LESSON 19

Transformations of Data

GOAL Determine how adding or multiplying a set of data by a constant affects the measures of central tendency and dispersion.

Activity

Transforming Data

1. Find the mean, median, mode(s), range, interquartile range, and standard deviation of the following set of data.

   −9, −8, −7, −6, −3, −3, −1, 2, 2, 2, 2, 5, 5, 6, 8, 9

2. Add 6 to each value in the set of data. Then find the mean, median, mode(s), range, interquartile range, and standard deviation of the new set of data. How do the measures of central tendency and measures of dispersion change when 6 is added to each data value?
3. Multiply each value in the set of data by 3. Then find the mean, median, mode(s), range, interquartile range, and standard deviation of the new set of data. How do the measures of central tendency and measures of dispersion change when each data value is multiplied by 3?

Adding a constant to all of the values in a set of data adds the constant to the measures of central tendency (mean, median, and mode) and the quartiles, but does not change the measures of dispersion (range, interquartile range, and standard deviation).

EXAMPLE 1 Adding a Constant to Data

The following measures of central tendency and dispersion are for the number of years that each of the nine U.S. Supreme Court justices had served as of 1999.

Mean: 13.4 years **Median:** 11 years **Mode:** There is no mode.

Range: 22 years **IQR:** 14 years **Standard deviation:** 7.44 years

In 2002, all nine justices were still serving. Find the measures of central tendency and dispersion for the number of years the justices had served as of 2002.

SOLUTION

To find the measures of central tendency for the number of years the justices had served as of 2002, add 3 to each measure of central tendency for the 1999 data.

Mean: $13.4 + 3 = 16.4$ years

Median: $11 + 3 = 14$ years

Mode: There is still no mode.

The measures of dispersion are not affected by adding 3 to each data value.

Range: 22 years **IQR:** 14 years **Standard deviation:** 7.44 years

CHECK **Example 1**

In Exercises 1 and 2, use the following information.

The list below shows the marked prices (in dollars) for shirts in a department store. The store is having a sale in which all shirts are \$5.00 off the marked price.

27.50, 27.50, 30.00, 30.00, 30.00, 34.50, 35.00, 38.00, 48.00, 49.50, 52.50

1. Find the mean, median, mode(s), range, interquartile range, and standard deviation of the sale prices.

2. Make a double box-and-whisker plot using the marked prices and sale prices.

Multiplying all of the values in a set of data by a positive constant multiplies the measures of central tendency (mean, median, and mode) and the quartiles as well as the measures of dispersion (range, interquartile range, and standard deviation) by the constant.

EXAMPLE 2 Multiplying Data by a Constant

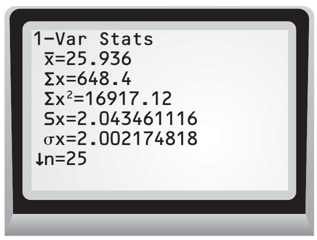
The list below shows the winning distances (in feet) in the men's Olympic long jump from 1896 to 2000. Find the mean and standard deviation of the winning distances *in meters*. (*Hint:* 1 foot = 0.3048 meter)

20.8, 23.6, 24.1, 23.6, 24.5, 24.9, 23.5, 24.4, 25.4, 25.1, 26.5, 25.7, 24.8,

25.7, 26.6, 26.5, 29.2, 27.0, 27.4, 28.0, 28.0, 28.6, 28.5, 27.9, 28.1

SOLUTION

Use a calculator to find the mean and standard deviation of the data.



$\bar{x} \approx 25.94$ ft $\qquad \sigma \approx 2.002$ ft

Because multiplying all of the values in a set of data by a constant multiplies the measures of central tendency and dispersion by the constant, you must multiply the mean and standard deviation by 0.3048 to find the measures in meters.

Mean: $\bar{x} \times 0.3048 \approx 25.94 \text{ ft} \times \dfrac{0.3048 \text{ m}}{1 \text{ ft}}$

≈ 7.907 m

Standard deviation: $\sigma \times 0.3048 \approx 2.002 \text{ ft} \times \dfrac{0.3048 \text{ m}}{1 \text{ ft}}$

≈ 0.6102 m

CHECK **Example 2**

In Exercises 3 and 4, use the Olympic long jump data in Example 2.

3. Find the mean and standard deviation of the winning distances *in yards*.

4. Find the median and range of the winning distances *in yards*.

EXERCISES

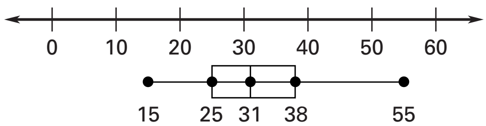
Describe the effect that the transformation will have on the mean, median, interquartile range, and standard deviation of the original data.

1. 100 is added to each data value.
2. 15 is subtracted from each data value.
3. Each data value is multiplied by 5.
4. Each data value is halved.

In Exercises 5–7, use the following information.

The box-and-whisker plot shows the original prices of board games at a toy store. Make a double box-and-whisker plot of the data for the original prices and the prices with the given sale.

5. Take $10 off the original price.
6. Take half off the original price.
7. Take 10% off the original price.



In Exercises 8–11, use the following information.

The list below shows the current hourly wages (in dollars) of 13 employees at a small fast-food restaurant.

6.00, 6.25, 6.75, 7.00, 7.25, 7.50, 7.90, 8.25, 9.75, 9.75, 10.50, 19.00, 25.00

8. Find the mean, median, range, interquartile range, and standard deviation of the current wages.
9. What will the mean, median, range, interquartile range, and standard deviation be if all of the wages are increased by $2.50?
10. Make a double box-and-whisker plot using the current wages and the wages after the $2.50 raise.
11. What will the mean, median, range, interquartile range, and standard deviation be if all of the current wages are increased by 5%?

In Exercises 12–15, use the following information.

A small health clinic employs nurses at salaries between $19,000 and $48,000. Suppose that every nurse is given a $1000 raise.

12. By how much will the mean salary increase? By how much will the median salary increase?
13. Will the $1000 raise increase the interquartile range?
14. Will the $1000 raise increase the standard deviation?
15. Suppose that the nurses each receive a 4% raise. The amount of the raise will vary from $760 to $1920, depending on the current salary. Will the 4% raise increase the interquartile range? Will it increase the standard deviation? Explain your reasoning.

In Exercises 16 and 17, explain how the transformation of the data will affect the measures of central tendency and dispersion of the original data.

16. The math test was very difficult and the teacher is thinking of "curving" the grades by adding 10 points to each student's score.

17. You weigh bags of candy in ounces. You decide to change all the measurements to pounds. (*Hint:* 16 ounces = 1 pound)

18. You have been recording the weight of your dog every day for the past two weeks. You found that the mean of the weights was 60.3 pounds with a standard deviation of 1.1 pounds. Then you noticed that the scale was set incorrectly. Instead of being set at zero when nothing is being weighed, it was set at 2 pounds. What are the actual mean and standard deviation of the weights?

In Exercises 19–22, use the data sets below.

Data set 1: 85, 89, 95, 97, 99

Data set 2: −8, −4, 2, 4, 6

19. What is the relationship between the two sets of data?

20. Find the mean of data set 2. Without calculating, predict the mean of data set 1. Test your prediction by calculating the mean.

21. Find the median and the range of each data set. How do they compare?

22. Calculate the standard deviation of each data set by hand. Which set was easier to work with?

In Exercises 23–26, use the following information.

The list below shows the temperature (in degrees Fahrenheit) of a school's pool taken at 13 different times.

74.5, 81.9, 72.5, 67.4, 75.2, 77.7, 73.4, 72.5, 78.1, 73.8, 79.8, 85.0, 63.0

23. Use a calculator to find the mean, median, interquartile range, and standard deviation of the temperatures. Round your answers to the nearest hundredth.

24. Convert all of the temperatures to degrees Celsius using the equation $C = \frac{5}{9}(F - 32)$ where C is the temperature in degrees Celsius and F is the temperature in degrees Fahrenheit. Round your answers to the nearest hundredth.

25. Use a calculator to find the mean, median, interquartile range, and standard deviation of the Celsius temperatures.

26. What effect did converting the Fahrenheit temperatures to Celsius have on the measures of central tendency and dispersion?

Describe the effect that the transformation will have on the mean and standard deviation of the original data.

27. 8 is added to each data value and then each resulting value is tripled.

28. Each data value is tripled and then 8 is added to each resulting value.