Measures of Center

by Sophia

1. Measures of Center

There are multiple measures of center:

- The mean
- The median
- The mode

Each of these measurements was previously covered in their own tutorials. However, you may be wondering which measure of center you should use for a given situation.

1a. Mean

The **mean** is the default measurement that you should use if there's no compelling reason to use anything else.

The mean is the best measurement to use if possible because it's the most versatile measure of center and therefore, the most appropriate one in the vast majority of cases. However, there are certain situations in which the mean is not an appropriate gauge for center. In those cases, you should use the median. You will rarely use the mode.

TERM TO KNOW

Mean

The average number in a quantitative data set; the sum of all the values, divided by the number of values.

1b. Median

Sometimes the mean is a poor representation of where the center really is, so the **median** becomes a better measure of center.

Consider this table of one company's employees and their salaries:

Title	Number of Employees	Salary
Boss	1	$200,000
Manager	3	$55,000
Shift Worker	8	$42,000

The mean of this set of data is about $58,000, but how many of the employees actually make more than $58,000 and how many make less than $58,000?

Eleven of the 12 people make less than $58,000. Only one employee makes more than that, and that employee makes substantially more. The boss's $200,000 salary is an outlier. Therefore, the mean doesn't make very much sense as a measurement of the middle.

In this case, a better measure of center would be the median. If you took all the salaries and wrote them out from least to greatest, the median (the one in the middle) would be $42,000. That more accurately describes what a typical worker makes.

~~$42,000~~
~~$42,000~~
~~$42,000~~
~~$42,000~~
~~$42,000~~
$42,000
$42,000
~~$42,000~~
~~$55,000~~
~~$55,000~~
~~$55,000~~
~~$200,000~~

⭐ **BIG IDEA**

In the presence of outliers, which are very few high or very few low values, the mean won't give an accurate representation of center. Use the median in cases like these.
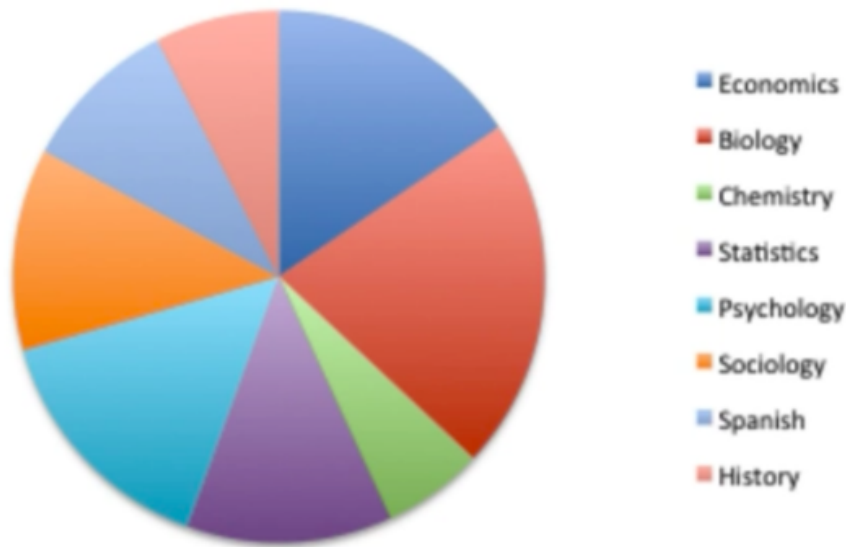
📄 **TERM TO KNOW**

Median

The value that is in the "middle" of a data set when the set is arranged from least to greatest.

1c. Mode

When should you use mode? The **mode** isn't used very often. It's used mainly for qualitative data sets, to determine the category that has the most values in it. Consider the graph below:

In this case, the mode is biology, the red section.

The mode is also used to describe the peak of a distribution, such as in a histogram.

📄 **TERM TO KNOW**

Mode
The most frequently appearing number in a set of quantitative data or most frequently occurring category in a set of qualitative data.

☑ **SUMMARY**

The mean is our default measure of center. It's the preferred one and the most versatile. However, sometimes if we have outliers or a few values that can skew the mean towards them--either on the high side or the low side--the mean won't accurately represent center anymore. In those cases, the median should be used instead. Typically we reserve the mode for qualitative distributions.

Good luck!

Source: THIS TUTORIAL WAS AUTHORED BY JONATHAN OSTERS FOR SOPHIA LEARNING. PLEASE SEE OUR **TERMS OF USE**.

Mean

The average number in a quantitative data set; the sum of all the values, divided by the number of values.

Median

The value that is in the "middle" of a data set when the set is arranged from least to greatest.

Mode

The most frequently appearing number in a set of quantitative data or most frequently occurring category in a set of qualitative data.