four cars sold at each company.

Company A: \$16,000; \$20,000; \$25,000; \$35,000;

Company B: \$21,000, \$23,000, \$36,000, \$48,000

- a.** Find the mean selling price at each company.

- b.** Find the median selling price at each company.

- c. Communicate Mathematical Ideas** At either company, Yuko will get paid a commission of 20% of the sale price of each car she sells. Based on the data, where do you recommend she take a job? Why?

LESSON 16.2 Mean Absolute Deviation

COMMON CORE 6.SP.5c

Summarize numerical data sets in relation to their context, such as by giving quantitative measures of ...variability (...mean absolute deviation)....



ESSENTIAL QUESTION

How can you determine and use the mean absolute deviation of a set of data points?

EXPLORE ACTIVITY



COMMON CORE 6.SP.5c

Understanding Mean Absolute Deviation

A **measure of variability** is a single number used to describe the spread of a data set. It can also be called a measure of spread. One measure of variability is the **mean absolute deviation (MAD)**, which is the mean of the distances between the data values and the mean of the data set.

**The data represent the height, in feet, of various buildings.
Find the mean absolute deviation for each data set.**

- A** 60, 58, 54, 56, 63, 65, 62, 59, 56, 58

Calculate the mean. Round to the nearest whole number.

Complete the table.

Height (ft)	60	58	54	56	63	65	62	59	56	58
Distance from mean										

Calculate the MAD by finding the mean of the values in the second row of the table. Round to the nearest whole number.

- B** 46, 47, 56, 48, 46, 52, 57, 52, 45

Find the mean. Round to the nearest whole number.

EXPLORE ACTIVITY (cont'd)

Complete the table.

Height (ft)	46	47	56	48	46	52	57	52	45
Distance from mean									

Calculate the MAD. Round to the nearest whole number.

Reflect

- Analyze Relationships** Compare the MADs. How do the MADs describe the distribution of the heights in each group?

Math Talk

Mathematical Practices

What is the difference between a measure of center and a measure of variability?



Math On the Spot

my.hrw.com

Using Mean Absolute Deviation

The mean absolute deviation can be used to answer statistical questions in the real world. Many of these questions may have implications for the operation of various businesses.

EXAMPLE 1



COMMON
CORE

6.SP.5c

A chicken farmer wants her chickens to all have about the same weight. She is trying two types of feed to see which type produces the best results. All the chickens in Pen A are fed Premium Growth feed, and all the chickens in Pen B are fed Maximum Growth feed. The farmer records the weights of the chickens in each pen in the tables below. Which chicken feed produces less variability in weight?



Pen A: Premium Growth Weights (lb)

5.8	6.1	5.5	6.6	7.3	5.9	6.3	5.7	6.8	7.1
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Pen B: Maximum Growth Weights (lb)

7.7	7.4	5.4	7.8	6.1	5.2	7.5	7.9	6.3	5.6
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

STEP 1

Find the mean weight of the chickens in each pen. Round your answers to the nearest tenth.

$$\text{Pen A: } \frac{5.8 + 6.1 + 5.5 + 6.6 + 7.3 + 5.9 + 6.3 + 5.7 + 6.8 + 7.1}{10} \approx 6.3$$

$$\text{Pen B: } \frac{7.7 + 7.4 + 5.4 + 7.8 + 6.1 + 5.2 + 7.5 + 7.9 + 6.3 + 5.6}{10} \approx 6.7$$

STEP 2

Find the distance from the mean for each of the weights.

The distances from the mean for Pen A are the distance of each weight from 6.3 lb.

Pen A: Premium Growth										
Weight (lb)	5.8	6.1	5.5	6.6	7.3	5.9	6.3	5.7	6.8	7.1
Distance from mean	0.5	0.2	0.8	0.3	1.0	0.4	0	0.6	0.5	0.8

The distances from the mean for Pen B are the distance of each weight from 6.7 lb.

Pen B: Maximum Growth										
Weight (lb)	7.7	7.4	5.4	7.8	6.1	5.2	7.5	7.9	6.3	5.6
Distance from mean	1.0	0.7	1.3	1.1	0.6	1.5	0.8	1.2	0.4	1.1

STEP 3

Calculate the MAD for the chickens in each pen. Round your answers to the nearest tenth.

Pen A: $\frac{0.5 + 0.2 + 0.8 + 0.3 + 1.0 + 0.4 + 0 + 0.6 + 0.5 + 0.8}{10} \approx 0.5 \text{ lb}$

Pen B: $\frac{1.0 + 0.7 + 1.3 + 1.1 + 0.6 + 1.5 + 0.8 + 1.2 + 0.4 + 1.1}{10} \approx 1.0 \text{ lb}$

Since Pen A's MAD is less, Premium Growth feed produces less variability in weight.

Math Talk**Mathematical Practices**

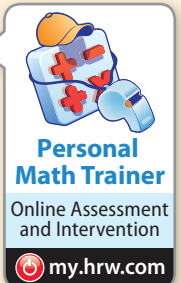
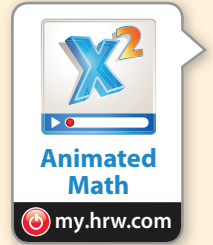
Will a smaller mean always signal less variability?

YOUR TURN

2. Two waiters at a cafe each served 10 large fruit smoothies. The amount in each large smoothie is shown below. Which waiter's smoothies showed less variability?

Amounts in Waiter A's Large Smoothies (oz)									
19.1	20.1	20.9	19.6	20.9	19.5	19.2	19.4	20.3	20.9

Amounts in Waiter B's Large Smoothies (oz)									
20.1	19.6	20.0	20.5	19.8	20.0	20.1	19.7	19.9	20.4





Math On the Spot

my.hrw.com

Using a Spreadsheet to Find MAD

Spreadsheets can be used to find the mean absolute deviation of a data set.

EXAMPLE 2

COMMON
CORE

6.SP.5c

A paper mill is testing two paper-cutting machines. Both are set to produce pieces of paper with a width of 8.5 inches. The actual widths of 8 pieces of paper cut by each machine are shown. Use a spreadsheet to determine which machine has less variability and, thus, does a better job.

Widths of Pieces of Paper Cut by Machine A (in.)

8.502	8.508	8.499	8.501	8.492	8.511	8.505	8.491
-------	-------	-------	-------	-------	-------	-------	-------

Widths of Pieces of Paper Cut by Machine B (in.)

8.503	8.501	8.498	8.499	8.498	8.504	8.496	8.502
-------	-------	-------	-------	-------	-------	-------	-------

STEP 1

Enter the data values for Machine A into row 1 of a spreadsheet, using cells A to H.

	A	B	C	D	E	F	G	H
1	8.502	8.508	8.499	8.501	8.492	8.511	8.505	8.491
2								
3								

STEP 2

Enter "mean =" into cell A2 and the formula =AVERAGE(A1:H1) into cell B2.

	A	B	C	D	E	F	G	H
1	8.502	8.508	8.499	8.501	8.492	8.511	8.505	8.491
2	mean =	8.501125						
3								

STEP 3

Enter "MAD =" into cell A3 and the formula =AVEDEV(A1:H1) into cell B3.

	A	B	C	D	E	F	G	H
1	8.502	8.508	8.499	8.501	8.492	8.511	8.505	8.491
2	mean =	8.501125						
3	MAD =	0.005375						

The MAD for Machine A is about 0.0054 in.

STEP 4

Repeat Steps 1–3 with the data values for Machine B.

	A	B	C	D	E	F	G	H
1	8.503	8.501	8.498	8.499	8.498	8.504	8.496	8.502
2	mean =	8.500125						
3	MAD =	0.002375						

The MAD for Machine B is about 0.0024 in.

Machine B has less variability, so it does a better job.

My Notes

YOUR TURN

3. Two aspirin-making devices are set to make tablets containing 0.35 gram of aspirin. The actual amounts in 8 tablets from each device are shown. Use a spreadsheet to determine which device has less variability.

Amounts of Aspirin in Tablets Made by Device A (g)							
0.353	0.351	0.350	0.352	0.349	0.348	0.350	0.346

Amounts of Aspirin in Tablets Made by Device B (g)							
0.349	0.341	0.347	0.358	0.359	0.354	0.339	0.343



**Personal
Math Trainer**

Online Assessment
and Intervention

my.hrw.com

Guided Practice

1. A bus route takes about 45 minutes. The company's goal is a MAD of less than 0.5 minute. One driver's times for 9 runs of the route are shown. Did the bus driver meet the goal? ([Explore Activity and Example 1](#))

Times to Complete Bus Route (min)								
44.2	44.9	46.1	45.8	44.7	45.2	45.1	45.3	44.6

- a. Calculate the mean of the bus times. _____
- b. Calculate the MAD to the nearest tenth. _____

The bus driver **did / did not** meet the company's goal.

2. Below are a different driver's times on the same route. Find the mean and the MAD using a spreadsheet. Enter the data values into row 1 using cells A to I. Enter "mean =" into cell A2 and "MAD =" into cell A3. ([Example 2](#))

Times to Complete Bus Route (min)								
44.4	43.8	45.6	45.9	44.1	45.6	44.0	44.9	45.8

The mean is _____ minutes, and the MAD is _____ minutes.

This time, the bus driver **did / did not** meet the company's goal.

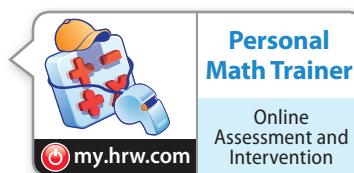


ESSENTIAL QUESTION CHECK-IN

3. What is the mean absolute deviation and what does it tell you about data sets?

16.2 Independent Practice

COMMON CORE 6.SP.5c



Frank wants to know how many people live in each household in his town. He conducts a random survey of 10 people and asks how many people live in their household. His results are shown in the table.

Number of People per Household									
1	6	2	4	4	3	5	5	2	8

- Calculate the mean number of people per household. _____
- Calculate the MAD of the number of people per household. _____
- What conclusions can you draw about the “typical” number of people in each household? Explain.

Teachers are being trained to standardize the scores they give to students’ essays. The same essay was scored by 10 different teachers at the beginning and at the end of their training. The results are shown in the tables.

Scores for Essay at Beginning of Teachers’ Training									
76	81	85	79	89	86	84	80	88	79

Scores for Essay at End of Teachers’ Training									
79	82	84	81	77	85	82	80	78	83

- Calculate the MADs for the teachers’ scores. Did the teachers make progress in standardizing their scores?

- What If?** What would it mean if the teachers had a MAD of 0?

The annual rainfall for Austin, Texas, and San Antonio, Texas, in each of the years from 2002 to 2011 are shown in the tables. Use the data for 9–11.



Annual Rainfall for Austin, Texas (in.)									
36.00	21.41	52.27	22.33	34.70	46.95	16.07	31.38	37.76	19.68

Annual Rainfall for San Antonio, Texas (in.)									
46.27	28.45	45.32	16.54	21.34	47.25	13.76	30.69	37.39	17.58

9. Use a spreadsheet to find the mean for the two cities' annual rainfalls. In which city does it rain more in a year, on average?
10. Use your spreadsheet to find the MADs. Use the MADs to compare the distribution of annual rainfall for the two cities.
11. **Make a Conjecture** Does the information allow you to predict how the future amounts of rainfall for the two cities will compare? Explain.
12. **Critical Thinking** The life spans of 10 adult mayflies have a mean of 4 hours and a MAD of 2 hours. Fill in the table with possible values for the life spans. You can use the same value more than once.

Life Spans of Ten Mayflies (h)									

Can any one of the 10 mayflies in the group live for 1 full day? Justify your answer.

**FOCUS ON HIGHER ORDER THINKING****Work Area**

- 13. Multistep** In a spreadsheet, before entering any data values, first enter “mean =” into cell A2 and the formula =AVERAGE(A1:J1) into cell B2. Next, enter “MAD =” into cell A3 and the formula =AVEDEV(A1:J1) into cell B3. You should see #DIV/0! in cell B2 and #NUM! in cell B3 as shown. Now do the following:

	A	B
1		
2	mean =	#DIV/0!
3	MAD =	#NUM!

- a. Enter “1” into cell A1. What do you get for the mean and the MAD of the data set? Explain why this makes sense.

- b. Enter “2” into cell B1. What do you get for the mean and the MAD of the data set this time? Explain why this makes sense.

- c. Enter the numbers 3 through 10 into cells C1 to J1 and watch the mean and the MAD change. Do they increase, decrease, or stay the same? Explain why this makes sense.

- 14. Make a Conjecture** Each of the values in a data set is increased by 10. Does this affect the MAD of the data set? Why or why not?

- 15. What If?** Suppose a data set contains all whole numbers. Would the MAD for the data set also be a whole number? Explain.

LESSON

16.3 Box Plots

COMMON
CORE

6.SP.4

Display numerical data in plots on a number line, including ... box plots. Also 6.SP.5c



ESSENTIAL QUESTION

How can you use a box plot and measures of spread to describe a data set?

Using a Box Plot

A **box plot** is a display that shows how the values in a data set are distributed, or spread out.

To make a box plot, first find five values for the data set:

- the least value
- the **lower quartile** — the median of the lower half of the data
- the median
- the **upper quartile** — the median of the upper half of the data
- the greatest value

EXAMPLE 1



COMMON
CORE

6.SP.4

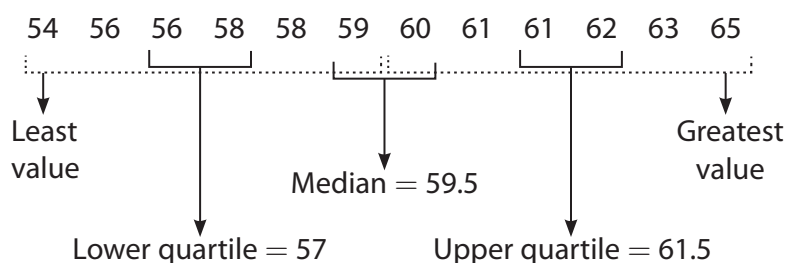
The heights of several students are shown. Make a box plot for the data.

Students' Heights (in.)

60	58	54	56	63	61
65	61	62	59	56	58

STEP 1

Order the data and find the needed values.



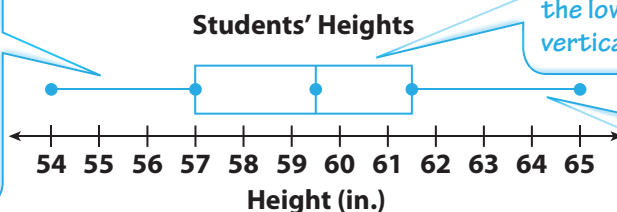
STEP 2

Draw the box plot.

Draw a number line that includes all the data values.

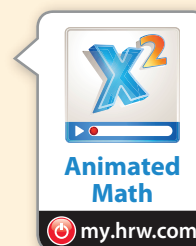
On the number line, draw dots above the least value, the lower quartile, the median, the upper quartile, and the greatest value.

Draw a segment connecting the least value to the lower quartile.



Draw a box whose ends pass through the lower and upper quartiles. Draw a vertical segment through the median.

Draw a segment connecting the upper quartile to the greatest value.



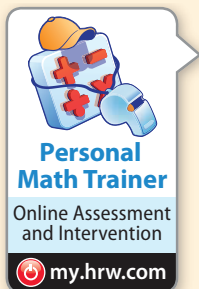
Math Talk

Mathematical Practices

How many data values are in the upper half of the data set? How many in the lower half?

Reflect

1. In the example, what percent of the data values are included in the box portion? What percent are included in each of the “whiskers” on the ends of the box? _____

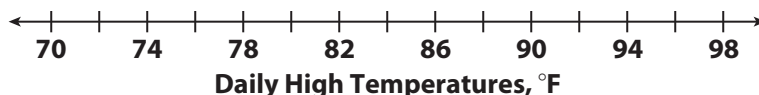


YOUR TURN

2. The daily high temperatures for some days last month are shown. Make a box plot of the data.

Daily High Temperatures (°F)

85	78	92	88	78	84
80	94	89	75	79	83



Finding the Interquartile Range

A **measure of spread** is a single number that describes the spread of a data set. One measure of spread is the *interquartile range*. The **interquartile range (IQR)** is the difference of the upper quartile and the lower quartile.

EXAMPLE 2



COMMON CORE 6.SP.5c

The box plots compare the ages of dancers in two different dance troupes.



- A Find the IQR for each set of data.

Group A: IQR = Upper quartile – Lower quartile

$$= 24 - 20 = 4$$

Group B: IQR = Upper quartile – Lower quartile

$$= 26 - 21.5 = 4.5$$

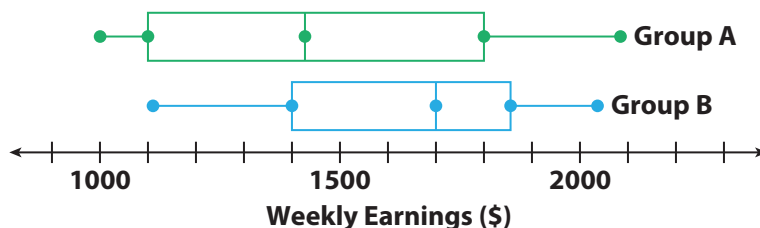
- B Compare the IQRs. How do the IQRs describe the distribution of the ages in each group?

The IQR of group B is slightly greater than the IQR of group A. The ages in the middle half of group B are slightly more spread out than in group A.



YOUR TURN

3. The box plots compare the weekly earnings of two groups of salespeople from different clothing stores. Find and compare the IQRs of the box plots.



Finding the Range

Another measure that describes the spread of a set of data is the *range*. The **range** is the difference of the greatest value and the least value in a set of data.

EXAMPLE 3



COMMON
CORE

6.SP.5c

The data sets show the ages of the players on two professional baseball teams. Find the range of each set of data.

Team A	36, 27, 28, 31, 39, 39, 28, 29, 24, 29, 30, 31, 29, 29, 28, 29, 31, 29, 32, 25, 37, 21, 26, 33, 29
Team B	25, 25, 26, 30, 27, 24, 29, 21, 27, 28, 26, 27, 25, 31, 22, 23, 29, 28, 25, 26, 28, 30, 23, 28, 29

STEP 1 Arrange the data sets in order from least to greatest.

Team A: 21, 24, 25, 26, 27, 28, 28, 28, 29, 29, 29, 29, 29, 29, 30, 31, 31, 31, 32, 33, 36, 37, 39, 39

Team B: 21, 22, 23, 23, 24, 25, 25, 25, 25, 26, 26, 26, 27, 27, 27, 28, 28, 28, 28, 29, 29, 29, 30, 30, 31

STEP 2 Find the range of the data. Subtract the least value from the greatest value in each data set.

Team A: $39 - 21 = 18$

Team B: $31 - 21 = 10$

The range of ages for team A is 18 years, while the range of ages for team B is 10 years.



Personal
Math Trainer

Online Assessment
and Intervention

my.hrw.com



Math On the Spot

my.hrw.com



Math Talk

Mathematical Practices

How can you find the range of a set of data represented by a box plot?

YOUR TURN

4. Find the range of each set of data. Which city's data has a greater range?

Average Monthly High Temperature (°F)	
Miami, FL	76, 78, 80, 83, 87, 90, 91, 91, 89, 86, 82, 78, 84
Chicago, IL	31, 35, 47, 59, 70, 80, 84, 82, 75, 62, 48, 35, 59

Guided Practice

The RBIs (runs batted in) for 15 players from the 2010 Seattle Mariners are shown. Use this data set for 1–7.

Mariners' RBIs
15 51 35 25 58 33 64 43 33 29 14 13 11 4 10

1. Order the data from least to greatest. (Example 1)

2. Find the median. (Example 1) _____

3. Find the lower quartile. (Example 1) _____

4. Find the upper quartile. (Example 1) _____

5. Make a box plot for the data. (Example 1)



6. Find the IQR. (Example 2) _____

7. Find the range. (Example 3) _____

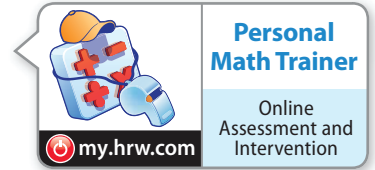


ESSENTIAL QUESTION CHECK-IN

8. How is the range of a set of data different from the IQR?

16.3 Independent Practice

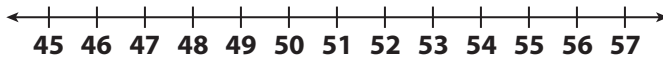
COMMON CORE 6.SP.4, 6.SP.5c



For 9–12, use the data set of the heights of several different students.

Students' Heights (in.)					
46	47	48	48	56	48
46	52	57	52	45	

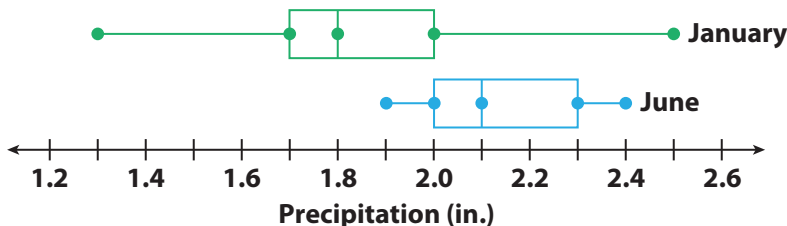
9. Draw a box plot of the data.



10. How many students are included in the data set? _____
11. What method could have been used to collect the data?

12. **Represent Real-World Problems** What other data could you collect from the students to create a box plot? Provide several examples with units of measurement, if applicable.

For 13–15, use the box plots of the total precipitation for the same group of cities for the months of January and June.



13. Calculate the IQR for each month.
January = _____ inches June = _____ inches
14. Calculate the range for each month.
January = _____ inches June = _____ inches
15. Compare the IQRs. What can you conclude about the two data sets?

16. Compare the ranges. What can you conclude about the two data sets?

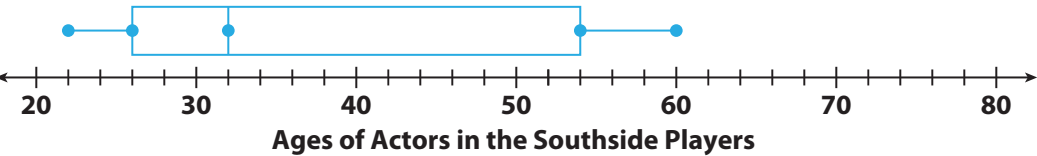


FOCUS ON HIGHER ORDER THINKING

Work Area

17. Analyze Relationships Can two box plots have the same range and IQR and yet represent completely different data? Explain.

18. Multiple Representations Matthew collected data about the ages of the actors in two different community theater groups. He drew a box plot for one of the sets of data.



Ages of Actors in the Northside Players	71, 62, 63, 21, 63, 39, 25, 26, 30
---	------------------------------------

a. Find the median, range, and IQR for each set of data.

Theater Group	Median	Range	IQR
Northside Players			
Southside Players			

b. Suppose you were to draw a second box plot for the Northside Players using the same number line as for the Southside Players. Which box plot would be longer overall? Which would have the longest box portion?

c. **Critique Reasoning** Mandy assumes that because nine data values are shown for the Northside Players, nine data values were used to make the box plot for the Southside Players. Explain why this is not necessarily true.

LESSON 16.4 Dot Plots and Data Distribution

COMMON
CORE

6.SP.4

Display numerical data in plots on a number line, including dot plots. . . . Also 6.SP.1, 6.SP.2, 6.SP.5c, 6.SP.5d



ESSENTIAL QUESTION

How can you summarize and display numeric data?

EXPLORE ACTIVITY

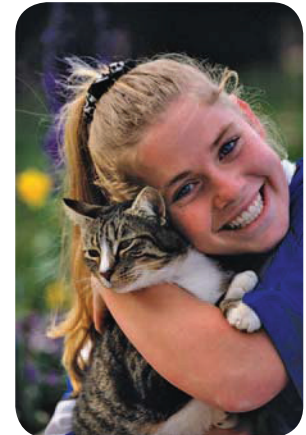


COMMON
CORE

6.SP.1

Variable Data and Statistical Questions

The question “How much does a typical cat weigh?” is an example of a statistical question. A **statistical question** is a question that has many different, or variable, answers.



A Decide whether each of the situations below could yield variable data.

1. Your sister wants to know the typical weight for an adult cat.

2. You want to know how tall your friend is. _____
3. You want to know how far your house is from school. _____
4. A car owner wants to know how much money people usually pay for a new tire. _____
5. How many students were in line for lunch at the cafeteria today at 12:30? _____

B For which of the situations in part **A** can you write a statistical question? Write questions for these situations.

Reflect

1. Choose one of the questions you wrote in part **B**. How might you find answers to this question? What units would you use for the answers?



Math On the Spot

my.hrw.com

Making a Dot Plot

Statistical questions are answered by collecting and analyzing data. One way to understand a set of data is to make a visual display. A **dot plot** is a visual display in which each piece of data is represented by a dot above a number line. A dot plot shows the frequency of each data value.

EXAMPLE 1



COMMON CORE 6.SP.4

A baseball team manager records the number of runs scored by the team in each game for several weeks. Use the data to make a dot plot.

1, 3, 1, 7, 2, 0, 11, 2, 2, 3, 1, 3, 4, 2, 2, 4, 5, 2, 6

The team usually scores between 0 and 7 runs in a game, but in one game they scored 11 runs.

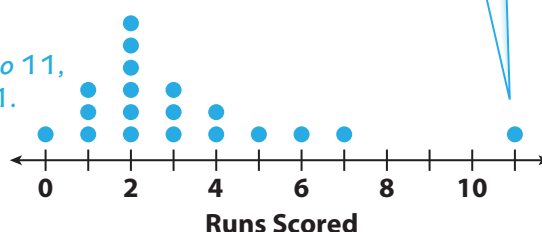
STEP 1

Make a number line.

Data values range from 0 to 11, so use a scale from 0 to 11.

STEP 2

Draw a dot above the number line for each data value.



Reflect

2. How many games did the team play during the season? How can you tell from looking at the dot plot?

3. At how many games did the team score 2 runs or fewer? How do you know?

YOUR TURN

4. A different baseball team scores the following numbers of runs in its games for several weeks:
4, 4, 6, 1, 2, 4, 1, 2, 5, 3, 3, 5, 4, 2



Use the data to make a dot plot. Tell how many games the team played, and identify the data value with the greatest frequency.



Personal Math Trainer

Online Assessment and Intervention

my.hrw.com

Interpreting a Dot Plot

A dot plot can give you a visual picture of the spread, center, and shape of a data distribution.

You can describe the spread of a data set by identifying the least and greatest values. You can also look for **outliers** which are data values that are either much greater or much less than the other data values.

You can describe the center and shape of a data set in terms of *peaks*, *clusters*, or *symmetry*. A symmetric distribution has approximately the same number of data values on either side of the center.



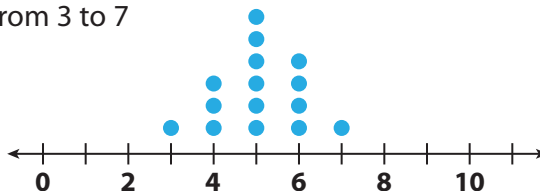
EXAMPLE 2

COMMON CORE 6.SP.2

Describe the spread, center, and shape of each data distribution.

- A** The data values are spread out from 3 to 7 with no outliers.

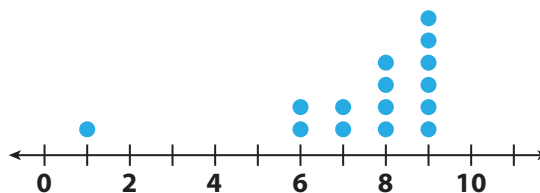
The data has a cluster from 3 to 7 with one peak at 5, which is the center of the distribution.



The distribution is symmetric. The data values are clustered around the center of the distribution.

- B** The data values are spread out from 1 to 9. The data value 1 appears to be an outlier.

The data has a cluster from 6 to 9 with one peak at 9, which is the greatest value in the data set.

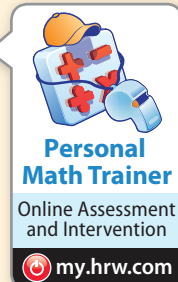


The distribution is not symmetric. The data values are clustered at one end of the distribution.

My Notes

YOUR TURN

5. Describe the spread, center, and shape of the data distribution from Example 1.





Math On the Spot

my.hrw.com

Finding Measures from a Dot Plot

You can also find and calculate measures of center and spread from a dot plot.

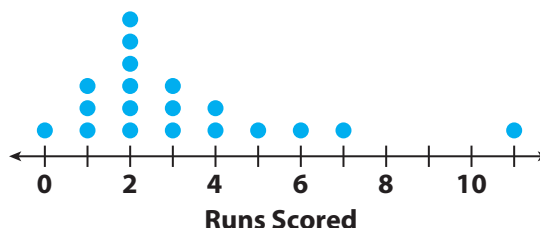
EXAMPLE 3



COMMON
CORE

6.SP.5d, 6.SP.5c

The dot plot shows the number of runs scored by a baseball team in each game for several weeks from Example 1.



- A** Find the mean, median, and range of the data.

STEP 1

To find the mean, find the sum of the data values and divide by the number of data values.

$$\frac{1(0) + 3(1) + 6(2) + 3(3) + 2(4) + 1(5) + 1(6) + 1(7) + 1(11)}{19} = \frac{61}{19} \approx 3.2$$

The mean is about 3.2.

STEP 2

To find the median, count the dots from left to right until you find the middle value. You may need to find the mean of two middle values.

The median is 2.

STEP 3

To find the range, read the least and greatest values from the dot plot. Subtract the least value from the greatest.

$$11 - 0 = 11$$

The range is 11.

Math Talk

Mathematical Practices

Why is the question in **B** a statistical question?

- B** How many runs does the team typically score in a game? Explain.

The mean number of runs is 3.2. The median number of runs is 2.

The shape of the dot plot suggests that the outlier 11 may be affecting these measures of center. To see if that is the case, find the mean and median without including the outlier. Compare these values with the original values.

STEP 1

Find the mean without including the outlier.

$$\frac{1(0) + 3(1) + 6(2) + 3(3) + 2(4) + 1(5) + 1(6) + 1(7)}{18} = \frac{50}{18} \approx 2.8$$

Without the outlier, the mean is 2.8, which is less than the original mean of 3.2.

STEP 2

Find the median without including the outlier.

Counting from left to right, the median is still 2.

Given that it is not affected by the outlier, the median may be more typical of the data. The team typically scores two runs per game.


YOUR TURN

6. Find the mean, median, and range of the data from Your Turn question 4. What is the typical number of runs the team scores in a game? Justify your answer.



**Personal
Math Trainer**

Online Assessment
and Intervention

 my.hrw.com

Guided Practice

Tell whether the situation could yield variable data. If possible, write a statistical question. (Explore Activity)

1. The town council members want to know how much recyclable trash a typical household in town generates each week.

Kate asked some friends how many movies they saw last winter. Use her data for 2 and 3.

Movies Seen Last Winter

0, 1, 1, 2, 2, 3, 3, 3, 4, 4, 4, 4, 5, 5, 5, 5,
6, 6, 7, 7, 7, 8, 8, 9, 9, 17

2. Make a dot plot of the data. (Example 1)



3. Find the mean, median, and range of the data. (Example 3)

4. Describe the spread, center, and shape of the data. (Example 2)

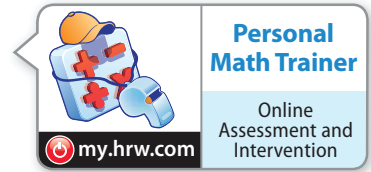


ESSENTIAL QUESTION CHECK-IN

5. What are some measures of center and spread that you can find from a dot plot? How can making a dot plot help you visualize a data distribution?

16.4 Independent Practice

COMMON CORE 6.SP.1, 6.SP.2, 6.SP.4, 6.SP.5c, 6.SP.5d



- 6. Vocabulary** Describe how a statistical question yields an answer with variability. Give an example.

For 7–10, determine whether the question is a statistical question. If it is a statistical question, identify the units for the answer.

- 7.** An antique collector wants to know the age of a particular chair in a shop.

- 8.** How tall do the people in your immediate and extended family tend to be?



- 9.** How tall is Sam?

- 10.** How much did your classmates typically spend on music downloads last year?

For 11–14, use the following data. The data give the number of days of precipitation per month during one year in a city.

12 10 11 9 9 10 12 9 8 7 9 10

- 11.** Make a dot plot of the data.



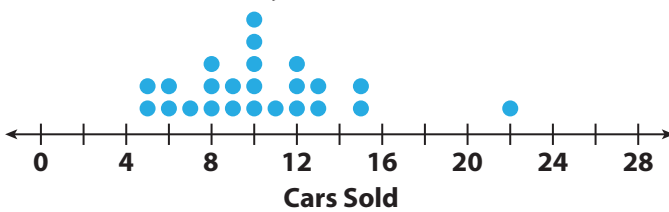
- 12.** What does each dot represent? How many months are represented?

- 13.** Describe the shape, center, and spread of the data distribution. Are there any outliers?

- 14.** Find the mean, median, and range of the data.

- 15. What If?** During one month there were 7 days of precipitation. What if there had only been 3 days of precipitation that month? How would that change the measures of center?

For 16–20, use the dot plot of the number of cars sold at a car dealership per week during the first half of the year.



- 16.** Find the mean, median, and range.

Mean = _____ Median = _____

Range = _____

- 17.** The owner of the car dealership decides to treat the value 22 as an outlier. Which measure of center or spread is affected the most if the owner removes this outlier? Explain.

- 18.** How many cars are sold in a typical week at the dealership? Explain.

- 19.** Write an expression that represents the total number of cars sold during the first half of the year.

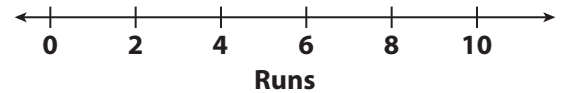
- 20.** Describe the spread, center, and shape of the data distribution.

- 21. Vocabulary** Explain how you can tell the frequency of a data value by looking at a dot plot.

For 22–26 use the following data. The data give the number of runs scored by opponents of the Boston Red Sox in June 2010.

4, 4, 9, 0, 2, 4, 1, 2, 11, 8, 2, 2, 5, 3, 2, 5, 6, 4, 0

- 22.** Make a dot plot for the data.



- 23.** How many games did the Boston Red Sox play in June 2010? Explain.

- 24.** Which data value in your dot plot has the greatest frequency? Explain what that frequency means for this data.

- 25.** Find the mean, median, and range of the data.

- 26.** What is a statistical question that you could answer using the dot plot? Answer your question and justify your response.

**FOCUS ON HIGHER ORDER THINKING**

Work Area

- 27.** A pediatrician records the ages of the patients seen in one day:
1, 2, 5, 7, 9, 17, 13, 16, 18, 12, 3, 5, 1.

a. Explain the Error Assuming that some of the patients are infants who are less than 1 year old, what information did the pediatrician forget to write down?

b. Critical Thinking Can you make a dot plot of the pediatrician's data? Can you find the mean, median, and range? Why or why not?

- 28. Multistep** A nurse measured a patient's heart rate at different times over several days.

Heart Rate (beats per minute)
86, 87, 89, 87, 86, 88, 90, 85, 82, 86, 83, 85, 84, 86



- a.** Make a dot plot.
- b.** Describe the shape, center, and spread of the data. Then find the mean, median, range, and IQR for the data.

- c. What If?** The nurse collected the data when the patient was resting. How might the dot plot and the measures change if the nurse collects the data when the patient is exercising?

LESSON 16.5 Histograms

COMMON CORE 6.SP.4

Display numerical data in plots on a number line, including ... histograms. ... Also 6.SP.5



ESSENTIAL QUESTION

How can you display data in a histogram?

EXPLORE ACTIVITY



COMMON CORE 6.SP.5, 6.SP.4

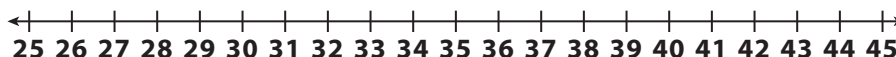
Grouping Data in Intervals

The members of the high-school basketball team practice free throws. Each player attempts 50 free throws. The number of free throws made by each player is listed below.

25, 29, 29, 30, 33, 34, 35, 35, 36, 39, 42, 44



- A** Use a dot plot to represent the data.



- B** On your dot plot, circle the dots that are in each interval of the frequency table below. Then complete the frequency table.

Interval	Frequency
20–29	
30–39	
40–49	

Enter the number of data values for the interval 30–39.

- C** Analyze the data. How were the data collected? How many data values are there? What are the mean, median, range, and IQR of the data?

Reflect

1. Can you use the dot plot to find the mean and the median of the data?
Can you use the frequency table? Why or why not?
2. How do you find the number of data values in a data set from a dot plot?
How can you find the number of data values from a frequency table?



Math On the Spot

my.hrw.com

Using a Histogram

A **histogram** is a type of bar graph whose bars represent the frequencies of numeric data within intervals.

EXAMPLE 1

COMMON CORE 6.SP.4, 6.SP.5

A birdwatcher counts and records the number of birds at a birdfeeder every morning at 9:00 for several days.

12, 3, 8, 1, 1, 6, 10, 14, 3, 6, 2, 1, 3, 2, 7

Make a histogram of the data.

STEP 1

Make a frequency table.

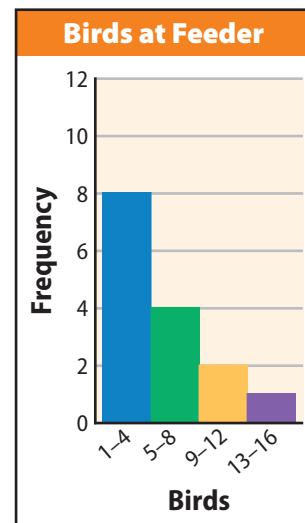
Divide the data into equal-sized intervals of 4. Make a frequency table.

Interval	Frequency
1–4	8
5–8	4
9–12	2
13–16	1

STEP 2

Make a histogram.

The intervals are listed along the horizontal axis. The vertical axis shows the frequencies. For each interval, draw a bar to show the number of days in that interval. The bars should have equal widths. They should touch but not overlap.

**Math Talk****Mathematical Practices**

How does the histogram show the total number of days the birdwatcher counted birds?

Reflect

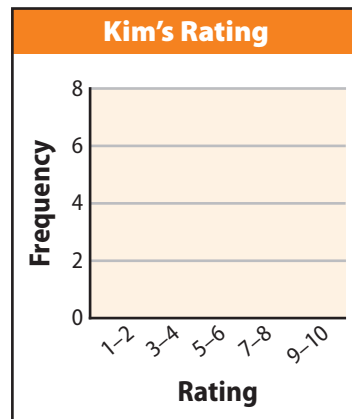
3. **What If?** Suppose the birdwatcher continues his observation for three more days and collects these new data values: 5, 18, and 2. How could you change the histogram to include the data?

YOUR TURN

4. Kim has started rating each movie she sees using a scale of 1 to 10 on an online site. Here are her ratings so far:

6, 9, 8, 5, 7, 4, 8, 8, 3, 7, 8, 7, 5, 1, 10

Make a histogram of the data.



Personal Math Trainer

Online Assessment and Intervention

my.hrw.com

Analyzing a Histogram

By grouping data in intervals, a histogram gives a picture of the distribution of a data set.

EXAMPLE 2



COMMON CORE 6.SP.5

Use the histogram from Example 1. What are some conclusions about the data that can you make from the shape of the distribution?

The highest bar is for the interval 1–4, which means that on more than half the days (8 out of 15), the birdwatcher saw only 1–4 birds. The bars decrease in height from left to right, showing that it was more likely for the birdwatcher to see a small number of birds rather than a large number on any given day.

YOUR TURN

5. Use your histogram from Your Turn 4. What are some conclusions you can make about Kim's movie ratings from the shape of the distribution?



Math On the Spot

my.hrw.com



Personal Math Trainer

Online Assessment and Intervention

my.hrw.com

Guided Practice

1. Wendy kept track of the number of text messages she sent each day for three weeks. Complete the frequency table.

(Explore Activity)

0, 5, 5, 7, 11, 12, 15, 20, 22, 24, 25,
25, 27, 27, 29, 29, 32, 33, 34, 35, 35

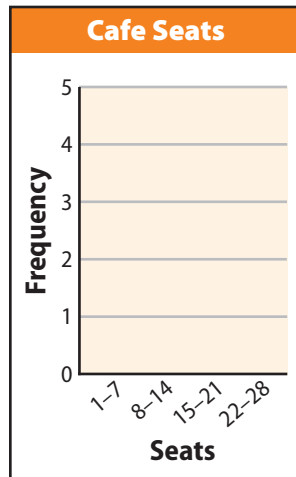
Interval	Frequency
0–9	

Ed counted the number of seats available in each cafe in his town. Use his data for 2–3.

18, 20, 22, 26, 10, 12, 16, 18, 7, 8

2. Complete the frequency table and the histogram. (Example 1)

Interval	Frequency
1–7	
8–14	
15–21	
22–28	



3. What are some conclusions you can make about the distribution of the data? (Example 2)

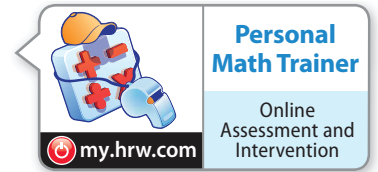


ESSENTIAL QUESTION CHECK-IN

4. How can you display data in a histogram?

16.5 Independent Practice

COMMON CORE 6.SP.4, 6.SP.5



An amusement park employee records the ages of the people who ride the new roller coaster during a fifteen-minute period.

Ages of riders: 47, 16, 16, 35, 45, 43, 11, 29, 31, 50, 23, 18, 18, 20, 29, 17, 18, 48, 56, 24, 18, 21, 38, 12, 23.

5. Complete the frequency table. Then make a histogram of the data.

Interval	Frequency
10–19	

Roller Coaster Riders

6. Describe two things you know about the riders who are represented by the data.

7. **Multiple Representations** West Middle School has classes of many different sizes during first period. The number of students in each class is shown.

9, 23, 18, 14, 20, 26, 14, 18, 18, 12, 8, 13, 21, 22, 28, 10, 7, 19, 24, 20

- a. Hank made a histogram using intervals of 6–10, 11–15, and so on. How many bars did his histogram have? What was the height of the highest bar? _____
- b. Lisa made a histogram using intervals of 0–9, 10–19, and so on. How many bars did her histogram have? What was the height of the highest bar? _____
- c. Besides a histogram, what are some other ways you could display these data? _____



8. **Communicate Mathematical Ideas** Can you find the mean or median of a set of data from a histogram? Explain.

9. **Multistep** A theater owner keeps track of how many people come to see movies on 21 different Saturdays.

Saturday Moviegoers

Interval	Frequency
60–69	1
70–79	3
80–89	10
90–99	7

Saturday Moviegoers



- a. Use the data to make a histogram.
- b. **Make a Prediction** The theater owner asks, “How many moviegoers come to the theater on a typical Saturday?” What would you tell the theater owner? Use your histogram to support your answer.

- c. **Communicate Mathematical Ideas** Is the theater owner’s question a statistical question? Why or why not?

10. **Explain the Error** Irina says she can find the range of a set of data from a histogram. Is she correct? Justify your answer.

Ready to Go On?



**Personal
Math Trainer**

Online Assessment
and Intervention

 my.hrw.com

16.1 Measures of Center

1. Find the mean and median of these data: 2, 5, 9, 11, 17, 19. _____

16.2 Mean Absolute Deviation

2. Find the distance of each data value in Exercise 1 from the mean. Then find the mean absolute deviation of the data. _____

16.3 Box Plots

3. Make a box plot for the data set.

36	42	44	52	61	70	78
----	----	----	----	----	----	----



16.4 Dot Plots and Data Distribution

A baseball team scored the following number of runs over a 10-game period:

6, 6, 8, 5, 4, 6, 4, 3, 8, 4

4. Make a dot plot for the data.

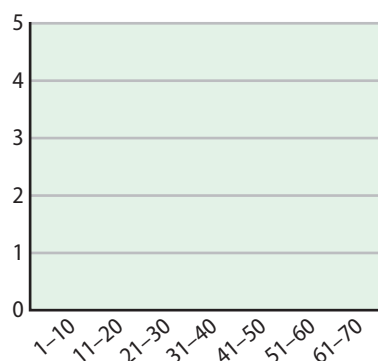
5. Find the mean, median, and range.



16.5 Histograms

6. Make a histogram for the data set.

23	45	62	19
48	10	39	54
39	16	48	12
25	32	18	4



ESSENTIAL QUESTION

7. How can you represent and summarize data in a dot plot?

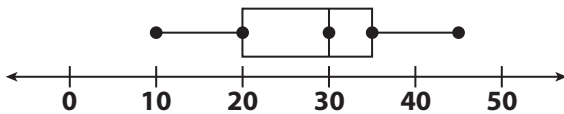


Assessment Readiness



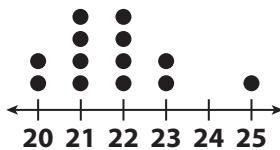
Selected Response

1. What is the interquartile range of the data represented by the box plot shown below?



- (A) 15 (C) 35
(B) 20 (D) 40

The dot plot shows the ages of quiz show contestants.



2. What is the median of the data?
- (A) 21 (C) 22
(B) 21.5 (D) 25
3. Which inequalities describe the possible ages of the contestants in the dot plot?
- (A) $a > 20$ and $a < 25$
(B) $a \geq 20$ and $a \leq 25$
(C) $a < 20$ and $a > 25$
(D) $a \leq 20$ and $a \geq 25$
4. Suppose a new data value, 30, is included in the dot plot. Which statement describes the effect on the median?
- (A) The median would increase.
(B) The median would decrease.
(C) The median would stay the same.
(D) The median would equal the mean.

5. Andrea recorded the points she scored in her last eight basketball games. What is the mean absolute deviation of the scores?

28, 32, 47, 16, 40, 35, 38, 54

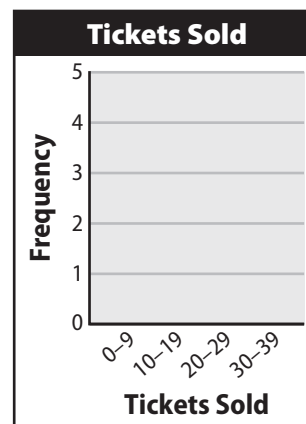
- (A) 8.5 (C) 17.75
(B) 36.25 (D) 38

Mini-Task

6. The frequency table shows data about how many tickets were sold by students.

Tickets Sold	Frequency
0–9	2
10–19	4
20–29	3
30–39	1

- a. Use the frequency table to make a histogram.



- b. How many students sold tickets?

- c. What percent of the students sold 20 or more tickets?