Statistics and Sampling

MODULE 5

Random Samples and Populations

FL 7.RP.1.2c, 7.SP.1.1, 7.SP.1.2

MODULE 6

Analyzing and Comparing Data

FL 7.SP.2.3, 7.SP.2.4

CAREERS IN MATH

Entomologist An entomologist is a biologist who studies insects. These scientists analyze data and use mathematical models to understand and predict the behavior of insect populations.

If you are interested in a career in entomology, you should study these mathematical subjects:
- Algebra
- Trigonometry
- Probability and Statistics
- Calculus

Research other careers that require the analysis of data and use of mathematical models.

Unit 3 Performance Task

At the end of the unit, check out how **entomologists** use math.

Vocabulary Preview

Use the puzzle to preview key vocabulary from this unit. Unscramble the circled letters to answer the riddle at the bottom of the page.
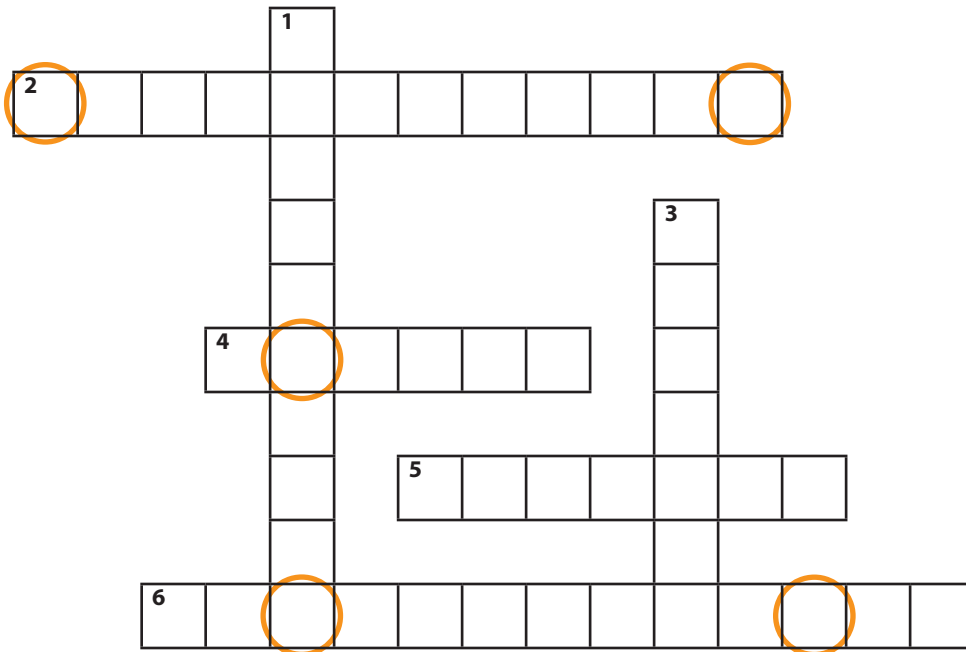
Across

2. A sample in which every person, object, or event has an equal chance of being selected (2 words). (Lesson 5.1)

4. Part of a population chosen to represent the entire group. (Lesson 5.1)

5. A display in which each piece of data is represented by a dot above a number line (2 words). (Lesson 6.2)

6. An integer generated by chance, such as by rolling a number cube or using a graphing calculator. (Lesson 5.1)

Down

1. The entire group of objects, individuals, or events in a set of data. (Lesson 5.1)

3. A display that shows how the values in a data set are distributed (2 words). (Lesson 6.3)

Q: Where do cowboys who love statistics live?

A: on the ___ ___ ___ ___ ___!

Random Samples and Populations

? ESSENTIAL QUESTION

How can you use random samples and populations to solve real-world problems?

Real-World Video

Scientists study animals like dolphins to learn more about characteristics such as behavior, diet, and communication. Acoustical data (recordings of dolphin sounds) can reveal the species that made the sound.

my.hrw.com

© Houghton Mifflin Harcourt Publishing Company • Image Credits: ©Kevin Schafer/Alamy

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.

Solve Proportions

EXAMPLE

$$\frac{a}{1} = \frac{30}{1.5}$$

$a \times 1.5 = 1 \times 30$ Write the cross products.

$1.5a = 30$ Simplify.

$\frac{1.5a}{1.5} = \frac{30}{1.5}$ Divide both sides by 1.5.

$a = 20$

Solve for *x*.

1. $\frac{x}{16} = \frac{45}{40}$ _____

2. $\frac{x}{5} = \frac{1}{4}$ _____

3. $\frac{2.5}{10} = \frac{x}{50}$ _____

4. $\frac{x}{6} = \frac{2}{9}$ _____

Find the Range

EXAMPLE   29, 26, 21, 30, 32, 19   Order the data from least to greatest.
19, 21, 26, 29, 30, 32

range = 32 − 19   The range is the difference between the greatest and the least data items.
= 13

Find the range of the data.

5. 52, 48, 57, 47, 49, 60, 59, 51 _____

6. 5, 9, 13, 6, 4, 5, 8, 12, 12, 6 _____

7. 97, 106, 99, 97, 115, 95, 108, 100 _____

8. 27, 13, 35, 19, 71, 12, 66, 47, 39 _____

Find the Mean

EXAMPLE   21, 15, 26, 19, 25, 14

mean $= \frac{21 + 15 + 26 + 19 + 25 + 14}{6}$ The mean is the sum of the data items divided by the number of items.

$= \frac{120}{6}$

$= 20$

Find the mean of each set of data.

9. 3, 5, 7, 3, 6, 4, 8, 6, 9, 5 _____

10. 8.1, 9.4, 11.3, 6.7, 6.2, 7.5 _____

Reading Start-Up

Visualize Vocabulary

Use the ✔ words to complete the right column of the chart.

Box Plots to Display Data	
Definition	**Review Word**
A display that uses values from a data set to show how the values are spread out.	
The middle value of a data set.	
The median of the lower half of the data.	
The median of the upper half of the data.	

Understand Vocabulary

Complete each sentence, using the preview words.

1. An entire group of objects, individuals, or events is a

   _____.

2. A _____ is part of the population chosen to represent the entire group.

3. A sample that does not accurately represent the population is a

   _____.

Vocabulary

Review Words
- ✔ box plot *(diagrama de caja)*
-   data *(datos)*
-   dot plot *(diagrama de puntos)*
-   interquartile range *(rango entre cuartiles)*
- ✔ lower quartile *(cuartil inferior)*
- ✔ median *(mediana)*
-   spread *(dispersión)*
-   survey *(estudio)*
- ✔ upper quartile *(cuartil superior)*

Preview Words
-   biased sample *(muestra sesgada)*
-   population *(población)*
-   random sample *(muestra aleatoria)*
-   sample *(muestra)*

Active Reading

Tri-Fold Before beginning the module, create a tri-fold to help you learn the concepts and vocabulary in this module. Fold the paper into three sections. Label the columns "What I Know," "What I Need to Know," and "What I Learned." Complete the first two columns before you read. After studying the module, complete the third column.

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.

FL 7.SP.1.1

Understand that statistics can be used to gain information about a population by examining a sample of the population; generalizations about a population from a sample are valid only if the sample is representative of that population. Understand that random sampling tends to produce representative samples and support valid inferences.

What It Means to You

You will learn how a random sample can be representative of a population.

UNPACKING EXAMPLE 7.SP.1.1

Avery wants to survey residents who live in an apartment building. She writes down all of the apartment numbers on slips of paper, and draws slips from a box without looking to decide who to survey. Will this produce a random sample?

The population is all of the residents or people who live in the apartment building. The sample is a valid random sample because every apartment number has the same chance of being selected.

FL 7.SP.1.2

Use data from a random sample to draw inferences about a population with an unknown characteristic of interest. Generate multiple samples (or simulated samples) of the same size to gauge the variation in estimates or predictions.

Key Vocabulary

population *(población)*
    The entire group of objects or individuals considered for a survey.

sample *(muestra)*
    A part of the population.

What It Means to You

You will use data collected from a random sample to make inferences about a population.

UNPACKING EXAMPLE 7.SP.1.2

Alexi surveys a random sample of 80 students at his school and finds that 22 of them usually walk to school. There are 1,760 students at the school. Predict the number of students who usually walk to school.

$$\frac{\text{number in sample who walk}}{\text{size of sample}} = \frac{\text{number in population who walk}}{\text{size of population}}$$

$$\frac{22}{80} = \frac{x}{1,760}$$

$$x = \frac{22}{80} \cdot 1,760$$

$$x = \frac{38,720}{80} = 484$$

Approximately 484 students usually walk to school.

Visit **my.hrw.com** to see all **Florida Math Standards** unpacked.

my.hrw.com

Populations and Samples

? **ESSENTIAL QUESTION**

How can you use a sample to gain information about a population?

EXPLORE ACTIVITY Real World   FL **7.SP.1.1**

Random and Non-Random Sampling

When information is being gathered about a group, the entire group of objects, individuals, or events is called the **population**. A **sample** is part of the population that is chosen to represent the entire group.

A vegetable garden has 36 tomato plants arranged in a 6-by-6 array. The gardener wants to know the average number of tomatoes on the plants. Each white cell in the table represents a plant. The number in the cell tells how many tomatoes are on that particular plant.

						First Number Cube
8	9	13	18	24	15	1
34	42	46	20	13	41	2
29	21	14	45	27	43	3
22	45	46	41	22	33	4
12	42	44	17	42	11	5
18	26	43	32	33	26	6
Second Number Cube						
1	2	3	4	5	6	

Because counting the number of tomatoes on all of the plants is too time-consuming, the gardener decides to choose plants at random to find the average number of tomatoes on them.

To simulate the random selection, roll two number cubes 10 times. Find the cell in the table identified by the first and second number cubes. Record the number in each randomly selected cell.

A What is the average number of tomatoes on the 10 plants that were randomly selected?

_____

B Alternately, the gardener decides to choose the plants in the first row. What is the average number of tomatoes on these plants?

_____

Math Talk
Mathematical Practices
How do the averages you got with each sampling method compare?

Reflect

1. How do the averages you got with each sampling method compare to the average for the entire population, which is 28.25?

   _____

   _____

2. Why might the first method give a closer average than the second method?

   _____

   _____

Random Samples and Biased Samples

A sample in which every person, object, or event has an equal chance of being selected is called a **random sample**. A random sample is more likely to be representative of the entire population than other sampling methods. When a sample does not accurately represent the population, it is called a **biased sample**.

EXAMPLE 1   *Real World*   **FL**   **7.SP.1.1**

Identify the population. Determine whether each sample is a random sample or a biased sample. Explain your reasoning.

A Roberto wants to know the favorite sport of adults in his hometown. He surveys 50 adults at a baseball game.

The population is adults in Roberto's hometown.

The sample is biased.

Think: People who don't like baseball will not be represented in this sample.

B Paula wants to know the favorite type of music for students in her class. She puts the names of all students in a hat, draws 8 names, and surveys those students.

The population is students in Paula's class.

The sample is random.

Think: Each student has an equal chance of being selected.

Math Talk

Mathematical Practices

Why do you think samples are used? Why not survey each member of the population?

Reflect

3. How might you choose a sample of size 20 to determine the preferred practice day of all the players in a soccer league?

   _____

   _____

YOUR TURN

4. For a survey, a company manager assigned a number to each of the company's 500 employees, and put the numbers in a bag. The manager chose 20 numbers and surveyed the employees with those numbers. Did the manager choose a random sample?

_____

_____

Bias in Survey Questions

Once you have selected a representative sample of the population, be sure that the data is gathered without bias. Make sure that the survey questions themselves do not sway people to respond a certain way.

EXAMPLE 2 Real World   FL   7.SP.1.1

In Madison County, residents were surveyed about a new skateboard park. Determine whether each survey question may be biased. Explain.

A Would you like to waste the taxpayers' money to build a frivolous skateboard park?

This question is biased. It discourages residents from saying yes to a new skateboard park by implying it is a waste of money.

B Do you favor a new skateboard park?

This question is not biased. It does not include an opinion on the skateboard park.

C Studies have shown that having a safe place to go keeps kids out of trouble. Would you like to invest taxpayers' money to build a skateboard park?

This question is biased. It leads people to say yes because it mentions having a safe place for kids to go and to stay out of trouble.

YOUR TURN

Determine whether each question may be biased. Explain.

5. When it comes to pets, do you prefer cats?

_____

6. What is your favorite season?

_____

_____

1. Follow each method described below to collect data to estimate the average shoe size of seventh grade boys. (Explore Activity)

Method 1

A Randomly select 6 seventh grade boys and ask each his shoe size. Record your results in a table like the one shown.

_____

B Find the mean of this data. Mean:

_____

Random Sample of Seventh Grade Male Students	
Student	Shoe Size

Method 2

A Find the 6 boys in your math class with the largest shoes and ask their shoe size. Record your results in a table like the one shown in Method 1.

_____

B Find the mean of this data. Mean: _____

2. Method 1 produces results that are [**more / less**] representative of the entire student population because it is a [**random / biased**] sample. (Example 1)

3. Method 2 produces results that are [**more / less**] representative of the entire student population because it is a [**random / biased**] sample. (Example 1)

4. Heidi decides to use a random sample to determine her classmates' favorite color. She asks, "Is green your favorite color?" Is Heidi's question biased? If so, give an example of an unbiased question that would serve Heidi better. (Example 2)
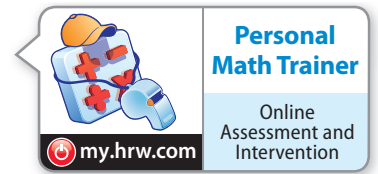
_____

? ESSENTIAL QUESTION CHECK-IN

5. How can you select a sample so that the information gained represents the entire population?

_____

_____

5.1 Independent Practice

Personal Math Trainer

Online Assessment and Intervention

my.hrw.com

6. Paul and his friends average their test grades and find that the average is 95. The teacher announces that the average grade of all of her classes is 83. Why are the averages so different?

_____

_____

_____

7. Nancy hears a report that the average price of gasoline is $2.82. She averages the prices of stations near her home. She finds the average price of gas to be $3.03. Why are the averages different?

_____

_____

_____

_____

For 8–10, determine whether each sample is a random sample or a biased sample. Explain.

8. Carol wants to find out the favorite foods of students at her middle school. She asks the boys' basketball team about their favorite foods.

_____

_____

9. Dallas wants to know what elective subjects the students at his school like best. He surveys students who are leaving band class.

_____

_____

_____

10. To choose a sample for a survey of seventh graders, the student council puts pieces of paper with the names of all the seventh graders in a bag, and selects 20 names.

_____

_____

_____

11. Members of a polling organization survey 700 of the 7,453 registered voters in a town by randomly choosing names from a list of all registered voters. Is their sample likely to be representative?

_____

_____

For 12–13, determine whether each question may be biased. Explain.

12. Joey wants to find out what sport seventh grade girls like most. He asks girls, "Is basketball your favorite sport?"

_____

_____

_____

13. Jae wants to find out what type of art her fellow students enjoy most. She asks her classmates, "What is your favorite type of art?"

_____

_____

_____

14. Draw Conclusions Determine which sampling method will better represent the entire population. Justify your answer.

Student Attendance at Football Games	
Sampling Method	**Results of Survey**
Collin surveys 78 students by randomly choosing names from the school directory.	63% attend football games.
Karl surveys 25 students that were sitting near him during lunch.	82% attend football games.

_____

_____

_____

_____

_____

_____

15. Multistep Barbara surveyed students in her school by looking at an alphabetical list of the 600 student names, dividing them into groups of 10, and randomly choosing one from each group.

a. How many students did she survey? What type of sample is this?

_____

b. Barbara found that 35 of the survey participants had pets. About what percent of the students she surveyed had pets? Is it safe to believe that about the same percent of students in the school have pets? Explain your thinking.

_____

_____

_____

16. Communicating Mathematical Ideas Carlo said a population can have more than one sample associated with it. Do you agree or disagree with his statement? Justify your answer.

_____

_____

_____

Making Inferences from a Random Sample

How can you use a sample to gain information about a population?

EXPLORE ACTIVITY 1 Real World   FL 7.SP.1.2, 7.SP.1.1

Using Dot Plots to Make Inferences

After obtaining a random sample of a population, you can make inferences about the population. Random samples are usually representative and support valid inferences.

Rosee asked students on the lunch line how many books they had in their backpacks. She recorded the data as a list: 2, 6, 1, 0, 4, 1, 4, 2, 2. Make a dot plot for the books carried by this sample of students.

STEP 1 Order the data from least to greatest. Find the least and greatest values in the data set.

STEP 2 Draw a number line from 0 to 6. Place a dot above each number on the number line for each time it appears in the data set.

Notice that the dot plot puts the data values in order.

Math Talk
Mathematical Practices
No students in Rosee's sample carry 3 books. Do you think this is true of all the students at the school? Explain.

Reflect

1. **Critical Thinking** How are the number of dots you plotted related to the number of data values?

   _____

2. **Draw Conclusions** Complete each qualitative inference about the population.

   Most students have _____ 1 book in their backpacks.

   Most students have fewer than _____ books in their backpacks.

   Most students have between _____ books in their backpacks.

3. **Analyze Relationships** What could Rosee do to improve the quality of her data?

   _____

Using Box Plots to Make Inferences

You can also analyze box plots to make inferences about a population.

The number of pets owned by a random sample of students at Park Middle school is shown below. Use the data to make a box plot.
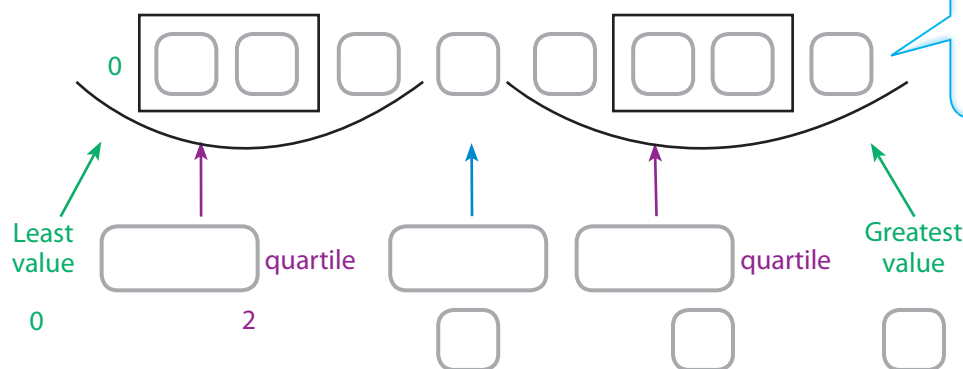
9, 2, 0, 4, 6, 3, 3, 2, 5

STEP 1 Order the data from least to greatest. Then find the least and greatest values, the median, and the lower and upper quartiles.

STEP 2 The lower and upper quartiles can be calculated by finding the medians of each "half" of the number line that includes all the data.

0

> The lower quartile is the mean of 2 and 2. The upper quartile is the mean of 5 and 6.

Least value

0

quartile

2

quartile

Greatest value

Draw a number line that includes all the data values.

Plot a point for each of the values found in Step 1.

STEP 3 Draw a box from the lower to upper quartile. Inside the box, draw a vertical line through the median. Finally, draw the whiskers by connecting the least and greatest values to the box.

Reflect

4. **Draw Conclusions** Complete each qualitative inference about the population.

A good measure for the most likely number of pets is _____.

50% of the students have between _____ and 3 pets.

Almost every student in Parkview has at least _____ pet.

Math Talk
Mathematical Practices

What can you see from a box plot that is not readily apparent in a dot plot?

© Houghton Mifflin Harcourt Publishing Company • Image Credits: Sergey Galushko/Alamy

Using Proportions to Make Inferences

You can use data based on a random sample, along with proportional reasoning, to make inferences or predictions about the population.

EXAMPLE 1   Real World   FL   7.SP.1.2, 7.RP.1.2c

A shipment to a warehouse consists of 3,500 MP3 players. The manager chooses a random sample of 50 MP3 players and finds that 3 are defective. How many MP3 players in the shipment are likely to be defective?

It is reasonable to make a prediction about the population because this sample is random.

STEP 1   Set up a proportion.

$$\frac{\text{defective MP3s in sample}}{\text{size of sample}} = \frac{\text{defective MP3s in population}}{\text{size of population}}$$

STEP 2   Substitute values into the proportion.

$$\frac{3}{50} = \frac{x}{3,500}$$ *Substitute known values. Let x be the number of defective MP3 players in the population.*

$$\frac{3 \cdot 70}{50 \cdot 70} = \frac{x}{3,500}$$ *50 · 70 = 3,500, so multiply the numerator and denominator by 70.*

$$\frac{210}{3,500} = \frac{x}{3,500}$$

$$210 = x$$

Based on the sample, you can predict that 210 MP3 players in the shipment would be defective.

YOUR TURN

5. **What If?** How many MP3 players in the shipment would you predict to be damaged if 6 MP3s in the sample had been damaged?

   _____

Reflect

6. **Check for Reasonableness** How could you use estimation to check if your answer is reasonable?

   _____

   _____

Patrons in the children's section of a local branch library were randomly selected and asked their ages. The librarian wants to use the data to infer the ages of all patrons of the children's section so he can select age appropriate activities. In 3–6, complete each inference. (Explore Activity 1 and 2)

7, 4, 7, 5, 4, 10, 11, 6, 7, 4

1. Make a dot plot of the sample population data.

←―┼→

2. Make a box plot of the sample population data.

←―┼―┼―┼―┼―┼―┼―┼―┼―┼―┼―┼―┼―┼―┼―┼―┼―┼―┼―┼→

3. The most common ages of children that use the library are _____ and _____.

4. The range of ages of children that use the library is from _____ to _____.

5. The median age of children that use the library is _____.

6. A manufacturer fills an order for 4,200 smart phones. The quality inspector selects a random sample of 60 phones and finds that 4 are defective. How many smart phones in the order are likely to be defective? (Example 1)

About _____ smart phones in the order are likely to be defective.

7. Part of the population of 4,500 elk at a wildlife preserve is infected with a parasite. A random sample of 50 elk shows that 8 of them are infected. How many elk are likely to be infected? (Example 1)
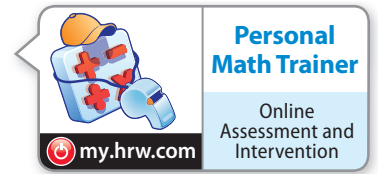
_____

? ESSENTIAL QUESTION CHECK-IN

8. How can you use a random sample of a population to make predictions?

_____

_____

_____

_____

_____

5.2 Independent Practice

Personal Math Trainer

Online Assessment and Intervention

my.hrw.com

9. A manager samples the receipts of every fifth person who goes through the line. Out of 50 people, 4 had a mispriced item. If 600 people go to this store each day, how many people would you expect to have a mispriced item?

_____

10. Jerry randomly selects 20 boxes of crayons from the shelf and finds 2 boxes with at least one broken crayon. If the shelf holds 130 boxes, how many would you expect to have at least one broken crayon?

_____

11. A random sample of dogs at different animal shelters in a city shows that 12 of the 60 dogs are puppies. The city's animal shelters collectively house 1,200 dogs each year. About how many dogs in all of the city's animal shelters are puppies?

_____

12. Part of the population of 10,800 hawks at a national park are building a nest. A random sample of 72 hawks shows that 12 of them are building a nest. Estimate the number of hawks building a nest in the population.

_____

13. In a wildlife preserve, a random sample of the population of 150 raccoons was caught and weighed. The results, given in pounds, were 17, 19, 20, 21, 23, 27, 28, 28, 28 and 32. Jean made the qualitative statement, "The average weight of the raccoon population is 25 pounds." Is her statement reasonable? Explain.

_____

_____

_____

14. Greta collects the number of miles run each week from a random sample of female marathon runners. Her data are shown below. She made the qualitative statement, "25% of female marathoners run 13 or more miles a week." Is her statement reasonable? Explain. Data: 13, 14, 18, 13, 12, 17, 15, 12, 13, 19, 11, 14, 14, 18, 22, 12

_____

_____

_____

15. A random sample of 20 of the 200 students at Garland Elementary is asked how many siblings each has. The data are ordered as shown. Make a dot plot of the data. Then make a qualitative statement about the population. Data: 0, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 3, 3, 3, 3, 4, 4, 4, 6

_____

_____

16. Linda collects a random sample of 12 of the 98 Wilderness Club members' ages. She makes an inference that most wilderness club members are between 20 and 40 years old. Describe what a box plot that would confirm Linda's inference should look like.

_____

_____

_____

17. **What's the Error?** Kudrey was making a box plot. He first plotted the least and greatest data values. He then divided the distance into half, and then did this again for each half. What did Kudrey do wrong and what did his box plot look like?

_____

_____

_____

_____

18. **Communicating Mathematical Ideas** A dot plot includes all of the actual data values. Does a box plot include any of the actual data values?

_____

_____

_____

_____

19. **Make a Conjecture** Sammy counted the peanuts in several packages of roasted peanuts. He found that the bags had 102, 114, 97, 85, 106, 120, 107, and 111 peanuts. Should he make a box plot or dot plot to represent the data? Explain your reasoning.

_____

_____

_____

_____

_____

20. **Represent Real-World Problems** The salaries for the eight employees at a small company are $20,000, $20,000, $22,000, $24,000, $24,000, $29,000, $34,000 and $79,000. Make a qualitative inference about a typical salary at this company. Would an advertisement that stated that the average salary earned at the company is $31,500 be misleading? Explain.

_____

_____

_____

Work Area

© Houghton Mifflin Harcourt Publishing Company

LESSON
5.3

FL 7.SP.1.2

Use data from a random sample ... Generate multiple samples (or simulated samples) of the same size to gauge the variation in estimates or predictions.

Generating Random Samples

How can you generate and use random samples to represent a population?

EXPLORE ACTIVITY 1 Real World   FL 7.SP.1.2

Generating a Random Sample Using Technology

In an earlier lesson, you generated random samples by rolling number cubes. You can also generate random samples by using technology. In Explore Activity 1, you will generate samples using a graphing calculator.

Each of the 200 students in a school will have a chance to vote on one of two names, Tigers or Bears, for the school's athletic teams. A group of students decides to select a random sample of 20 students and ask them for which name they intend to vote. How can the group choose a random sample to represent the entire population of 200 students?

A One way to identify a random sample is to use a graphing calculator to generate random integers.

To simulate choosing 20 students at random from among 200 students:

- Press MATH, scroll right and select **PRB**, then select **5: randInt(**.

- Enter the least value, comma, greatest possible value.

In this specific case, the students will enter **randInt** (⬜ , ⬜)

because there are _____ students in school.

- Hit ENTER _____ times to generate _____ random numbers.

randInt (1, 200)
43
93
75
178

The group gets a list of all the students in the school and assigns a number to each one. The group surveys the students with the given numbers.

Of the 20 students surveyed, 9 chose Tigers. The percent choosing

Tigers was _____. What might the group infer?

_____

_____
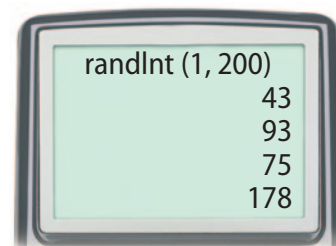
B You can simulate multiple random samples to see how much statistical measures vary for different samples of size 20.

Assume that the 200 students are evenly divided among those voting for Tigers and those voting for Bears. You can generate random numbers and let each number represent a vote. Let numbers from 1 to 100 represent votes for Tigers, and numbers from 101 to 200 represent votes for Bears. For each simulated sample, use randInt(1, 200) and generate 20 numbers.

Perform the simulation 10 times and record how many numbers from 1 to 100 are generated. How many of the samples indicated that there were 9 or fewer votes for Tigers?

_____

Combine your results with those of your classmates. Make a dot plot showing the number of numbers from 1 to 100 generated in each simulation.

_____

Reflect

1. **Communicate Mathematical Reasoning** Assume that it was accurate to say that the 200 students are evenly divided among those voting for Tigers and those voting for Bears. Based on your results, does it seem likely that in a sample of size 20, there would be 9 or fewer votes for Tigers?

_____

2. **Make a Prediction** Based on your answers, do you think it is likely that Tigers will win? Explain.

_____

3. **Multiple Representations** Suppose you wanted to simulate a random sample for the situation in Explore Activity 1 without using technology. One way would be to use marbles of two different colors to represent students choosing the different names. Describe how you could perform a simulation.

_____

_____

_____

_____

_____

_____

Generating a Random Sample without Technology

A tree farm has a 100 acre square field arranged in a 10-by-10 array. The farmer wants to know the average number of trees per acre. Each cell in the table represents an acre. The number in each cell represents the number of trees on that acre.

22	24	27	29	31	24	27	29	30	25
37	22	60	53	62	42	64	53	41	62
61	54	57	34	44	66	39	60	65	40
45	33	64	36	33	51	62	66	42	42
37	34	57	33	47	43	66	33	61	66
66	45	46	67	60	59	51	46	67	48
53	46	35	35	55	56	61	46	38	64
55	51	54	62	55	58	51	45	41	53
61	38	48	48	43	59	64	48	49	47
41	53	53	59	58	48	62	53	45	59

The farmer decides to choose a random sample of 10 of the acres.

A To simulate the random selection, number the table columns 1–10 from left to right, and the rows 1–10 from top to bottom. Write the numbers 1–10 on identical pieces of paper. Place the pieces into a bag. Draw one at random, replace it, and draw another. Let the first number represent a table column, and the second represent a row. For instance, a draw of 2 and then 3 represents the cell in the second column and third row of the table, an acre containing 54 trees. Repeat this process 9 more times.

B Based on your sample, predict the average number of trees per acre. How does your answer compare with the actual mean number, 48.4?

_____

_____

C Compare your answer to **B** with several of your classmates' answers. Do they vary a lot? Is it likely that you can make a valid prediction about the average number of trees per acre? Explain.

_____

Reflect

4. **Communicate Mathematical Ideas** Suppose that you use the method in **A** to collect a random sample of 25 acres. Do you think any resulting prediction would be more or less reliable than your original one? Explain.

_____

_____

5. **Multiple Representations** How could you use technology to select the acres for your sample?

_____

_____

Guided Practice

A manufacturer gets a shipment of 600 batteries of which 50 are defective. The store manager wants to be able to test random samples in future shipments. She tests a random sample of 20 batteries in this shipment to see whether a sample of that size produces a reasonable inference about the entire shipment. (Explore Activities 1 and 2)

1. The manager selects a random sample using the formula

   randInt $\left(\boxed{}, \boxed{} \right)$ to generate _____ random numbers.

2. She lets numbers from 1 to _____ represent defective batteries, and

   _____ to _____ represent working batteries. She generates this list: 120, 413, 472, 564, 38, 266, 344, 476, 486, 177, 26, 331, 358, 131, 352, 227, 31, 253, 31, 277.

3. Does the sample produce a reasonable inference?

   _____

   _____

? ESSENTIAL QUESTION CHECK-IN

4. What can happen if a sample is too small or is not random?

   _____

   _____

5.3 Independent Practice

Personal
Math Trainer

Online
Assessment and
Intervention

my.hrw.com

Maureen owns three bagel shops. Each shop sells 500 bagels per day. Maureen asks her store managers to use a random sample to see how many whole-wheat bagels are sold at each store each day. The results are shown in the table. Use the table for 5–7.

	Total bagels in sample	Whole-wheat bagels
Shop A	50	10
Shop B	100	23
Shop C	25	7

5. If you assume the samples are representative, how many whole-wheat bagels might you infer are sold at each store?

_____

6. Rank the samples for the shops in terms of how representative they are likely to be. Explain your rankings.

_____

_____

7. Which sample or samples should Maureen use to tell her managers how many whole-wheat bagels to make each day? Explain.

_____

_____

_____

8. In a shipment of 1,000 T-shirts, 75 do not meet quality standards. The table below simulates a manager's random sample of 20 T-shirts to inspect. For the simulation, the integers 1 to 75 represent the below-standard shirts.

124	876	76	79	12	878	86	912	435	91
340	213	45	678	544	271	714	777	812	80

In the sample, how many of the shirts are below quality standards? _____

If someone used the sample to predict the number of below standard shirts in the shipment, how far off would the prediction be?

_____

_____

9. Multistep A 64-acre coconut farm is arranged in an 8-by-8 array. Mika wants to know the average number of coconut palms on each acre. Each cell in the table represents an acre of land. The number in each cell tells how many coconut palms grow on that particular acre.

56	54	40	34	44	66	43	65
66	33	42	36	33	51	62	63
33	34	66	33	47	43	66	61
46	35	48	67	60	59	52	67
46	32	64	35	55	47	61	38
45	51	53	62	55	58	51	41
48	38	47	48	43	59	64	54
53	67	59	59	58	48	62	45

a. The numbers in green represent Mika's random sample of 10 acres. What is the average number of coconut palms on the randomly selected acres?

_____

b. Project the number of palms on the entire farm.

_____

H.O.T. FOCUS ON HIGHER ORDER THINKING

10. Draw Conclusions A random sample of 15 of the 78 competitors at a middle school gymnastics competition are asked their height. The data set lists the heights in inches: 55, 57, 57, 58, 59, 59, 59, 59, 59, 61, 62, 62, 63, 64, 66. What is the mean height of the sample? Do you think this is a reasonable prediction of the mean height of all competitors? Explain.

_____

_____

_____

11. Critical Thinking The six-by-six grid contains the ages of actors in a youth Shakespeare festival. Describe a method for randomly selecting 8 cells by using number cubes. Then calculate the average of the 8 values you found.

12	15	16	9	21	11
9	10	14	10	13	12
16	21	14	12	8	14
16	20	9	16	19	18
17	14	12	15	10	15
12	20	14	10	12	9

_____

_____

_____

_____

_____

12. Communicating Mathematical Ideas Describe how the size of a random sample affects how well it represents a population as a whole.

_____

_____

Work Area

Ready to Go On?

5.1 Populations and Samples

1. A company uses a computer to identify their 600 most loyal customers from its database and then surveys those customers to find out how they like their service. Identify the population and determine whether the sample is random or biased.

_____

_____

5.2 Making Inferences from a Random Sample

2. A university has 30,330 students. In a random sample of 270 students, 18 speak three or more languages. Predict the number of students at the university who speak three or more languages.

_____

5.3 Generating Random Samples

A store receives a shipment of 5,000 MP3 players. In a previous shipment of 5,000 MP3 players, 300 were defective. A store clerk generates random numbers to simulate a random sample of this shipment. The clerk lets the numbers 1 through 300 represent defective MP3 players, and the numbers 301 through 5,000 represent working MP3 players. The results are given.

13 2,195 3,873 525 900 167 1,094 1,472 709 5,000

3. Based on the sample, how many of the MP3 players might the clerk predict would be defective?

_____

4. Can the manufacturer assume the prediction is valid? Explain.

_____

? ESSENTIAL QUESTION

5. How can you use random samples to solve real-world problems?

_____

_____

Assessment Readiness

Selected Response

1. A farmer is using a random sample to predict the number of broken eggs in a shipment of 3,000 eggs. Using a calculator, the farmer generates the following random numbers. The numbers 1–250 represent broken eggs.

477	2,116	1,044	81	619	755
2,704	900	238	1,672	187	1,509

Based on this sample, how many broken eggs might the farmer expect?

- Ⓐ 250 broken eggs
- Ⓑ 375 broken eggs
- Ⓒ 750 broken eggs
- Ⓓ 900 broken eggs

2. A middle school has 490 students. Mae surveys a random sample of 60 students and finds that 24 of them have pet dogs. How many students are likely to have pet dogs?

- Ⓐ 98
- Ⓑ 196
- Ⓒ 245
- Ⓓ 294

3. A circular garden has a diameter of 18 meters. What is the circumference of the garden?

- Ⓐ 28.26 meters
- Ⓑ 56.52 meters
- Ⓒ 254.34 meters
- Ⓓ 1,017.36 meters

4. Which of the following is a random sample?

- Ⓐ A radio DJ asks the first 10 listeners who call in if they liked the last song.
- Ⓑ 20 customers at a chicken restaurant are surveyed on their favorite food.
- Ⓒ A polling organization numbers all registered voters, then generates 800 random integers. The polling organization interviews the 800 voters assigned those numbers.
- Ⓓ Rebecca used an email poll to survey 100 students about how often they use the internet.

Mini-Task

5. Each cell in the table represents the number of people who work in one 25-square-block section of the town of Middleton. The mayor uses a random sample to estimate the average number of workers per block.

47	61	56	48	ⓢ56
ⓢ60	39	63	60	46
51	58	49	63	45
55	58	ⓢ50	ⓢ43	48
ⓢ62	ⓢ53	44	66	55

a. The circled numbers represent the mayor's random sample. What is the mean number of workers in this sample?

_____

b. Predict the number of workers in the entire 25-block section of Middleton.

_____

Analyzing and Comparing Data

? ESSENTIAL QUESTION

How can you use solve real-world problems by analyzing and comparing data?

▶

Real-World Video

Scientists place radio frequency tags on some animals within a population of that species. Then they track data, such as migration patterns, about the animals.

🔴 **my.hrw.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.

Fractions, Decimals, and Percents

EXAMPLE Write $\frac{13}{20}$ as a decimal and a percent.

$$\begin{array}{r} 0.65 \\ 20\overline{)13.00} \\ -12\ 0 \\ \hline 1\ 00 \\ -1\ 00 \\ \hline 0 \end{array}$$

Write the fraction as a division problem.
Write a decimal point and zeros in the dividend.
Place a decimal point in the quotient.

$0.65 = 65\%$ Write the decimal as a percent.

Write each fraction as a decimal and a percent.

1. $\frac{7}{8}$ _____   **2.** $\frac{4}{5}$ _____   **3.** $\frac{1}{4}$ _____   **4.** $\frac{3}{10}$ _____

5. $\frac{19}{20}$ _____   **6.** $\frac{7}{25}$ _____   **7.** $\frac{37}{50}$ _____   **8.** $\frac{29}{100}$ _____

Find the Median and Mode

EXAMPLE 17, 14, 13, 16, 13, 11
11, 13, 13, 14, 16, 17

Order the data from least to greatest.

$\text{median} = \frac{13 + 14}{2} = 13.5$

The median is the middle item or the average of the two middle items.

$\text{mode} = 13$

The mode is the item that appears most frequently in the data.

Find the median and the mode of the data.

9. 11, 17, 7, 6, 7, 4, 15, 9 _____   **10.** 43, 37, 49, 51, 56, 40, 44, 50, 36 _____

Find the Mean

EXAMPLE 17, 14, 13, 16, 13, 11

$\text{mean} = \frac{17 + 14 + 13 + 16 + 13 + 11}{6}$

$= \frac{84}{6}$

$= 14$

The mean is the sum of the data items divided by the number of items.

Find the mean of the data.

11. 9, 16, 13, 14, 10, 16, 17, 9 _____   **12.** 108, 95, 104, 96, 97, 106, 94 _____

Reading Start-Up

Visualize Vocabulary

Use the ✔ words to complete the right column of the chart.

Statistical Data		
Definition	**Example**	**Review Word**
A group of facts.	Grades on history exams: 85, 85, 90, 92, 94	
The middle value of a data set.	85, 85, 90, 92, 94	
A value that summarizes a set of values, found through addition and division.	Results of the survey show that students typically spend 5 hours a week studying.	

Understand Vocabulary

Complete each sentence using the preview words.

1. A display that uses values from a data set to show how the

   values are spread out is a _____.

2. A _____ uses a number line to display data.

Active Reading

Layered Book Before beginning the module, create a layered book to help you learn the concepts in this module. Label the first flap with the module title. Label the remaining flaps with the lesson titles. 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.

FL 7.SP.2.3

Informally assess the degree of visual overlap of two numerical data distributions with similar variabilities, measuring the difference between the centers by expressing it as a multiple of a measure of variability.

Key Vocabulary

measure of center *(medida de centro)*
A measure used to describe the middle of a data set; the mean and median are measures of center.

What It Means to You

You will compare two populations based on random samples.

UNPACKING EXAMPLE 7.SP.2.3

Melinda surveys a random sample of 16 students from two college dorms to find the average number of hours of sleep they get. Use the results shown in the dot plots to compare the two populations.

Average Daily Hours of Sleep

Anderson Hall   Jones Hall
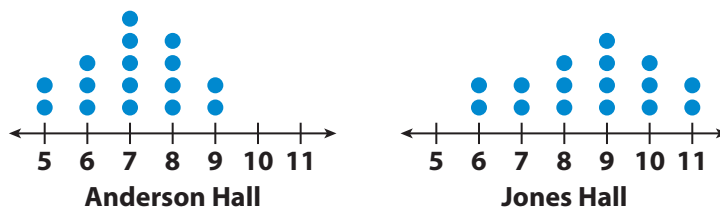
Students in Jones Hall tend to sleep more than students in Anderson Hall, but the variation in the data sets is similar.

FL 7.SP.2.3

Informally assess… distributions with similar variabilities, measuring the difference between the centers by expressing it as a multiple of a measure of variability.

Key Vocabulary

measure of spread *(medida de la dispersión)*
A measure used to describe how much a data set varies; the range, IQR, and mean absolute deviation are measures of spread.

Visit **my.hrw.com** to see all **Florida Math Standards** unpacked.

my.hrw.com

What It Means to You

You will compare two groups of data by comparing the difference in the means to the variability.

UNPACKING EXAMPLE 7.SP.2.3

The tables show the number of items that students in a class answered correctly on two different math tests. How does the difference in the means of the data sets compare to the variability?

Items Correct on Test 1
20, 13, 18, 19, 15, 18, 20, 20, 15, 15, 19, 18

Mean: 17.5; Mean absolute deviation: 2

Items Correct on Test 2
8, 12, 12, 8, 15, 16, 14, 12, 13, 9, 14, 11

Mean: 12; Mean absolute deviation: 2

The means of the two data sets differ by $\frac{17.5-12}{2}=2.75$ times the variability of the data sets.

Comparing Data Displayed in Dot Plots

? ESSENTIAL QUESTION

How do you compare two sets of data displayed in dot plots?

EXPLORE ACTIVITY Real World   FL 7.SP.2.4

Analyzing Dot Plots

You can use dot plots to analyze a data set, especially with respect to its center and spread.

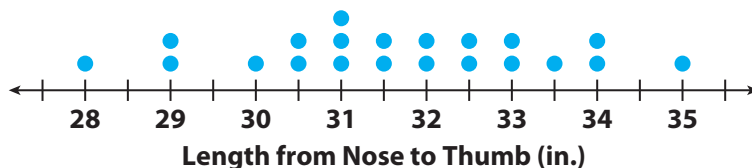**People once used body parts for measurements. For example, an inch was the width of a man's thumb. In the 12th century, King Henry I of England stated that a yard was the distance from his nose to his outstretched arm's thumb. The dot plot shows the different lengths, in inches, of the "yards" for students in a 7th grade class.**

Length from Nose to Thumb (in.)

A Describe the shape of the dot plot. Are the dots evenly distributed or grouped on one side?

_____

B Describe the center of the dot plot. What single dot would best represent the data?

_____

_____

C Describe the spread of the dot plot. Are there any outliers?

_____

_____

Reflect

1. Calculate the mean, median, and range of the data in the dot plot.

_____

Comparing Dot Plots Visually

You can compare dot plots visually using various characteristics, such as center, spread, and shape.

EXAMPLE 1

Real World

FL **7.SP.2.3**

The dot plots show the heights of 15 high school basketball players and the heights of 15 high school softball players.

5'0"   5'2"   5'4"   5'6"
Softball Players' Heights

5'2"   5'4"   5'6"   5'8"   5'10"   6'0"
Basketball Players' Heights
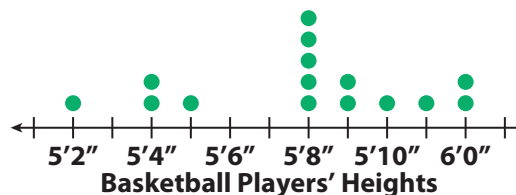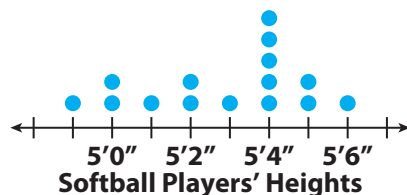
A Visually compare the shapes of the dot plots.

Softball: All the data is 5'6" or less.
Basketball: Most of the data is 5'8" or greater.
As a group, the softball players are shorter than the basketball players.

B Visually compare the centers of the dot plots.

Softball: The data is centered around 5'4".
Basketball: The data is centered around 5'8".
This means that the most common height for the softball players is 5 feet 4 inches, and for the basketball players 5 feet 8 inches.

C Visually compare the spreads of the dot plots.

Softball: The spread is from 4'11" to 5'6".
Basketball: The spread is from 5'2" to 6'0".
There is a greater spread in heights for the basketball players.

Math Talk

Mathematical Practices

How do the heights of field hockey players compare with the heights of softball and basketball players?

YOUR TURN

2. Visually compare the dot plot of heights of field hockey players to the dot plots for softball and basketball players.

5'0   5'2   5'4   5'6
Field Hockey Players' Heights
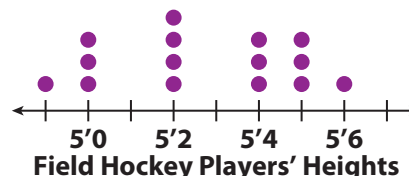
Shape: _____

_____

Center: _____

_____

Spread: _____

_____

Personal Math Trainer

Online Assessment and Intervention

my.hrw.com

Comparing Dot Plots Numerically

You can also compare the shape, center, and spread of two dot plots numerically by calculating values related to the center and spread. Remember that outliers can affect your calculations.

EXAMPLE 2   Real World   FL 7.SP.2.4

Numerically compare the dot plots of the number of hours a class of students exercises each week to the number of hours they play video games each week.

Exercise (h)
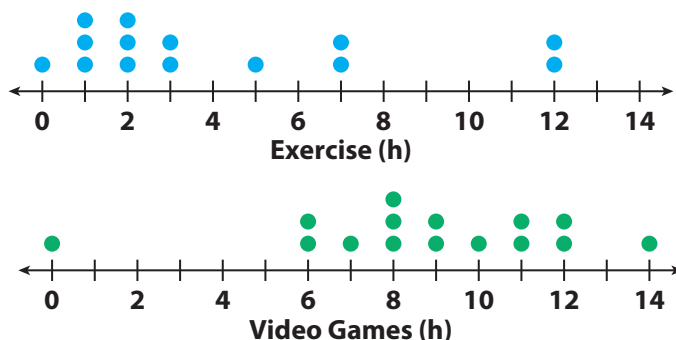
Video Games (h)

A Compare the shapes of the dot plots.

Exercise: Most of the data is less than 4 hours.
Video games: Most of the data is 6 hours or greater.

B Compare the centers of the dot plots by finding the medians.

Median for exercise: 2.5 hours. Even though there are outliers at 12 hours, most of the data is close to the median.
Median for video games: 9 hours. Even though there is an outlier at 0 hours, these values do not seem to affect the median.

C Compare the spreads of the dot plots by calculating the range.

Exercise range with outlier: $12 - 0 = 12$ hours
Exercise range without outlier: $7 - 0 = 7$ hours
Video games range with outlier: $14 - 0 = 14$ hours
Video games range without outlier: $14 - 6 = 8$ hours

Math Talk
Mathematical Practices
How do outliers affect the results of this data?

YOUR TURN

3. Calculate the median and range of the data in the dot plot. Then compare the results to the dot plot for Exercise in Example 2.
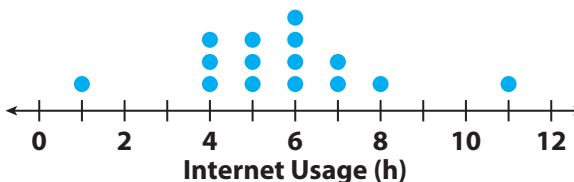
Internet Usage (h)

_____

_____

_____

The dot plots show the number of miles run per week for two different classes. For 1–5, use the dot plots shown.

Class A (mi)

Class B (mi)
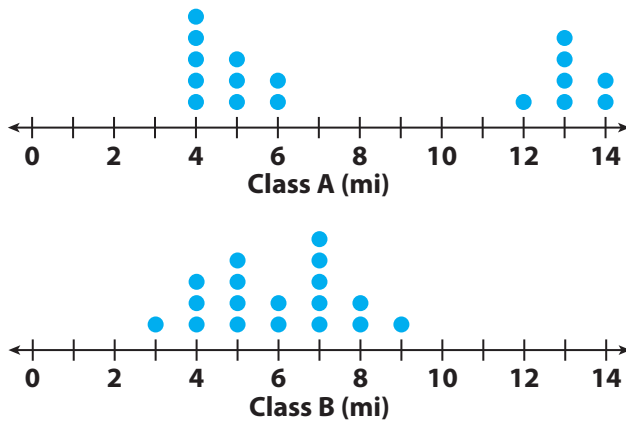
1. Compare the shapes of the dot plots.

_____

_____

2. Compare the centers of the dot plots.

_____

_____

3. Compare the spreads of the dot plots.

_____

_____

4. Calculate the medians of the dot plots.

_____

5. Calculate the ranges of the dot plots.

_____

? ESSENTIAL QUESTION CHECK-IN

6. What do the medians and ranges of two dot plots tell you about the data?

_____

_____

_____

6.1 Independent Practice

FL 7.SP.2.3, 7.SP.2.4

The dot plot shows the number of letters in the spellings of the 12 months. Use the dot plot for 7–10.

Number of Letters
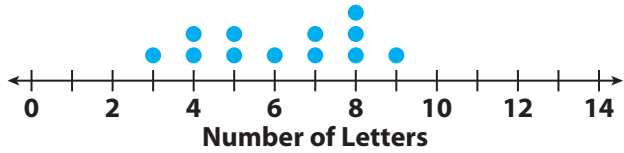
7. Describe the shape of the dot plot.

_____

8. Describe the center of the dot plot.

_____

9. Describe the spread of the dot plot.

_____

10. Calculate the mean, median, and range of the data in the dot plot.

_____

The dot plots show the mean number of days with rain per month for two cities.

Number of Days of Rain for Montgomery, AL

Number of Days of Rain for Lynchburg, VA
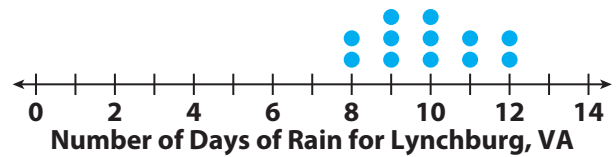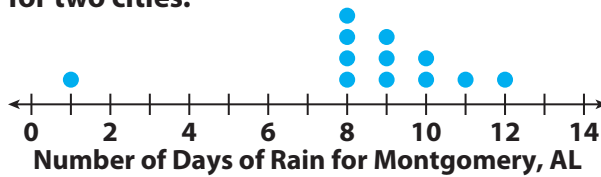
11. Compare the shapes of the dot plots.

_____

_____

12. Compare the centers of the dot plots.

_____

_____

13. Compare the spreads of the dot plots.

_____

_____

14. What do the dot plots tell you about the two cities with respect to their average monthly rainfall?

_____

_____

The dot plots show the shoe sizes of two different groups of people.

Group A Shoe Sizes

Group B Shoe Sizes
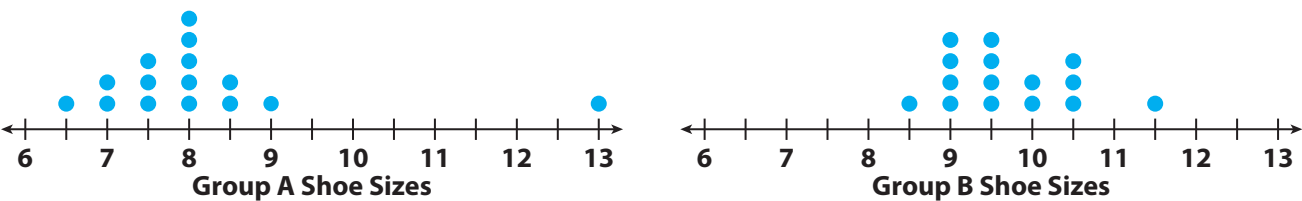
15. Compare the shapes of the dot plots.

_____

_____

16. Compare the medians of the dot plots.

_____

17. Compare the ranges of the dot plots (with and without the outliers).

_____

_____

18. Make A Conjecture Provide a possible explanation for the results of the dot plots.

_____

_____

H.O.T. **FOCUS ON HIGHER ORDER THINKING**

Work Area

19. Analyze Relationships Can two dot plots have the same median and range but have completely different shapes? Justify your answer using examples.

_____

_____

_____

_____

_____

20. Draw Conclusions What value is most affected by an outlier, the median or the range? Explain. Can you see these effects in a dot plot?

_____

_____

_____

Comparing Data Displayed in Box Plots

? **ESSENTIAL QUESTION**

How do you compare two sets of data displayed in box plots?

EXPLORE ACTIVITY *Real World*   FL 7.SP.2.4

Analyzing Box Plots

Box plots show five key values to represent a set of data, the least and greatest values, the lower and upper quartile, and the median. To create a box plot, arrange the data in order, and divide them into four equal-size parts or quarters. Then draw the box and the whiskers as shown.

The number of points a high school basketball player scored during the games he played this season are organized in the box plot shown.

Points Scored
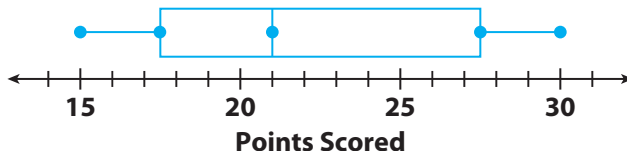
A Find the least and greatest values.

Least value: _____   Greatest value: _____

B Find the median and describe what it means for the data.

_____

_____

_____

C Find and describe the lower and upper quartiles.

_____

_____

_____

D The interquartile range is the difference between the lower and upper quartiles, which is represented by the length of the box. Find the interquartile range.

$Q_3 - Q_1 =$ _____ $-$ _____ $=$ _____

Math Talk
Mathematical Practices

How do the lengths of the whiskers compare? Explain what this means.

Reflect

1. Why is one-half of the box wider than the other half of the box?

_____

_____

Box Plots with Similar Variability

You can compare two box plots numerically according to their centers, or medians, and their spreads, or variability. Range and interquartile range (IQR) are both measures of spread. Box plots with similar variability should have similar boxes and whiskers.
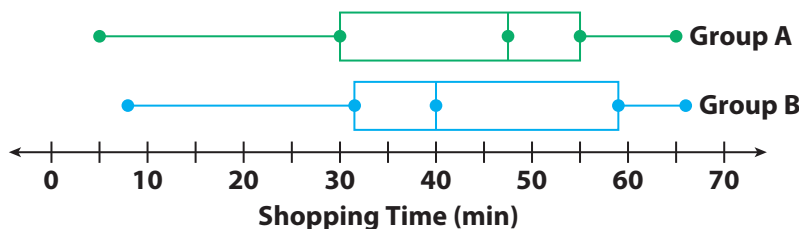
EXAMPLE 1 Real World   FL 7.SP.2.3

The box plots show the distribution of times spent shopping by two different groups.

A Compare the shapes of the box plots.

The positions and lengths of the boxes and whiskers appear to be very similar. In both plots, the right whisker is shorter than the left whisker.

B Compare the centers of the box plots.

Group A's median, 47.5, is greater than Group B's, 40. This means that the median shopping time for Group A is 7.5 minutes more.

C Compare the spreads of the box plots.

The box shows the interquartile range. The boxes are similar.

Group A: $55 - 30 = 25$ min   Group B: About $59 - 32 = 26$ min

The whiskers have similar lengths, with Group A's slightly shorter than Group B's.

Math Talk

Mathematical Practices

Which store has the shopper who shops longest? Explain how you know.

Reflect

2. Which group has the greater variability in the bottom 50% of shopping times? The top 50% of shopping times? Explain how you know.
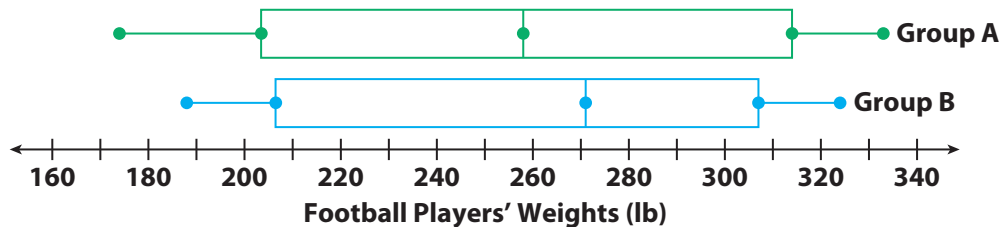
_____

_____

My Notes

YOUR TURN

3. The box plots show the distribution of weights in pounds of two different groups of football players. Compare the shapes, centers, and spreads of the box plots.

Football Players' Weights (lb)
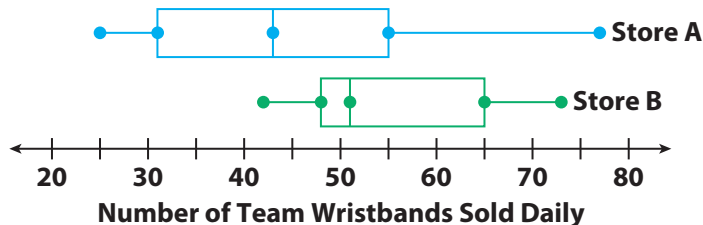
_____

_____

_____

_____

_____

Box Plots with Different Variability

You can compare box plots with greater variability, where there is less overlap of the median and interquartile range.

EXAMPLE 2 Real World FL 7.SP.2.4

The box plots show the distribution of the number of team wristbands sold daily by two different stores over the same time period.

Number of Team Wristbands Sold Daily

A Compare the shapes of the box plots.

Store A's box and right whisker are longer than Store B's.

B Compare the centers of the box plots.

Store A's median is about 43, and Store B's is about 51. Store A's median is close to Store B's minimum value, so about 50% of Store A's daily sales were less than sales on Store B's worst day.

C Compare the spreads of the box plots.

Store A has a greater spread. Its range and interquartile range are both greater. Four of Store B's key values are greater than Store A's corresponding value. Store B had a greater number of sales overall.
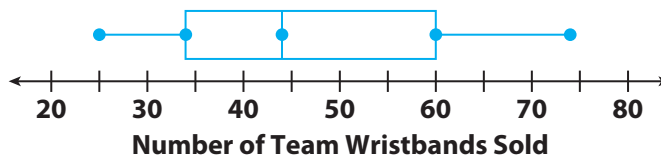
YOUR TURN

4. Compare the shape, center, and spread of the data in the box plot with the data for Stores A and B in the two box plots in Example 2.
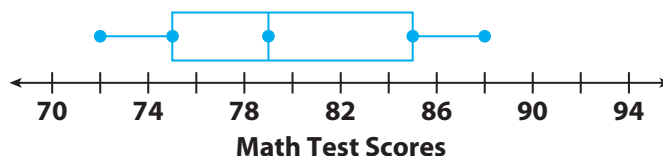
Number of Team Wristbands Sold

_____

_____

_____

Guided Practice

For 1–3, use the box plot Terrence created for his math test scores. Find each value. (Explore Activity)

1. Minimum = _____ Maximum = _____
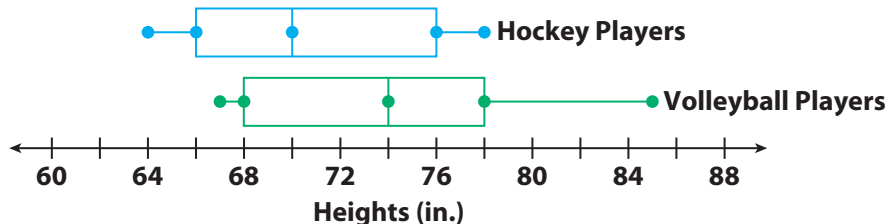
2. Median = _____

3. Range = _____ IQR = _____

Math Test Scores

For 4–7, use the box plots showing the distribution of the heights of hockey and volleyball players. (Examples 1 and 2)

Hockey Players

Volleyball Players

Heights (in.)

4. Which group has a greater median height? _____

5. Which group has the shortest player? _____

6. Which group has an interquartile range of about 10? _____

❓ ESSENTIAL QUESTION CHECK-IN

7. What information can you use to compare two box plots?
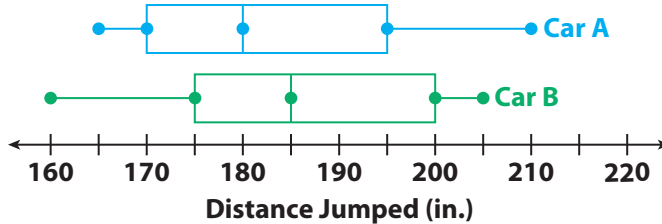
_____

6.2 Independent Practice

FL 7.SP.2.3, 7.SP.2.4

For 8–11, use the box plots of the distances traveled by two toy cars that were jumped from a ramp.

8. Compare the minimum, maximum, and median of the box plots.

_____

_____

_____

_____

_____

9. Compare the ranges and interquartile ranges of the data in box plots.

_____

_____

_____

10. What do the box plots tell you about the jump distances of two cars?

_____

_____

_____

_____

_____

_____

11. Critical Thinking What do the whiskers tell you about the two data sets?

_____

_____

_____

_____

_____

_____

For 12–14, use the box plots to compare the costs of leasing cars in two different cities.
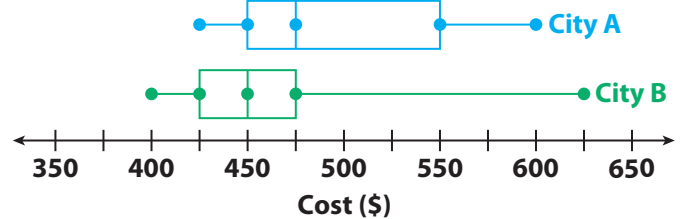
12. In which city could you spend the least amount of money to lease a car? The greatest?

_____

_____

_____

13. Which city has a higher median price? How much higher is it?

_____

14. Make a Conjecture In which city is it more likely to choose a car at random that leases for less than $450? Why?

_____

_____

_____

_____

15. Summarize Look back at the box plots for 12–14 on the previous page. What do the box plots tell you about the costs of leasing cars in those two cities?

_____

_____

_____

_____

16. Draw Conclusions Two box plots have the same median and equally long whiskers. If one box plot has a longer box than the other box plot, what does this tell you about the difference between the data sets?

_____

_____

17. Communicate Mathematical Ideas What you can learn about a data set from a box plot? How is this information different from a dot plot?

_____

_____

_____

_____

_____

_____

_____

_____

_____

18. Analyze Relationships In mathematics, *central tendency* is the tendency of data values to cluster around some central value. What does a measure of variability tell you about the central tendency of a set of data? Explain.

_____

_____

_____

Work Area

FL **7.SP.2.3**

Informally assess ... two numerical data distributions ... measuring the difference between the centers by expressing it as a multiple of a measure of variability. *Also 7.SP.2.4*

? ESSENTIAL QUESTION

How can you use statistical measures to compare populations?

Comparing Differences in Centers to Variability

Recall that to find the mean absolute deviation (MAD) of a data set, first find the mean of the data. Next, take the absolute value of the difference between the mean and each data point. Finally, find the mean of those absolute values.

Math On the Spot
my.hrw.com

EXAMPLE 1 Real World

FL **7.SP.2.3**

The tables show the number of minutes per day students in a class spend exercising and playing video games. What is the difference of the means as a multiple of the mean absolute deviations?

Minutes Per Day Exercising
0, 7, 7, 18, 20, 38, 33, 24, 22, 18, 11, 6

Minutes Per Day Playing Video Games
13, 18, 19, 30, 32, 46, 50, 34, 36, 30, 23, 19

STEP 1 Calculate the mean number of minutes per day exercising.

$0 + 7 + 7 + 18 + 20 + 38 + 33 + 24 + 22 + 18 + 11 + 6 = 204$

$204 \div 12 = 17$ *Divide the sum by the number of students.*

STEP 2 Calculate the mean absolute deviation for the number of minutes exercising.

$|0-17| = 17$ $|7-17| = 10$ $|7-17| = 10$ $|18-17| = 1$

$|20-17| = 3$ $|38-17| = 21$ $|33-17| = 16$ $|24-17| = 7$

$|22-17| = 5$ $|18-17| = 1$ $|11-17| = 6$ $|6-17| = 11$

Find the mean of the absolute values.

$17 + 10 + 10 + 1 + 3 + 21 + 16 + 7 + 5 + 1 + 6 + 11 = 108$

$108 \div 12 = 9$ *Divide the sum by the number of students.*

STEP 3 Calculate the mean number of minutes per day playing video games. Round to the nearest tenth.

$$13 + 18 + 19 + 30 + 32 + 46 + 50 + 34 + 36 + 30 + 23 + 19 = 350$$

$350 \div 12 \approx 29.2$ **Divide the sum by the number of students.**

STEP 4 Calculate the mean absolute deviation for the numbers of minutes playing video games.

$|13 - 29.2| = 16.2$	$|18 - 29.2| = 11.2$	$|19 - 29.2| = 10.2$
$|30 - 29.2| = 0.8$	$|32 - 29.2| = 2.8$	$|46 - 29.2| = 16.8$
$|50 - 29.2| = 20.8$	$|34 - 29.2| = 4.8$	$|36 - 29.2| = 6.8$
$|30 - 29.2| = 0.8$	$|23 - 29.2| = 6.2$	$|19 - 29.2| = 10.2$

Find the mean of the absolute values. Round to the nearest tenth.

$$16.2 + 11.2 + 10.2 + 0.8 + 2.8 + 16.8 + 20.8 + 4.8 + 6.8 + 0.8 + 6.2 + 10.2 = 107.6$$

$107.6 \div 12 \approx 9$ **Divide the sum by the number of students.**

STEP 5 Find the difference in the means.

$29.2 - 17 = 12.2$ **Subtract the lesser mean from the greater mean.**

STEP 6 Write the difference of the means as a multiple of the mean absolute deviations, which are similar but not identical.

$12.2 \div 9 \approx 1.36$ **Divide the difference of the means by the MAD.**

The means of the two data sets differ by about 1.4 times the variability of the two data sets.

YOUR TURN

1. The high jumps in inches of the students on two intramural track and field teams are shown below. What is the difference of the means as a multiple of the mean absolute deviations?

High Jumps for Students on Team 1 (in.)
44, 47, 67, 89, 55, 76, 85, 80, 87, 69, 47, 58

High Jumps for Students on Team 2 (in.)
40, 32, 52, 75, 65, 70, 72, 61, 54, 43, 29, 32

Personal Math Trainer

Online Assessment and Intervention

⏻ my.hrw.com
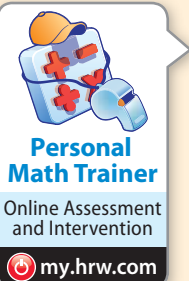
Using Multiple Samples to Compare Populations

Many different random samples are possible for any given population, and their measures of center can vary. Using multiple samples can give us an idea of how reliable any inferences or predictions we make are.
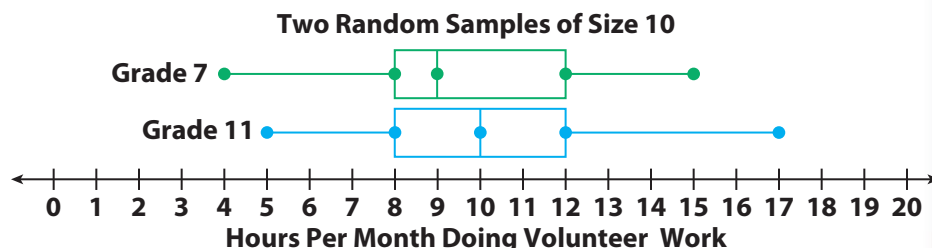
EXAMPLE 2 · Real World · FL · 7.SP.2.4

A group of about 250 students in grade 7 and about 250 students in grade 11 were asked, "How many hours per month do you volunteer?" Responses from one random sample of 10 students in grade 7 and one random sample of 10 students in grade 11 are summarized in the box plots.
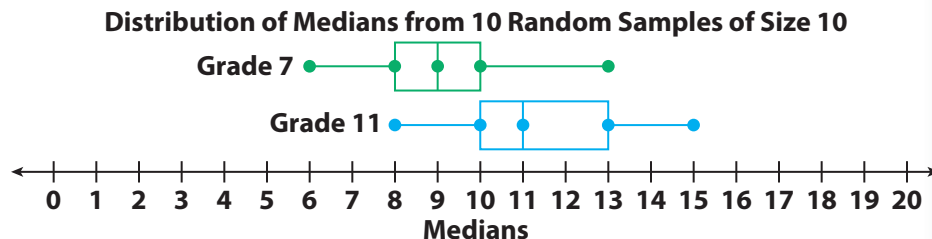
Two Random Samples of Size 10

Grade 7

Grade 11

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Hours Per Month Doing Volunteer Work

How can we tell if the grade 11 students do more volunteer work than the grade 7 students?

STEP 1 The median is higher for the students in grade 11. But there is a great deal of variation. To make an inference for the entire population, it is helpful to consider how the medians vary among multiple samples.

STEP 2 The box plots below show how the medians from 10 different random samples for each group vary.

Distribution of Medians from 10 Random Samples of Size 10

Grade 7

Grade 11

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

The medians vary less than the actual data. Half of the grade 7 medians are within 1 hour of 9. Half of the grade 11 medians are within 1 or 2 hours of 11. Although the distributions overlap, the middle halves of the data barely overlap. This is fairly convincing evidence that the grade 11 students volunteer more than the grade 7 students.

Math Talk
Mathematical Practices

Why doesn't the first box plot establish that students in grade 11 volunteer more than students in grade 7?

YOUR TURN

2. The box plots show the variation in the means for 10 different random samples for the groups in the example. Why do these data give less convincing evidence that the grade 11 students volunteer more?

Distribution of Means from 10 Random Samples of Size 10
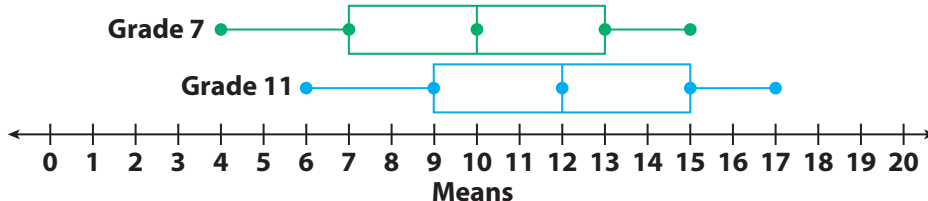
Guided Practice

The tables show the numbers of miles run by the students in two classes. Use the tables in 1–2. (Example 1)

Miles Run by Class 1 Students
12, 1, 6, 10, 1, 2, 3, 10, 3, 8, 3, 9, 8, 6, 8

Miles Run by Class 2 Students
11, 14, 11, 13, 6, 7, 8, 6, 8, 13, 8, 15, 13, 17, 15

1. For each class, what is the mean? What is the mean absolute deviation?

_____

2. The difference of the means is about _____ times the mean absolute deviations.

3. Mark took 10 random samples of 10 students from two schools. He asked how many minutes they spend per day going to and from school. The tables show the medians and the means of the samples. Compare the travel times using distributions of the medians and means. (Example 2)

School A
Medians: 28, 22, 25, 10, 40, 36, 30, 14, 20, 25
Means: 27, 24, 27, 15, 42, 36, 32, 18, 22, 29

School B
Medians: 22, 25, 20, 14, 20, 18, 21, 18, 26, 19
Means: 24, 30, 22, 15, 20, 17, 22, 15, 36, 27

_____

_____

? ESSENTIAL QUESTION CHECK-IN

4. Why is it a good idea to use multiple random samples when making comparative inferences about two populations?

_____

6.3 Independent Practice

FL **7.SP.2.3, 7.SP.2.4**

Personal Math Trainer

Online Assessment and Intervention

my.hrw.com

Josie recorded the average monthly temperatures for two cities in the state where she lives. Use the data for 5–7.

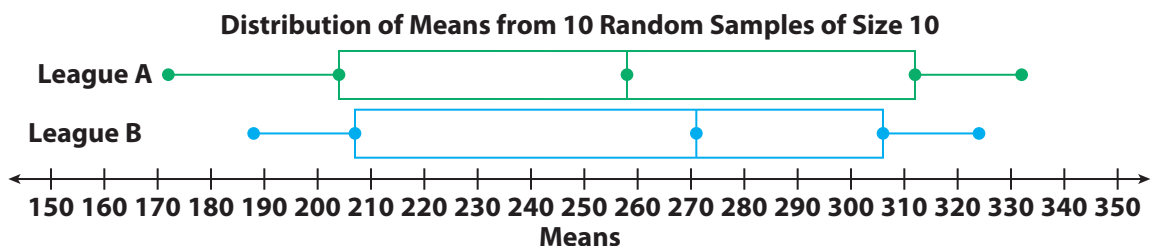| Average Monthly Temperatures for City 1 (°F) |
|---|
| 23, 38, 39, 48, 55, 56, 71, 86, 57, 53, 43, 31 |

Average Monthly Temperatures for City 2 (°F)
8, 23, 24, 33, 40, 41, 56, 71, 42, 38, 28, 16

5. For City 1, what is the mean of the average monthly temperatures? What is the mean absolute deviation of the average monthly temperatures?

_____

6. What is the difference between each average monthly temperature for

City 1 and the corresponding temperature for City 2? _____

7. **Draw Conclusions** Based on your answers to Exercises 5 and 6, what do you think the mean of the average monthly temperatures for City 2 is? What do you think the mean absolute deviation of the average monthly temperatures for City 2 is? Give your answers without actually calculating the mean and the mean absolute deviation. Explain your reasoning.

_____

_____

8. What is the difference in the means as a multiple of the mean absolute

deviations? _____

9. **Make a Conjecture** The box plots show the distributions of mean weights of 10 samples of 10 football players from each of two leagues, A and B. What can you say about any comparison of the weights of the two populations? Explain.

Distribution of Means from 10 Random Samples of Size 10

League A

League B

150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350
Means

_____

_____

10. Justify Reasoning Statistical measures are shown for the ages of middle school and high school teachers in two states.

State A: Mean age of middle school teachers = 38, mean age of high school teachers = 48, mean absolute deviation for both = 6

State B: Mean age of middle school teachers = 42, mean age of high school teachers = 50, mean absolute deviation for both = 4

In which state is the difference in ages between members of the two groups more significant? Support your answer.

_____

_____

11. Analyze Relationships The tables show the heights in inches of all the adult grandchildren of two sets of grandparents, the Smiths and the Thompsons. What is the difference in the medians as a multiple of the ranges?

Heights of the Smiths' Adult Grandchildren (in.)	Heights of the Thompsons' Adult Grandchildren (in.)
64, 65, 68, 66, 65, 68, 69, 66, 70, 67	75, 80, 78, 77, 79, 76, 75, 79, 77, 74

H.O.T. FOCUS ON HIGHER ORDER THINKING

12. Critical Thinking Jill took many samples of 10 tosses of a standard number cube. What might she reasonably expect the median of the medians of the samples to be? Why?

_____

_____

13. Analyze Relationships Elly and Ramon are both conducting surveys to compare the average numbers of hours per month that men and women spend shopping. Elly plans to take many samples of size 10 from both populations and compare the distributions of both the medians and the means. Ramon will do the same, but will use a sample size of 100. Whose results will probably produce more reliable inferences? Explain.

_____

_____

14. Counterexamples Seth believes that it is always possible to compare two populations of numerical values by finding the difference in the means of the populations as a multiple of the mean absolute deviations. Describe a situation that explains why Seth is incorrect.
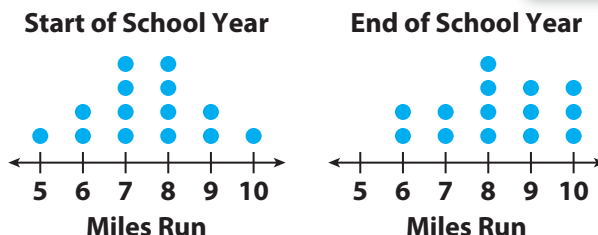
_____

_____

Work Area

Ready to Go On?

## 6.1 Comparing Data Displayed in Dot Plots

The two dot plots show the number of miles run by 14 students at the start and at the end of the school year. Compare each measure for the two dot plots. Use the data for 1–3.

Start of School Year

Miles Run

End of School Year

Miles Run

1. means _____

2. medians _____

3. ranges _____

6.2 Comparing Data Displayed in Box Plots

The box plots show lengths of flights in inches flown by two model airplanes. Use the data for 4–5.

Airplane A

Airplane B

Length of Flight (in.)
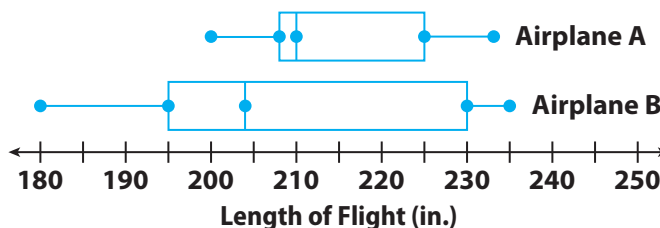
4. Which has a greater median flight length? _____

5. Which has a greater interquartile range? _____

6.3 Using Statistical Measures to Compare Populations

6. Roberta grows pea plants, some in shade and some in sun. She picks 8 plants of each type at random and records the heights.

Shade plant heights (in.)	7	11	11	12	9	12	8	10
Sun plant heights (in.)	21	24	19	19	22	23	24	24

Express the difference in the means as a multiple of their ranges.

_____

? ESSENTIAL QUESTION

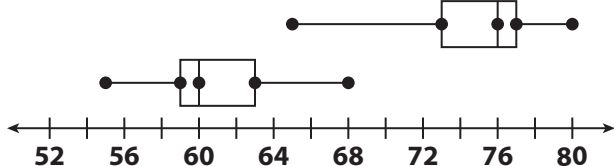**7.** How can you use and compare data to solve real-world problems?

_____

Assessment Readiness

Selected Response

1. Which statement about the data is true?

(A) The difference between the medians is about 4 times the range.

(B) The difference between the medians is about 4 times the IQR.

(C) The difference between the medians is about 2 times the range.

(D) The difference between the medians is about 2 times the IQR.

2. Which is a true statement based on the box plots below?
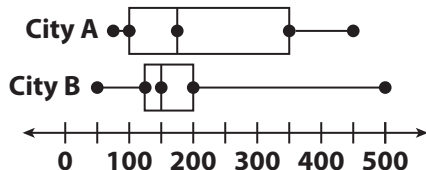
City A
City B

(A) The data for City A has the greater range.

(B) The data for City B is more symmetric.

(C) The data for City A has the greater interquartile range.

(D) The data for City B has the greater median.

3. Which of the following is **not** a possible cross section of a cylinder?

(A) a circle

(B) an oval

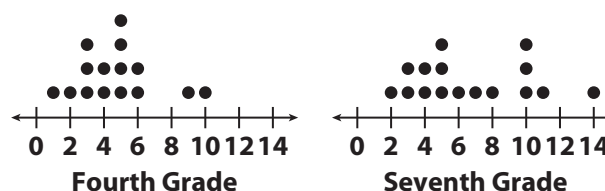**(C)** a rectangle

(D) a triangle

4. Which is a true statement based on the dot plots below?
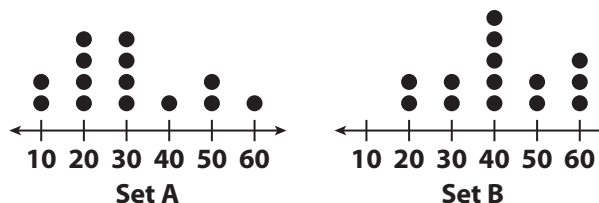
Set A   Set B

(A) Set A has the lesser range.

(B) Set B has the greater median.

(C) Set A has the greater mean.

(D) Set B is less symmetric than Set A.

Mini-Task

5. The dot plots show the lengths of a random sample of words in a fourth-grade book and a seventh-grade book.

Fourth Grade   Seventh Grade

a. Compare the shapes of the plots.

_____

_____

_____

b. Compare the ranges of the plots. Explain what your answer means in terms of the situation.

_____

_____

_____

_____

Study Guide Review

Random Samples and Populations

Key Vocabulary
biased sample *(muestra sesgada)*
population *(población)*
random sample *(muestra aleatoria)*
sample *(muestra)*

? ESSENTIAL QUESTION

How can you use random samples and populations to solve real-world problems?

EXAMPLE 1

An engineer at a lightbulb factory chooses a random sample of 100 lightbulbs from a shipment of 2,500 and finds that 2 of them are defective. How many lightbulbs in the shipment are likely to be defective?

$$\frac{\text{defective lightbulbs}}{\text{size of sample}} = \frac{\text{defective lightbulbs in population}}{\text{size of population}}$$

$$\frac{2}{100} = \frac{x}{2,500}$$

$$\frac{2 \cdot 25}{100 \cdot 25} = \frac{x}{2,500}$$

$$x = 50$$

In a shipment of 2,500 lightbulbs, 50 are likely to be defective.

EXAMPLE 2

The 300 students in a school are about to vote for student body president. There are two candidates, Jay and Serena, and each candidate has about the same amount of support. Use a simulation to generate a random sample. Interpret the results.

Step 1: Write the digits 0 through 9 on 10 index cards, one digit per card. Draw and replace a card three times to form a 3-digit number. For example, if you draw 0-4-9, the number is 49. If you draw 1-0-8, the number is 108. Repeat this process until you have a sample of 30 3-digit numbers.

Step 2: Let the numbers from 1 to 150 represent votes for Jay and the numbers from 151 to 300 represent votes for Serena. For example:

Jay: 83, 37, 16, 4, 127, 93, 9, 62, 91, 75, 13, 35, 94, 26, 60, 120, 36, 73

Serena: 217, 292, 252, 186, 296, 218, 284, 278, 209, 296, 190, 300

Step 3: Notice that 18 of the 30 numbers represent votes for Jay. The results suggest that Jay will receive $\frac{18}{30} = 60\%$ of the 300 votes, or 180 votes.

Step 4: Based on this one sample, Jay will win the election. The results of samples can vary. Repeating the simulation many times and looking at the pattern across the different samples will produce more reliable results.

EXERCISES

1. Molly uses the school directory to select, at random, 25 students from her school for a survey on which sports people like to watch on television. She calls the students and asks them, "Do you think basketball is the best sport to watch on television?" (Lesson 5.1)

a. Did Molly survey a random sample or a biased sample of the students at her school?

_____

b. Was the question she asked an unbiased question? Explain your answer.

_____

_____

2. There are 2,300 licensed dogs in Clarkson. A random sample of 50 of the dogs in Clarkson shows that 8 have ID microchips implanted. How many dogs in Clarkson are likely to have ID microchips implanted? (Lesson 5.2)

_____

3. A store gets a shipment of 500 MP3 players. Twenty-five of the players are defective, and the rest are working. A graphing calculator is used to generate 20 random numbers to simulate a random sample of the players. (Lesson 5.3)

A list of 20 randomly generated numbers representing MP3 players is:

474	77	101	156	378	188	116	458	230	333
78	19	67	5	191	124	226	496	481	161

a. Let numbers 1 to 25 represent players that are _____.

b. Let numbers 21 to 500 represent players that are _____.

c. How many players in this sample are expected to be

defective? _____

d. If 300 players are chosen at random from the shipment, how many are expected to be defective based on the sample? Does the sample provide a reasonable inference? Explain.

_____
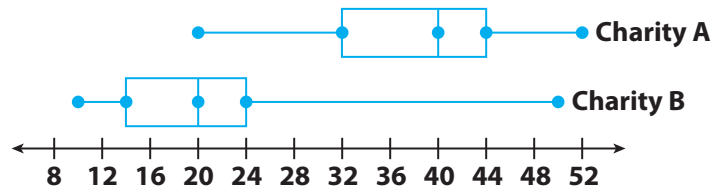
_____

_____

Analyzing and Comparing Data

? ESSENTIAL QUESTION

How can you solve real-world problems by analyzing and comparing data?

EXAMPLE

The box plots show amounts donated to two charities at a fundraising drive. Compare the shapes, centers and spreads of the box plots.

Charity A
Charity B
8 12 16 20 24 28 32 36 40 44 48 52

Shapes: The lengths of the boxes and overall plot lengths are fairly similar, but while the whiskers for Charity A are similar in length, Charity B has a very short whisker and a very long whisker.

Centers: The median for Charity A is $40, and for Charity B is $20.

Spreads: The interquartile range for Charity A is $44 - 32 = 12$. The interquartile range for Charity B is slightly less, $24 - 14 = 10$.

The donations varied more for Charity B and were lower overall.

EXERCISES

The dot plots show the number of hours a group of students spends online each week, and how many hours they spend reading. Compare the dot plots visually. (Lesson 6.1)

0 1 2 3 4 5 6 7
Time Online (h)

0 1 2 3 4 5 6 7
Time Reading (h)
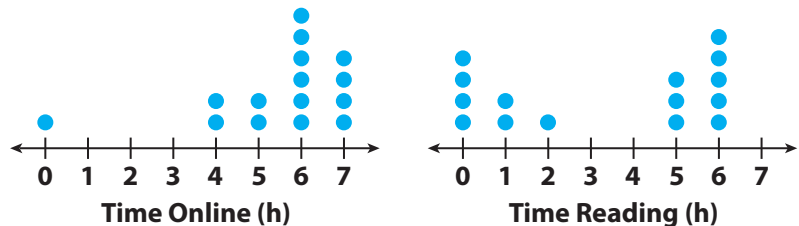
1. Compare the shapes, centers, and spreads of the dot plots.

Shape: _____

_____

Center: _____

_____

Spread: _____

_____

2. Calculate the medians of the dot plots. _____

3. Calculate the ranges of the dot plots. _____

4. The average times (in minutes) a group of students spends studying and watching TV per school day are given. **(Lesson 6.3)**

      Studying: 25, 30, 35, 45, 60, 60, 70, 75
   Watching TV: 0, 35, 35, 45, 50, 50, 70, 75

  a. Find the mean times for studying and for watching TV.

    _____

  b. Find the mean absolute deviations (MADs) for each data set.

    _____

  c. Find the difference of the means as a multiple of the MAD, to two decimal places.

    _____

Unit 3 Performance Tasks

1. **CAREERS IN MATH** **Entomologist** An entomologist is studying how two different types of flowers appeal to butterflies. The box-and-whisker plots show the number of butterflies that visited one of two different types of flowers in a field. The data were collected over a two-week period, for one hour each day.

Number of Butterflies
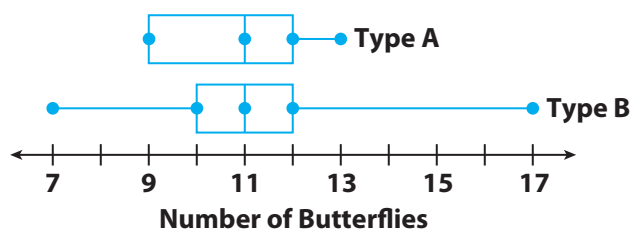
  a. Find the median, range, and interquartile range for each data set.

    _____

    _____

  b. Which measure makes it appear that flower type A had a more consistent number of butterfly visits? Which measure makes it appear that flower type B did? If you had to choose one flower as having the more consistent visits, which would you choose? Explain your reasoning.
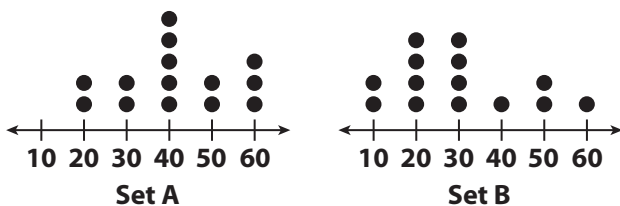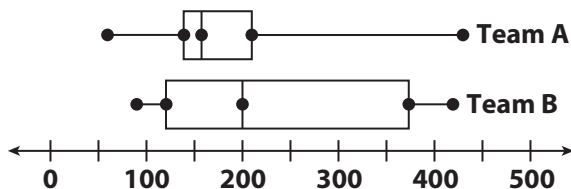
    _____

    _____

    _____

    _____

Selected Response

**1.** Which is a true statement based on the dot plots below?

Set A

Set B

Ⓐ Set B has the greater range.

Ⓑ Set B has the greater median.

Ⓒ Set B has the greater mean.

Ⓓ Set A is less symmetric than Set B.

2. Which is a solution to the equation $7g - 2 = 47$?

Ⓐ $g = 5$

Ⓑ $g = 6$

Ⓒ $g = 7$

Ⓓ $g = 8$

**3.** Which is a true statement based on the box plots below?

Team A

Team B

Ⓐ The data for Team B have the greater range.

Ⓑ The data for Team A are more symmetric.

Ⓒ The data for Team B have the greater interquartile range.

Ⓓ The data for Team A have the greater median.

4. Which is the best way to choose a random sample of people from a sold-out movie audience for a survey?

Ⓐ Survey all audience members who visit the restroom during the movie.

Ⓑ Assign each seat a number, write each number on a slip of paper, and then draw several slips from a hat. Survey the people in those seats.

Ⓒ Survey all of the audience members who sit in the first or last row of seats in the movie theater.

Ⓓ Before the movie begins, ask for volunteers to participate in a survey. Survey the first twenty people who volunteer.

5. What is the solution of $-3x + 2 < 11$?

Ⓐ $x < 3$ Ⓒ $x > 3$

Ⓑ $x < -3$ Ⓓ $x > -3$

6. A survey asked 100 students in a school to name the temperature at which they feel most comfortable. The box plot below shows the results for temperatures in degrees Fahrenheit. Which could you infer based on the box plot below?
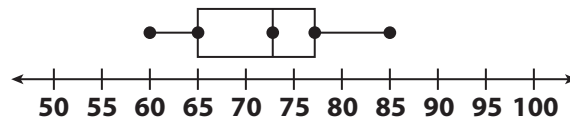
Ⓐ Most students prefer a temperature less than 65 degrees.

Ⓑ Most students prefer a temperature of at least 70 degrees.

Ⓒ Almost no students prefer a temperature of less than 75 degrees.

Ⓓ Almost no students prefer a temperature of more than 65 degrees.

© Houghton Mifflin Harcourt Publishing Company

7. The box plots below show data from a survey of students under 14 years old. They were asked on how many days in a month they read and draw. Based on the box plots, which is a true statement about students?
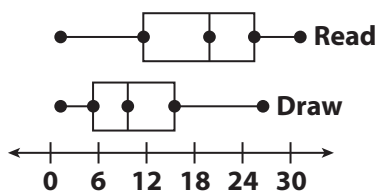

(A) Most students draw at least 12 days a month.

(B) Most students read less than 12 days a month.

(C) Most students read more often than they draw.

(D) Most students draw more often than they read.

> **Hot Tip!**
>
> **Use logic to eliminate answer choices that are incorrect. This will help you to make an educated guess if you are having trouble with the question.**

8. Which describes the relationship between ∠NOM and ∠JOK in the diagram?
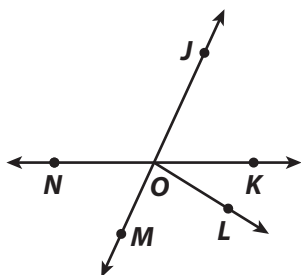
(A) adjacent angles

(B) complementary angles

(C) supplementary angles

(D) vertical angles

Mini-Task

9. The tables show the typical number of minutes spent exercising each week for a group of fourth-grade students and a group of seventh-grade students.

Weekly Exercising (minutes), 4th Grade
120, 75, 30, 30, 240, 90, 100, 180, 125, 300

Weekly Exercising (minutes), 7th Grade
410, 145, 240, 250, 125, 95, 210, 190, 245, 300

a. What is the mean number of minutes spent exercising for fourth graders? For seventh graders?

_____

b. What is the mean absolute deviation of each data set?

_____

c. Compare the two data sets with respect to their measures of center and their measures of variability.

_____

_____

_____

_____

d. How many times the MADs is the difference between the means, to the nearest tenth?

_____

_____

_____