

Measures of Central Tendency

A measure of central tendency uses one number to represent the center of all the values in a data set. This number is a value around which all the values in a data set tend to cluster. The measures of central tendency are mean, median, mode. In this prompt we will also define and discuss range and quartiles.

Mean

The mean is the arithmetic average of the values in a data set. The mean can be found by adding all the values in a data set and dividing by the number of values in that data set.

Median

The median is the middle value in a data set arranged in increasing or decreasing order. If the number of data values is odd, the median is the value in the middle of the list. If the number of data values is even the median is the average of the middle two values.

Mode

The mode is the value that appears most often in a data set. In a set of values, there may be one mode, more than one mode (if two or more values appear most often), or no mode at all (if all the values are different).

Outlier

Sometimes a data set will include a value that is noticeably larger or smaller than the rest of the data values. An extreme value like this is called an outlier. A single outlier can have an effect on the mean, but will not affect the median or the mode.

Range

Range is the difference between the highest value and the lowest value in a data set. Range is a measure of variability. It shows how far apart the values of a data set are spread.

Quartiles

In statistics, quartiles are those values that split a set of number into four equal groups. You have already encountered the median, which divides a set of numbers in two. The lower quartile (Q_1) is the median of the lower half. The upper quartile (Q_3) is the median of the upper half.

Interpreting data

Sometimes the conclusions drawn from data can be misleading. You need to carefully analyze and interpret the reliability and reasonableness of data.

Example:

A company recorded and analyzed the wait times that it took for its service representatives to answer customers' phone calls. The following table shows a sample of wait times.

Call	1	2	3	4	5	6	7	8	9	10
Waiting Period (min)	2.00	1.50	2.50	1.50	2.25	9.50	1.75	0.75	1.00	2.45

Find the mean of the wait times. Why might the mean wait time be misleading.

First, find the mean of the wait times. Recall that mean is the average of all the times so:

$$\frac{2.00 + 1.50 + 2.50 + 1.50 + 2.25 + 9.50 + 1.75 + 0.75 + 1.00 + 2.50}{10} = 2.52$$

Now, think about how the mean wait time might be misleading. Only one wait time was more than 2.52 minutes. The wait time of 9.50 is an outlier and greatly increased the mean. To better describe the typical wait time, it might be better to calculate the mean without the outlier or use the median of the data set.

Solution:

$$\frac{2.00 + 1.50 + 2.50 + 1.50 + 2.25 + 1.75 + 0.75 + 1.00 + 2.50}{9} = 1.75$$

or

0.75, 1.00, 1.50, 1.50, 1.75, 2.00, 2.25, 2.50, 2.50, 9.50

↑
Median

↓

$$\frac{1.75 + 2.00}{2} = 1.88$$