

Why is the mode important in statistics?

Authored by
stats writer

June 30, 2024

RECOMMENDED CITATION

stats writer (2024). *Why is the mode important in statistics?*. PSYCHOLOGICAL SCALES.
Retrieved from <https://scales.arabpsychology.com/?p=161426>

The mode is an important statistical measure that represents the most frequently occurring value in a data set. It provides valuable information about the central tendency of the data, allowing for a better understanding of the overall distribution. The mode is particularly useful in data sets with outliers or skewed distributions, as it is not influenced by extreme values. Additionally, the mode can be used to determine the most common category or group in categorical data. Overall, the mode plays a critical role in summarizing and interpreting data, making it an essential tool in statistical analysis.

Why is the Mode Important in Statistics?

The mode represents the value that occurs most often in a dataset.

A dataset can have no mode (if no value repeats), one mode, or multiple modes.

For example, the mode in the following dataset is 19:

Dataset: 3, 4, 11, 15, 19, 19, 19, 22, 22, 23, 23, 26

This is the value that occurs most often.

In statistics, the mode is important for the following reasons:

Reason 1: It lets us know which value(s) in a dataset is the most common.

Reason 2: It's useful for finding the most frequently

occurring value in categorical data when the mean and median can't be calculated.

Reason 3: It gives us an idea of where the "center" of a dataset is located, although the median and mean are more commonly used (as we'll see later in this article).

The following examples illustrate each of these reasons in practice.

Reason 1: The Mode Tells Us Which Value is Most Common

Suppose we have a dataset with 100,000 rows that contain the selling price of houses throughout the United States:

House	Selling Price
1	\$224,000
2	\$340,000
3	\$395,000
4	\$239,000
5	\$560,000
...	
100,000	\$400,000

Suppose we use some statistical software (like , , , etc.) to calculate the mode of this dataset and find that there

are three modes:

\$280,000 \$300,000 \$305,000

This immediately gives us an idea of which house prices occur most frequently in the dataset.

Calculating the mode is also much quicker than staring at thousands of rows of data and attempting to identify which house prices occur most often.

Reason 2: The Mode Finds the Most Common Value in Categorical Data

Suppose we have a dataset with 1,000 rows that tells us the color of the car owned by individuals in a certain neighborhood:

Car	Color
1	Red
2	Yellow
3	Black
4	Black
5	Red
...	
1,000	Blue

The variable "color" is a , which means the values fall

into categories ("red", "yellow", "black", etc.) so we can't calculate a quantitative value like the mean or median.

However, we can calculate the mode because this simply represents the most commonly occurring value in the dataset.

For example, we might use some statistical software to find that the mode of this dataset is "black" - which tells us that the most frequently occurring car color in this dataset is black.

Reason 3: The Mode Gives Us an Idea of Where the Center of a Dataset is Located

The mode is also considered a , which means it can give us an idea of where the "center" of the dataset is located.

For example, suppose we have the following dataset that shows the exam scores of 20 different students in a class:

Student	Score
1	68
2	72
3	72
4	74
5	77
6	79
7	80
8	81
9	82
10	82
11	82
12	84
13	85
14	87
15	88
16	90
17	91
18	92
19	95
20	98

The mode turns out to be 82 - this is the most common exam score. This also turns out to be a good indication of the where the "center" exam score value is located in this dataset.

However, suppose we instead had the following dataset of exam scores:

Student	Score
1	68
2	72
3	72
4	72
5	77
6	79
7	80
8	81
9	82
10	82
11	83
12	84
13	85
14	87
15	88
16	90
17	91
18	92
19	95
20	98

In this dataset, 72 is the mode exam score. However, this turns out to be a poor indication of where the "center" exam score is located.

The mean exam score is 82.9 and the median exam score is 82.5, which both give us a better idea of where the "center" value is located compared to the mode.

Summary

Here's a quick summary of the main points made in this article:

The mode represents the value(s) that occurs most often in a dataset. The mode tells us the most common value in categorical data when the mean and median can't be used. The mode gives us an idea of where the "center" of a dataset is located, but it can be misleading compared to the mean or median.

Additional Resources

The following tutorials provide additional information about the mean, median, and mode in statistics: