

Lesson 7: Mean

Daily Data Collection

Each student will record the following information on the board:

Backpack Weight & Your Age – BF/GF Age

Distribution Part 1: Center

The Mean [you learned the median earlier]

The most common measure of center is the mean. The mean is the arithmetic average. On the distribution graph, the mean is the balancing point (center of gravity) of the graph. To find the mean of a set of observations you use the following formula →

Σ is the Greek capital letter sigma. In math it means “Add ‘em up”

$$\bar{x} = \frac{1}{n} \sum x_i$$

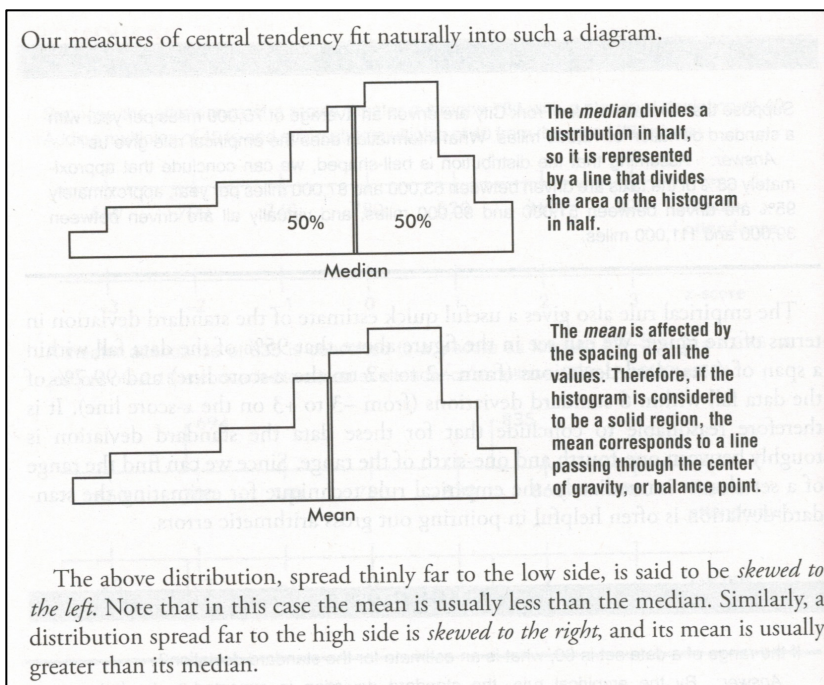
\bar{x} Represents the sample mean. It is called “x bar”

Comparing the Mean and Median

The mean and median in a symmetric distribution will be very close to each other. Although the mean is the most popular measure of center, it is not always the most appropriate measure of center. The mean is very sensitive to extreme observations (called outliers). If a distribution is skewed, it will pull the mean toward its skewness and the mean will not be an accurate measure of center. Because outliers affect the mean, we say that the mean is NOT a resistant measure of center. The median is not affected by extreme values and is a resistant measure of center.

Generally: If right skewed, then Mean > Median If left skewed, then Mean < Median

AP Question – When skewness or outliers exist, the median is a better predictor of center than the mean.



Class Data:

Find the mean and median for backpack weight and age difference [Yours – SO] and use these measures to describe if skewness exists for the data.

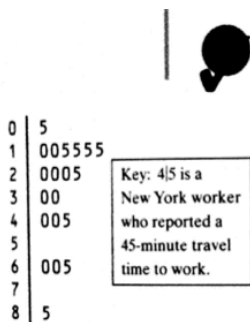
Guided Practice:

People say that it takes a long time to get to work in New York State due to the heavy traffic near big cities. What do the data say? Here are the travel times in minutes of 20 randomly chosen New York workers:

10 30 5 25 40 20 10 15 30 20 15 20 85 15 65 15 60 60 40 45

Find the mean and median for travel times and use these measures to describe if skewness exists for the data.

Choose one of the following to see if it verifies your interpretation of skewness: Dot-plot, Stem-plot, or Histogram



CHECK YOUR UNDERSTANDING

Questions 1 through 4 refer to the following setting. Here, once again, is the stemplot of travel times to work for 20 randomly selected New Yorkers. Earlier, we found that the median was 22.5 minutes.

1. Based only on the stemplot, would you expect the mean travel time to be less than, about the same as, or larger than the median? Why?
2. Use your calculator to find the mean travel time. Was your answer to Question 1 correct?
3. Interpret your result from Question 2 in context without using the words “mean” or “average.”
4. Would the mean or the median be a more appropriate summary of the center of this distribution of drive times? Justify your answer. _____