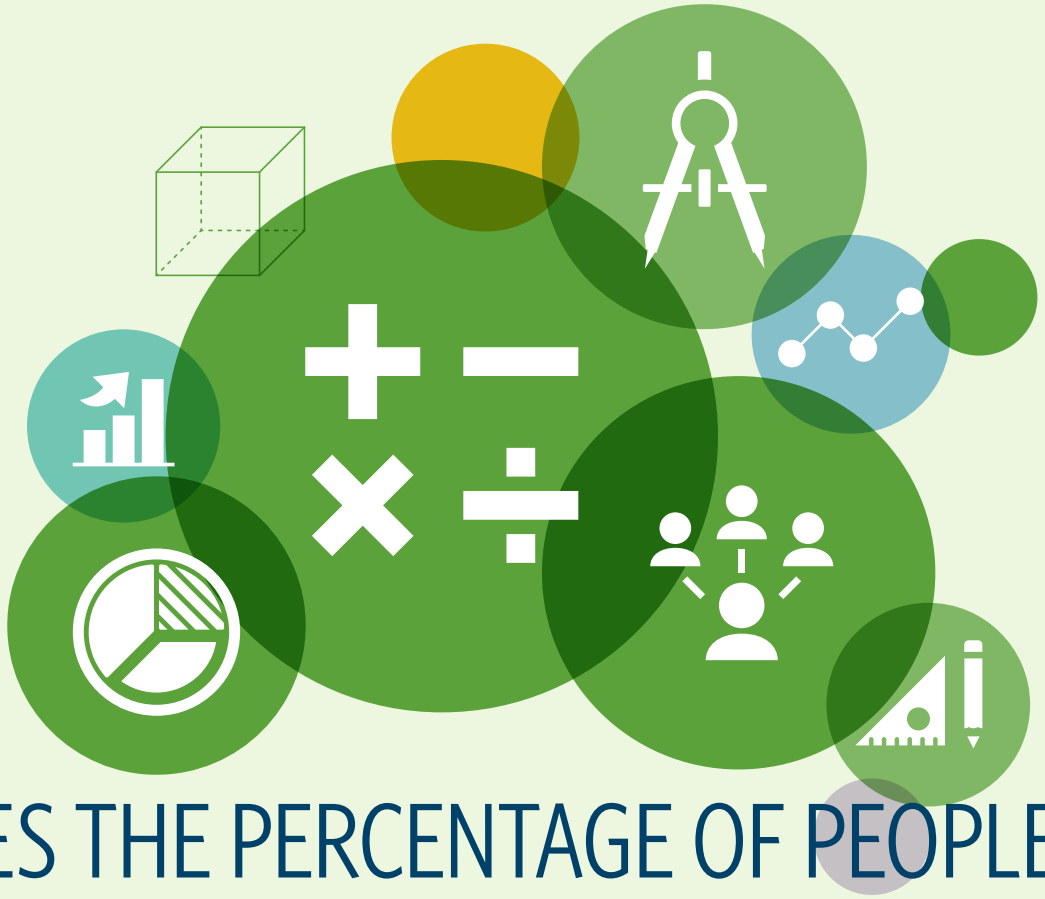


DOES THE PERCENTAGE OF PEOPLE WHO WALK TO WORK IN CITIES VARY WITH POPULATION SIZE?

TEACHER VERSION

Subject Level:
Middle School Math

Grade Level:
6–7

Approx. Time Required:
30-40 minutes

Learning Objectives:

- Students will be able to create box plots using census data.
- Students will be able to compare their box plots to investigate differences in center and variability.

United States Census Bureau

Classrooms Powered by Census Data

STATISTICS IN SCHOOLS

Activity Description

Students will create box plots to make inferences about the percentages of people who walk to work in cities of different population sizes (small, medium, and large). Students will use these findings to write a short report.

Suggested Grade Level:
6–7

Approximate Time Required:
30–40 minutes

Learning Objectives:
- Students will be able to create box plots using census data.
- Students will be able to compare their box plots to investigate differences in center and variability.

Topics:
- Box plots
- Center

Skills Taught:
- Calculating and comparing measures of center and variability
- Comparing and interpreting box plots

Materials Required

- The student version of this activity, 5 pages
- Rulers

A graphing calculator, graphing software (e.g., Microsoft Excel), or other graphing technology is optional.

Activity Item

The following item is part of this activity. The item, its data source, and instructions for viewing the source data online appear at the end of this teacher version.

- Item 1: Top 15 Walk-to-Work Cities in the United States, 2014

For more information to help you introduce your students to the U.S. Census Bureau, read "*Census Bureau 101 for Students*." This information sheet can be printed and passed out to your students as well.

Standards Addressed

See charts below. For more information, read "*Overview of Education Standards and Guidelines Addressed in Statistics in Schools Activities*."

Common Core State Standards for Mathematics

Standard	Domain	Cluster
CCSS.MATH.CONTENT.6.SP.B.4 Display numerical data in plots on a number line, including dot plots, histograms, and box plots.	6 SP – Statistics & Probability	Summarize and describe distributions.
CCSS.MATH.CONTENT.7.SP.B.4 Use measures of center and measures of variability for numerical data from random samples to draw informal comparative inferences about two populations. For example, decide whether the words in a chapter of a seventh-grade science book are generally longer than the words in a chapter of a fourth-grade science book.	7 SP – Statistics & Probability	Draw informal comparative inferences about two populations.

Common Core State Standards for Mathematical Practice

Standard
CCSS.MATH.PRACTICE.MP4. Model with mathematics. Students will use various mathematical measures to draw conclusions about census data.
CCSS.MATH.PRACTICE.MP5. Use appropriate tools strategically. Students will use graph templates or graphing technology to create their mathematical models so they can find patterns in the data.
CCSS.MATH.PRACTICE.MP6. Attend to precision. Students will work carefully with lists of percentages in decimal form to perform their calculations correctly.

National Council of Teachers of Mathematics' Principles and Standards for School Mathematics

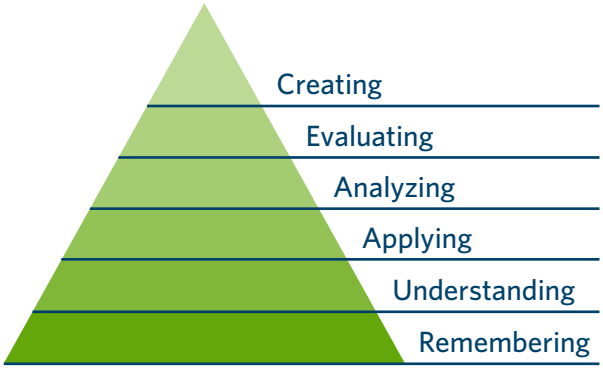
Content Standard	Students should be able to:	Expectation for Grade Band
Data Analysis and Probability	Formulate questions that can be addressed with data and collect, organize, and display relevant data to answer them.	Select, create, and use appropriate graphical representations of data, including histograms, box plots, and scatterplots.
Data Analysis and Probability	Select and use appropriate statistical methods to analyze data.	Find, use, and interpret measures of center and spread, including mean and interquartile range. Discuss and understand the correspondence between data sets and their graphical representations, especially histograms, stem-and-leaf plots, box plots, and scatterplots.
Data Analysis and Probability	Develop and evaluate inferences and predictions that are based on data.	Use observations about differences between two or more samples to make conjectures about the populations from which the samples were taken.

Guidelines for Assessment and Instruction in Statistics Education

GAISE	Level A	Level B	Level C
Formulate Questions		X	
Collect Data			
Analyze Data		X	
Interpret Results		X	

Bloom's Taxonomy

Students will ***apply*** what they know about box plots to ***analyze*** census data.



Teacher Notes

Before the Activity

Students must understand the following key terms:

- **First quartile (Q_1)** – also known as lower quartile, the value that divides a sorted data set into the smallest 25 percent of the data and the largest 75 percent
- **Third quartile (Q_3)** – also known as upper quartile, the value that divides a sorted data set into the smallest 75 percent of the data and the largest 25 percent
- **Interquartile range (IQR)** – a measure of variability in a set of numerical data that is the difference between the first and third quartiles of the data set
- **Outlier** – a data point that is well outside of the expected range of values or does not follow the overall pattern of the other data points
- **Population** – the number of people living in an area, representing the entire set of data from which samples can be taken
- **Sample** – a selection of observations from a population
- **Mean** – a measure of center in a set of numerical data, computed by adding the values in the list and then dividing by the number of values in the list
- **Median** – a measure of center in a set of numerical data, identified as the value appearing at the middle of a sorted version of the list (or the mean of the two middle values if the sorted list contains an even number of values)
- **Mode** – the number that appears most frequently in a data set
- **Box plot** – a graph that plots the median, quartiles, and extremes of the data set and uses a box to show the middle 50 percent of data

Students should have the following skills:

- Ability to create box plots
- Ability to calculate proportions

Teachers should review with students what skewness is and how it can affect the mean, specifically how that mean compares with the median.

Teachers may choose to have students use a graphing calculator or other graphing technology to create their box plots in part 1. Teachers should be aware that some software programs use different methods for calculating quartiles, which may generate different values than those provided in the sample student answers (derived from TI Teacher Software).

Teachers should survey the class to get students talking about transportation to work or school:

- How do the adults in your household get to work, if they work outside the home?
- How do you get to school — riding a bus, bicycling, skateboarding, riding in a car, walking, or another way?
- Do you and your family members each use the same transportation to get to work or school every day? If not, what factors affect your choice of transportation — weather, daily schedule, the availability of public transportation, or something else?

During the Activity

Teachers should monitor students as they work and have them work in groups for part 2.

After the Activity

Teachers should direct students to share their reports and should lead a discussion about the various ways that these data could be used in real-life situations.

Extension Ideas

- Teachers could divide the data by geographic region (instead of by city population size).
- Teachers could instead use the city bicycling or carpooling data from the data source for this activity *(COMMUTING CHARACTERISTICS BY SEX 2014: ACS 1-year Estimates)*. Teachers could also use this source data to have students compare the commuting patterns of men and women.
- To adapt this activity to GAISE Level A, teachers could have students collect and use data from the class survey about how students and adults get to school or work, instead of using the city data.
- To adapt this activity to GAISE Level C, teachers could have students discuss why the Census Bureau divided the sample into three city sizes and to propose other possible ways of dividing the data.

Student Activity

Click ***here*** to download a printable version for students.

Activity Item

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

- Item 1: Top 15 Walk-to-Work Cities in the United States, 2014

Student Learning Objectives:

- I will be able to create box plots using census data.
- I will be able to compare my box plots to investigate differences in center and variability.

In this activity, you will examine the percentages of people who walk to work in cities of different population sizes — small, medium, or large — to learn about variability.

Part 1 – Create Box Plots

Using the data in **Item 1: Top 15 Walk-to-Work Cities in the United States, 2014**, create three box plots, one for each size category, on the graph template below. Stack your box plots one above the other (it doesn't matter which is on top). Make sure that you:

- Round your data points to one decimal place, and write them in the spaces provided below the number line.
- Use a ruler to make each box plot neat.
- Title each box plot in the spaces provided next to the vertical axis, add an overall title for your graph, and label your horizontal axis.

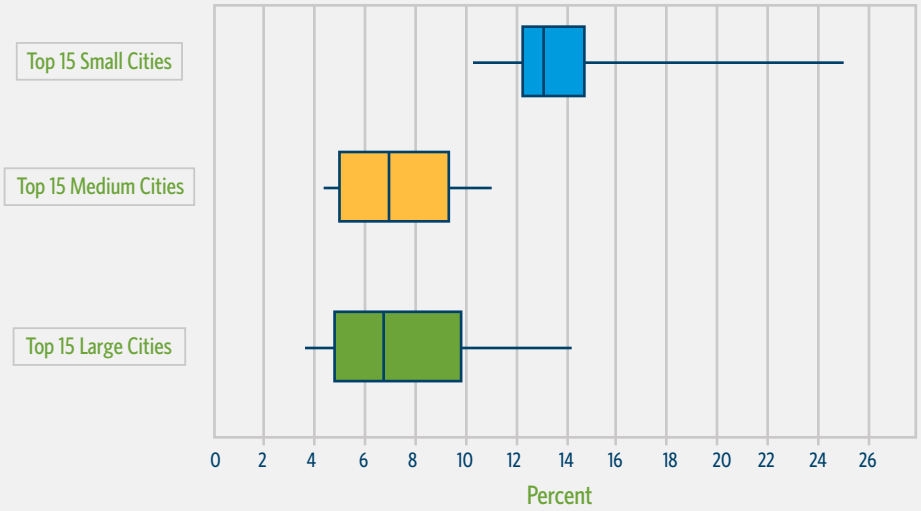
Small Cities: minimum = **10.3** Q_1 = **12.2** median = **13.1** Q_3 = **14.5** maximum = **24.9**

Medium Cities: minimum = **4.4** Q_1 = **5.0** median = **6.6** Q_3 = **9.4** maximum = **10.9**

Large Cities: minimum = **3.7** Q_1 = **4.6** median = **6.7** Q_3 = **9.9** maximum = **14.3**

Student graphs should look similar to:

Percentage of Residents Who Walk to Work in U.S. Cities of Different Population Sizes: 2014



Part 2 – Examine and Compare Your Box Plots

1. What is the unit of measurement used for the variables in all three box plots?

   The percentage of people in select cities who walk to work

2. How do the medians of your three box plots compare?

   The medians are very close for medium and large cities (6.6 percent and 6.7 percent, respectively), but the median is much higher, about twice as much, for small cities (13.1 percent).

3. How do you think the means for each box plot would compare with the medians for that same box plot? How about across box plots?

   Student answers will vary but could include: Because each data set is at least slightly skewed right, the means would probably be larger than the medians. Looking at all three box plots, I would expect the mean of small cities to be the largest and the mean of medium cities to be the smallest.

4. Calculate the interquartile range for each box plot:
   - Small Cities: **2.3** percentage points
   - Medium Cities: **4.4** percentage points
   - Large Cities: **5.3** percentage points

a. What do these numbers tell you, in terms of variability?

Student answers will vary but could include: There is less variability in the percentages of people who walk to work in the small cities compared with the percentages for the medium and large cities.

5. Do you think there are any outliers in the data for small cities? If so, list the data values and explain why you think they are outliers.

Student answers will vary but could include: Yes, I think the values for Cambridge, Massachusetts (24.9 percent); Columbia, South Carolina (21.4 percent); and Berkeley, California (19.6 percent), are outliers because they are extreme compared with the rest of the data for small cities.

6. When looking at the data in **Item 1**, do you notice any interesting features that are not captured in the box plots? If so, what are they?

Student answers will vary but could include features related to the shape of the data distribution, such as the closeness in the percentages of people who walk to work in 12 of the 15 small cities or the greater variability in the percentages of people who walk to work in the medium and large cities.

7. Write three statements that compare and contrast your box plots:

1.

2.

3.

Student responses will vary but could include:

- **Overall, greater percentages of people walk to work in the smaller cities than in the larger cities.**
- **The maximum percentage of people walking to work in large cities is higher than the maximum for medium cities.**
- **The minimums, first quartiles, medians, and third quartiles for the medium and large cities are fairly similar.**

8. How many of the small-city data points are greater than the medium-city data points? Write your answer as a proportion.

13/15

9. What does it mean that the first quartiles for medium and large cities are similar?

It means that the majority of the percentages of people who walk to work in the medium and large cities begin at a similar point: around 5 percent.

Part 3 – Write a Report About What You Learned

Write a three-paragraph report, with a title, about what you learned in this activity. In the first paragraph, summarize what you discovered about the percentages of people who walk to work in the different cities and how that information helps (or doesn't help) you answer this question: "Does the percentage of people who walk to work in cities vary with population size?" In the second paragraph, give specific examples to support your points, being sure to use good mathematical language. In the third paragraph, discuss how these data are important — specifically how city governments could use them.

Student reports will vary; a possible title could be: "Greater percentages of people in smaller cities walk to work."

Item 1: Top 15 Walk-to-Work Cities in the United States, 2014

Rank	Large Cities (>200,000 workers)	Percent of residents 16 and older who walk to work	Medium Cities (>100,000 and <200,000 workers)	Percent of residents 16 and older who walk to work	Small Cities (>20,000 and <100,000 workers)	Percent of residents 16 and older who walk to work
1	Boston, MA	14.3	Pittsburgh, PA	10.9	Cambridge, MA	24.9
2	Washington, DC	13.1	Madison, WI	10.3	Columbia, SC	21.4
3	San Francisco, CA	11.2	Newark, NJ	9.6	Berkeley, CA	19.6
4	New York, NY	9.9	Jersey City, NJ	9.4	Albany, NY	14.5
5	Seattle, WA	9.8	Honolulu, HI	9.1	Ann Arbor, MI	14.4
6	Philadelphia, PA	8.2	San Juan, PR	7.7	Jacksonville, NC	13.8
7	Minneapolis, MN	7.8	Norfolk, VA	6.7	Bloomington, IN	13.4
8	Chicago, IL	6.7	Buffalo, NY	6.6	Iowa City, IA	13.1
9	Baltimore, MD	6.6	Cincinnati, OH	6.4	New Haven, CT	13.0
10	Portland, OR	5.4	Salt Lake City, UT	5.8	Flagstaff, AZ	12.8
11	Milwaukee, WI	4.9	Cleveland, OH	5.4	Somerville, MA	12.2
12	Atlanta, GA	4.6	Richmond, VA	5.0	Syracuse, NY	12.2
13	Miami, FL	4.2	St. Louis, MO	5.0	Union, NJ	11.6
14	Denver, CO	4.1	New Orleans, LA	4.6	Evanston, IL	10.8
15	Detroit, MI	3.7	St. Paul, MN	4.4	Miami Beach, FL	10.3

Source: COMMUTING CHARACTERISTICS BY SEX 2014: ACS 1-year Estimates

Large Cities

Medium Cities

Small Cities