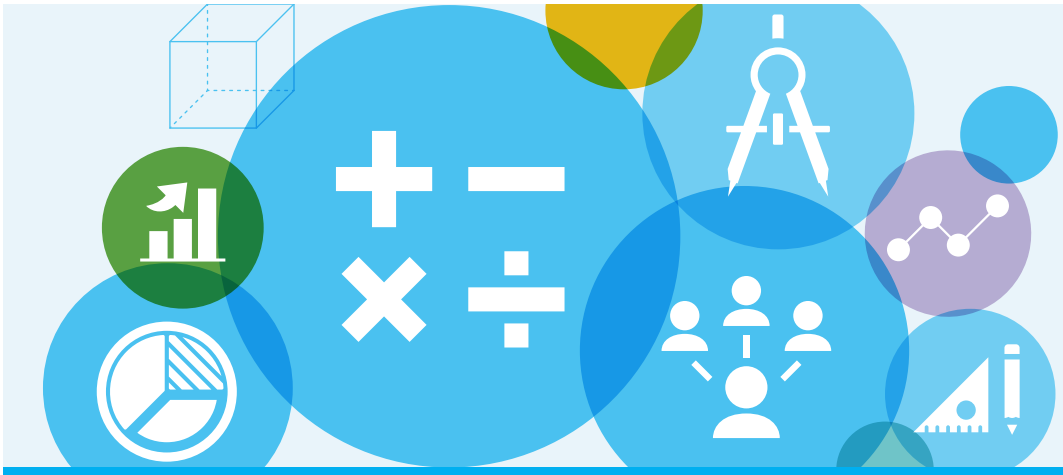


STUDENT VERSION

COMMUTING TO WORK: BOX PLOTS, CENTRAL TENDENCY, AND OUTLIERS

Activity Item

The following item is part of this activity and appears at the end of this student version.

- Item 1: Number of People Who Bike to Work in Select States

Student Learning Objectives

- I will be able to analyze and draw conclusions about the effect of an outlier on the mean, median, and shape of a data distribution.
- I will be able to assess the accuracy of measures of central tendency and variability in the presence of outliers.
- I will be able to assess the utility of outliers, specifically that they may illuminate influential factors relevant to the data set that may otherwise not have been evident.
- I will be able to make precise claims about the data set and nature of outliers.

NAME:_____ DATE:_____

Use **Item 1: Number of People Who Bike to Work in Select States** to answer the following questions and prompts.

1. Calculate the following measures of central tendency for this data set.
   a. Mean (round to the tenths place):

   b. Median:

2. Which measure of central tendency (above) is more representative of this data set? Explain.

3. Calculate the 5-number summary for this data set.
   a. Minimum:

b. Lower quartile (Q_1):

c. Median (from question 1):

d. Upper quartile (Q_3):

e. Maximum:

4. Calculate the following measures of variability for this data set.
   a. Range:

b. Interquartile range (IQR):

5. Which measure of variability (above) provides a better description of this data set? Explain.

6. Create a box plot in the space below to represent the data set, making sure to label all its components.

7. Assess the potential impact of an outlier in the data set by answering the questions and prompts below.
   a. By just looking at **Item 1: Number of People Who Bike to Work in Select States**, which state do you think might represent an outlier in this data set?

   b. For a value to be considered an outlier, it must be greater than (1.5*IQR) + Q_3 or less than Q_1 - (1.5*IQR). Use math to verify whether the state you identified above qualifies as an outlier for this data set.

   c. Describe the impact of this outlier on the mean, median, and shape of the data distribution.

   d. Why do you think the number of people who bike to work in the outlier state differs noticeably from that of the other states?

e. List additional states not included in **Item 1** that you suspect might also be high outliers in terms of the number of people who bike to work. Explain your reasoning.

f. Which of all the U.S. states would you predict has the greatest number of people who commute by bicycle. Why?

g. Based on the mathematical definition of an outlier, is it possible that a state not listed in **Item 1** is a low outlier with respect to the number of people who bike to work? Explain your reasoning.

8. Assess the potential utility of outliers by answering the questions and prompts below.
   a. How does the presence of an outlier in this particular data set help bring attention to significant factors that might influence the number of people who bike to work in a state? List one or two possible factors.

b. Do you think any of these significant factors could be identified if the outlier were not present? Explain your reasoning.

c. Thinking beyond this data set, what general role might outliers play in developing an understanding of data sets?

9. Assess the methodology of this data collection by answering the questions below.
   a. How were the data collected?

   b. Are the states included in this data set the appropriate states to compare? Explain your reasoning.

c. Thinking about how the number of bike commuters was calculated in this data set, what impact might the sheer population of a state have on the number of its residents who bike to work?

d. How else could the number of people who bike to work be calculated to account for the impact of state population?

Item 1: Number of People Who Bike to Work in Select States

State	Number of Bicycle Commuters
Alabama	1,950
Alaska	2,843
Connecticut	4,227
Delaware	1,221
Florida	53,759
Georgia	10,339
Idaho	6,860
Kansas	4,916
Maryland	9,237
Virginia	15,719

Source for data: *U.S. Census Bureau, Means of Transportation to Work. 2016-2020 American Community Survey 5-Year Estimates.*