

10.7 Comparing Populations

Essential Question How can you compare data sets that represent two populations?

1 ACTIVITY: Comparing Two Data Distributions

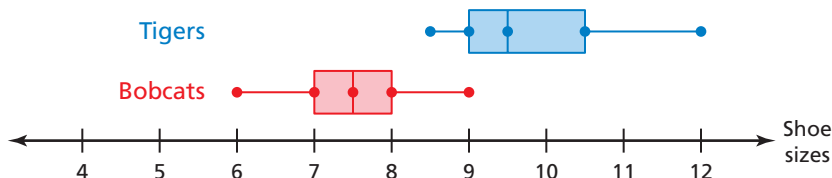
Work with a partner. You want to compare the shoe sizes of male students in two classes. You collect the data shown in the table.

Male Students in Eighth-Grade Class														
7	9	8	$7\frac{1}{2}$	$8\frac{1}{2}$	10	6	$6\frac{1}{2}$	8	8	$8\frac{1}{2}$	9	11	$7\frac{1}{2}$	$8\frac{1}{2}$
Male Students in Sixth-Grade Class														
6	$5\frac{1}{2}$	6	$6\frac{1}{2}$	$7\frac{1}{2}$	$8\frac{1}{2}$	7	$5\frac{1}{2}$	5	$5\frac{1}{2}$	$6\frac{1}{2}$	7	$4\frac{1}{2}$	6	6

- How can you display both data sets so that you can visually compare the measures of center and variation? Make the data display you chose.
- Describe the shape of each distribution.
- Complete the table.

	Mean	Median	Mode	Range	Interquartile Range (IQR)	Mean Absolute Deviation (MAD)
Male Students in Eighth-Grade Class						
Male Students in Sixth-Grade Class						

- Compare the measures of center for the data sets.
- Compare the measures of variation for the data sets. Does one data set show more variation than the other? Explain.
- Do the distributions overlap? How can you tell using the data display you chose in part (a)?
- The double box-and-whisker plot below shows the shoe sizes of the members of two girls basketball teams. Can you conclude that at least one girl from each team has the same shoe size? Can you conclude that at least one girl from the Bobcats has a larger shoe size than one of the girls from the Tigers? Explain your reasoning.



Probability and Statistics

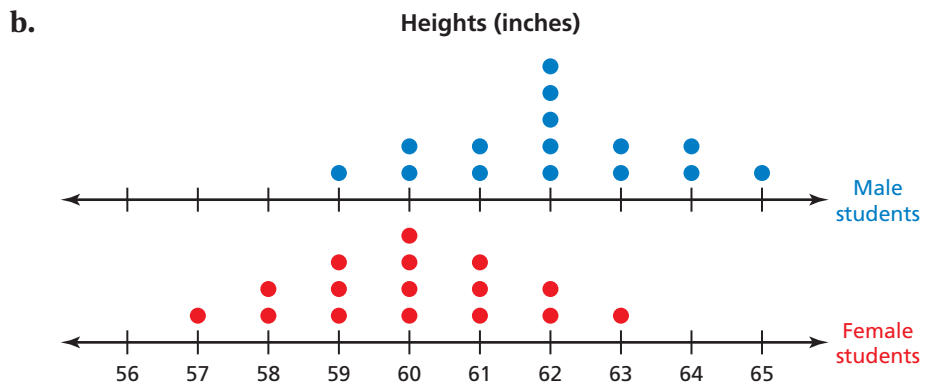
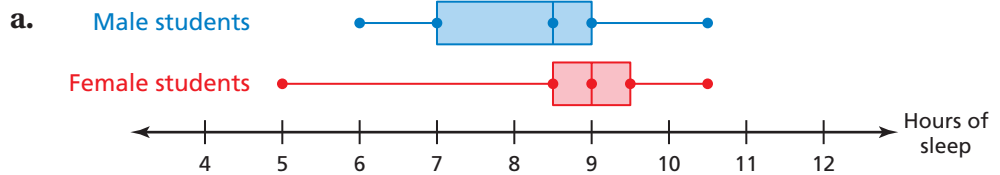
In this lesson, you will

- use measures of center and variation to compare populations.
- use random samples to compare populations.

2

ACTIVITY: Comparing Two Data Distributions

Work with a partner. Compare the shapes of the distributions. Do the two data sets overlap? Explain. If so, use measures of center and the least and the greatest values to describe the overlap between the two data sets.



Math Practice

Recognize Usefulness of Tools

How is each type of data display useful? Which do you prefer? Explain.

c. **Ages of People in Two Exercise Classes**

10:00 A.M. Class		8:00 P.M. Class
	1	8 9
	2	1 2 2 7 9 9
	3	0 3 4 5 7
9 7 3 2 2 2	4	0
7 5 4 3 1	5	
7 0 0	6	
0	7	

Key: 1 | 8 = 18

What Is Your Answer?

3. **IN YOUR OWN WORDS** How can you compare data sets that represent two populations?



Use what you learned about comparing data sets to complete Exercise 3 on page 452.

Recall that you use the mean and the mean absolute deviation (MAD) to describe symmetric distributions of data. You use the median and the interquartile range (IQR) to describe skewed distributions of data.

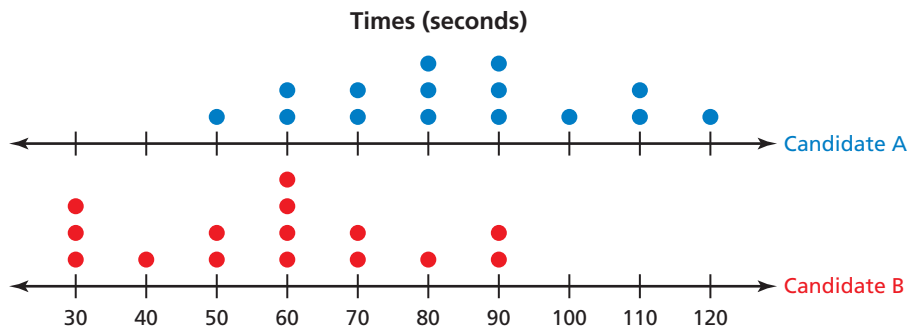
To compare two populations, use the mean and the MAD when both distributions are symmetric. Use the median and the IQR when either one or both distributions are skewed.

EXAMPLE 1 Comparing Populations

The double dot plot shows the time that each candidate in a debate spent answering each of 15 questions.

Study Tip

You can more easily see the visual overlap of dot plots that are aligned vertically.



- a. Compare the populations using measures of center and variation.

Both distributions are approximately symmetric, so use the mean and the MAD.

$$\begin{aligned} \text{Candidate A} \\ \text{Mean} &= \frac{1260}{15} = 84 \\ \text{MAD} &= \frac{244}{15} \approx 16 \end{aligned}$$

$$\begin{aligned} \text{Candidate B} \\ \text{Mean} &= \frac{870}{15} = 58 \\ \text{MAD} &= \frac{236}{15} \approx 16 \end{aligned}$$

- ∴ So, the variation in the times was about the same, but Candidate A had a greater mean time.

- b. Express the difference in the measures of center as a multiple of the measure of variation.

$$\frac{\text{mean for Candidate A} - \text{mean for Candidate B}}{\text{MAD}} = \frac{26}{16} \approx 1.6$$

- ∴ So, the difference in the means is about 1.6 times the MAD.

On Your Own

1. **WHAT IF?** Each value in the dot plot for Candidate A increases by 30 seconds. How does this affect the answers in Example 1? Explain.

Now You're Ready
Exercises 4–6

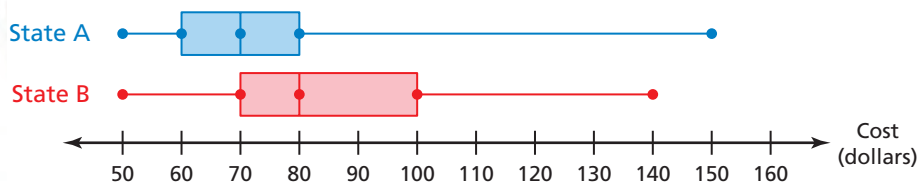
You do not need to have all the data from two populations to make comparisons. You can use random samples to make comparisons.

EXAMPLE 2 Using Random Samples to Compare Populations



You want to compare the costs of speeding tickets in two states.

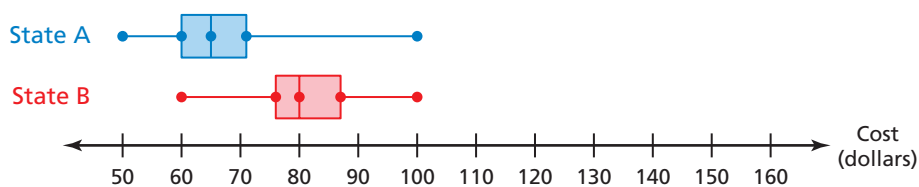
- a. The double box-and-whisker plot shows a random sample of 10 speeding tickets issued in two states. Compare the samples using measures of center and variation. Can you use this to make a valid comparison about speeding tickets in the two states? Explain.



Both distributions are skewed right, so use the median and the IQR.

- The median and the IQR for State A, 70 and 20, are less than the median and the IQR for State B, 80 and 30. However, the sample size is too small and the variability is too great to conclude that speeding tickets generally cost more in State B.

- b. The double box-and-whisker plot shows the medians of 100 random samples of 10 speeding tickets for each state. Compare the variability of the sample medians to the variability of the sample costs in part (a).



The IQR of the sample medians for each state is about 10.

- So, the sample medians vary much less than the sample costs.

- c. Make a conclusion about the costs of speeding tickets in the two states.

The sample medians show less variability. Most of the sample medians for State B are greater than the sample medians for State A.

- So, speeding tickets generally cost more in State B than in State A.

On Your Own

Now You're Ready
Exercise 8

2. **WHAT IF?** A random sample of 8 speeding tickets issued in State C has a median of \$120. Can you conclude that a speeding ticket in State C costs more than in States A and B? Explain.

Vocabulary and Concept Check

- REASONING** When comparing two populations, when should you use the mean and the MAD? the median and the IQR?
- WRITING** Two data sets have similar variabilities. Suppose the measures of center of the data sets differ by 4 times the measure of variation. Describe the visual overlap of the data.

Practice and Problem Solving

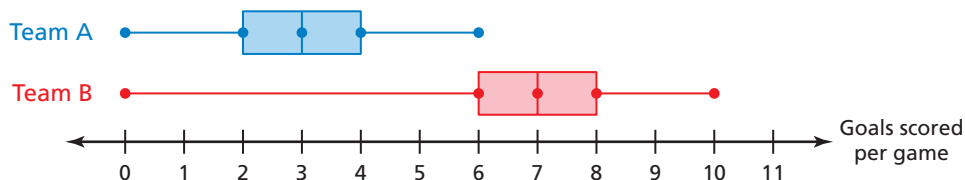
- SNAKES** The tables show the lengths of two types of snakes at an animal store.

Garter Snake Lengths (inches)					
26	30	22	15	21	24
28	32	24	25	18	35

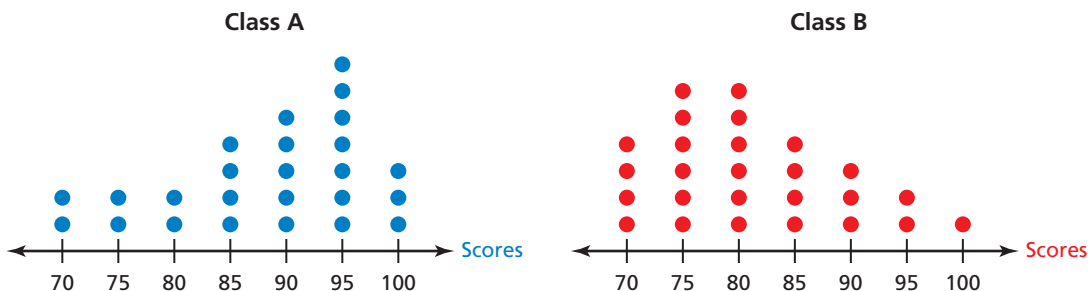
Water Snake Lengths (inches)					
34	25	24	35	40	32
41	27	37	32	21	30

- Find the mean, median, mode, range, interquartile range, and mean absolute deviation for each data set.
- Compare the data sets.

- HOCKEY** The double box-and-whisker plot shows the goals scored per game by two hockey teams during a 20-game season.



- Compare the populations using measures of center and variation.
 - Express the difference in the measures of center as a multiple of the measure of variation.
- TEST SCORES** The dot plots show the test scores for two classes taught by the same teacher.



- Compare the populations using measures of center and variation.
- Express the difference in the measures of center as a multiple of each measure of variation.

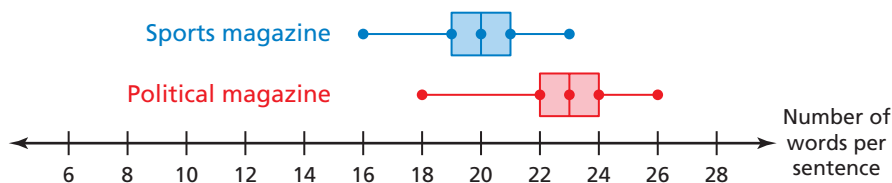
6. **ATTENDANCE** The tables show the attendances at volleyball games and basketball games at a school during the year.

Volleyball Game Attendance						
112	95	84	106	62	68	53
75	88	93	127	98	117	60
49	54	85	74	88	132	

Basketball Game Attendance						
202	190	173	155	169	188	195
176	141	152	181	198	214	179
163	186	184	207	219	228	

- a. Compare the populations using measures of center and variation.
 b. Express the difference in the measures of center as a multiple of each measure of variation.
7. **NUMBER SENSE** Compare the answers to Exercises 4(b), 5(b), and 6(b). Which value is the greatest? What does this mean?
- 2 8. **MAGAZINES** You want to compare the number of words per sentence in a sports magazine to the number of words per sentence in a political magazine.
- a. The data represent random samples of 10 sentences in each magazine. Compare the samples using measures of center and variation. Can you use this to make a valid comparison about the magazines? Explain.
 Sports magazine: 9, 21, 15, 14, 25, 26, 9, 19, 22, 30
 Political magazine: 31, 22, 17, 5, 23, 15, 10, 20, 20, 17

- b. The double box-and-whisker plot shows the means of 200 random samples of 20 sentences. Compare the variability of the sample means to the variability of the sample numbers of words in part (a).



- c. Make a conclusion about the numbers of words per sentence in the magazines.
9. **Project** You want to compare the average amounts of time students in sixth, seventh, and eighth grade spend on homework each week.
- a. Design an experiment involving random sampling that can help you make a comparison.
 b. Perform the experiment. Can you make a conclusion about which students spend the most time on homework? Explain your reasoning.



Fair Game Review what you learned in previous grades & lessons

Graph the inequality on a number line. (Section 4.1)

10. $x > 5$

11. $b \leq -3$

12. $n < -1.6$

13. $p \geq 2.5$

14. **MULTIPLE CHOICE** The number of students in the marching band increased from 100 to 125. What is the percent of increase? (Section 6.5)

(A) 20%

(B) 25%

(C) 80%

(D) 500%