Advantages & Disadvantages of Using Median in Statistics

Authored by

stats writer

November 18, 2025

The median is one of the three primary measures of central tendency used in statistics, representing the exact middle value of a quantitative dataset. Unlike the arithmetic mean, which is highly sensitive to the magnitude of every observation, the median is a measure of position.

Understanding the appropriate application of this metric is crucial for accurate data analysis. While the median offers significant advantages in certain analytical scenarios, particularly when data distributions are uneven, it also carries inherent limitations that analysts must consider.

Defining the Median: A Measure of Central Tendency

The median is formally defined as the 50th percentile--the point that separates the upper half of the data from the lower half. It provides a robust alternative to the arithmetic mean, particularly when describing typical values within complex or non-normal datasets.

To calculate the median, every observation within the dataset must first be arranged sequentially, either in ascending or descending order. This process of ordering is fundamental because the median relies entirely on the rank of the data points, not their numerical magnitude.

The Calculation of the Median

The methodology for calculating the median depends on the total number of observations, commonly denoted as N. The process ensures that the result is truly the central value of the distribution.

For a dataset containing an odd number of observations, the median is simply the value located at the position determined by the formula (N + 1) / 2. This position yields a single, identifiable data point that splits the dataset perfectly in half.

Conversely, if the dataset has an even number of observations, no single observation occupies the center. In this case, the median is calculated by taking the average of the two middle values. These two central values are averaged to produce a representative midpoint for the distribution.

Core Advantage 1: Robustness Against Outliers

One of the most powerful reasons to employ the median is its remarkable resistance to extreme values, known as outliers. An outlier is a data point that differs significantly from other observations, often skewing standard calculations like the mean.

Since the calculation of the median depends solely on the positional ranking of observations, the actual numerical magnitude of the values at the extremes--the smallest and largest--has no effect on the result. For instance, whether the lowest score in a test dataset is 10 or 50, the median remains unchanged, provided the middle value is preserved.

This property makes the median a highly desirable metric for calculating typical values in fields where extreme data points are common, such as household wealth, real estate pricing, or environmental measurements, thereby providing a much more stable and representative measure of the central tendency.

Core Advantage 2: Superiority in Skewed Distributions

A related advantage is the median's effectiveness when analyzing skewed distributions. A distribution is considered skewed if it is asymmetrical, meaning the data points cluster more heavily on one side of the center. Income data, for example, is typically right-skewed because a few extremely high earners pull the distribution's average far above the typical salary.

When a dataset is highly skewed, the mean is pulled toward the long tail--the direction of the skewness--and thus ceases to be a good representation of the "typical" observation. In contrast, the median's positional nature ensures it remains closer to the true center of the bulk of the data.

Therefore, for distributions that are not normally distributed, the median is statistically preferred as the measure of central tendency because it better reflects the true center experienced by the majority of the population captured in the dataset.

Primary Disadvantage 1: Loss of Information (Non-Exhaustive Calculation)

Despite its robustness, the median suffers from a key drawback: it does not incorporate all available information from the dataset during its calculation. By focusing purely on position, the median ignores the actual values and distances of observations beyond the central point.

In statistics, a general principle holds that using all data points available leads to more efficient and informative summary metrics. The fact that the median disregards the numerical value of outliers, while beneficial for stability, means that potentially valuable information about the data's range and spread is overlooked.

This lack of comprehensiveness can be problematic if an analyst needs a measure that captures the overall magnitude of the entire dataset, rather than just its center point. The mean, although susceptible to outliers, utilizes every single value in its computation, providing a calculation that is fully exhaustive of the dataset's numerical information.

Primary Disadvantage 2: Inability to Calculate Aggregate Sums

A significant practical limitation of the median is its inability to be used to easily derive the sum of all observations in the dataset. This constraint arises directly from the median's nature as a positional measure.

If we know the mean (average value) and the total sample size (N), we can quickly calculate the total sum of all values by multiplying the mean by N. This utility is invaluable in business and financial analysis where calculating aggregate totals (e.g., total sales, total expenditures) is often the primary goal.

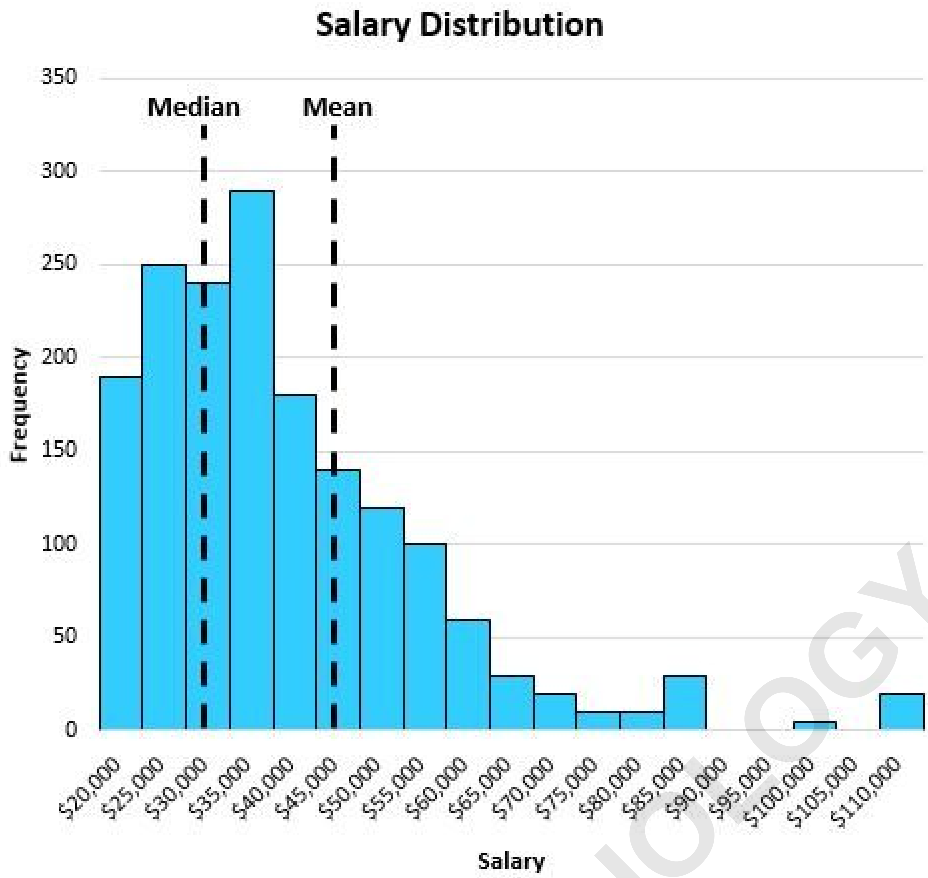
Conversely, knowing the median and the sample size N provides no direct mathematical method for determining the total sum of the data. This limitation means that if the research objective requires aggregation, the median is often inadequate as the sole measure of central tendency.

Example 1: The Advantages of Using the Median

Case Study 1: Analyzing Right-Skewed Salary Data

Consider a hypothetical company's salary distribution, which is typical of real-world financial data in that it is right skewed. This skewness means a majority of employees earn moderate salaries, while a small handful of executives or specialists earn extremely high incomes.

When we calculate both the mean and the median salary for this distribution, the differences become stark. Suppose the calculated mean is approximately $47,000 per year, while the calculated median is only $32,000 per year.

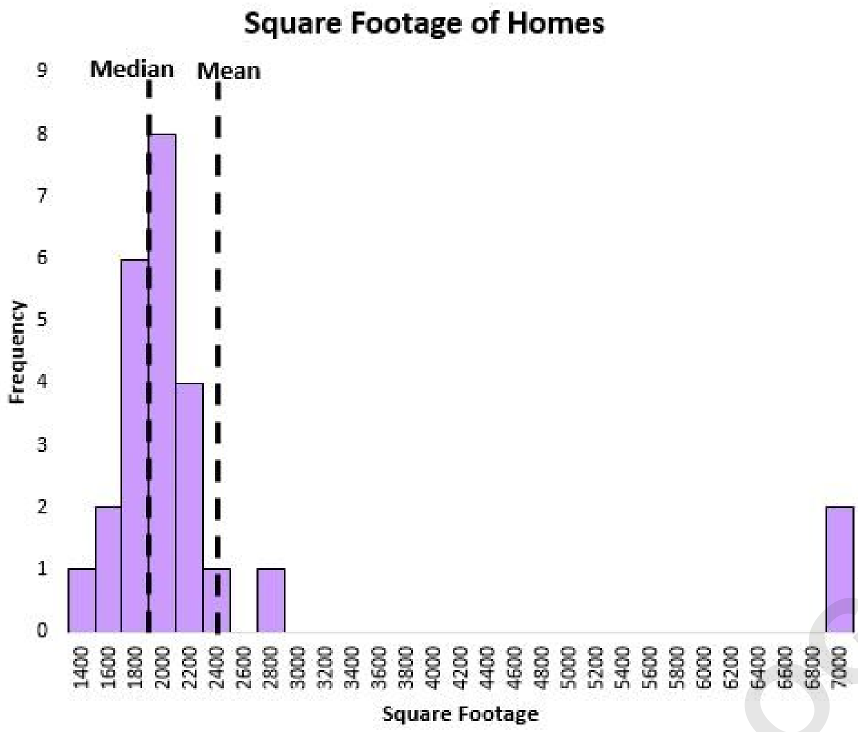


In this scenario, $47,000 does not accurately represent the typical employee salary; it is inflated by the high values on the right tail. The median value of $32,000, however, indicates that half of the employees earn less than this amount and half earn more, making it a much more representative and truthful measure of the typical individual's earning within that company.

Case Study 2: Understanding Robustness in Housing Square Footage

Let us examine another distribution containing data about the square footage of houses on a specific street. Most houses are standard sizes, but the street also includes a few exceptionally large, luxurious mansions--these are the outliers.

If we attempt to determine the typical house size using the mean, the few extremely large houses will dramatically inflate the average square footage, causing the mean to take on a value much larger than the size of the majority of homes.



In contrast, the median square footage will remain stable, providing a measure that truly represents the center of the residential market on that street, unaffected by the presence of a few extreme data points. This clearly illustrates the median's utility as a robust measure when data variability is high or skewness is present.

Example 2: The Disadvantages of Using the Median

Demonstrating Information Loss with Exam Scores

To recall the first major disadvantage: the median does not utilize all observations. This means that changes to extreme values can occur without any corresponding change in the calculated center.

Consider a dataset showing exam scores for students in a class, already arranged in ascending order:

Scores: 68, 70, 71, 75, 78, 82, **83**, 83, 85, 90, 91, 91, 92

With 13 scores, the median is the 7th score, which is 83. This value represents the middle performance level.

Now, consider an alternative scenario where the same class structure exists, but the lowest three exam scores are significantly lower due to non-attendance or other issues:

Scores: 22, 35, 38, 75, 78, 82, **83**, 83, 85, 90, 91, 91, 92

Even with these drastically lower scores (22, 35, 38 compared to 68, 70, 71), the median exam score remains exactly 83. The median ignores the large difference in the low scores because it is merely a measure of position, not magnitude. While the mean would drop significantly in the second scenario, the median provides an incomplete picture of the overall performance drop.

The Limitation in Aggregate Sales Summation

The second disadvantage highlights the difficulty in using the median for practical summation tasks. Suppose we analyze the total sales made by 11 different employees during a fiscal quarter:

Sales: 12, 12, 15, 19, 22, **24**, 28, 30, 32, 35, 38

We know that the median sales value is 24, and the total number of employees (N) is 11. Although we have these two figures, we cannot use them to deduce the total cumulative sales achieved by all employees combined.

If, by contrast, we had been provided with the information that the mean sales value was 24, we could immediately calculate the total aggregate sales: Mean * N = 24 * 11 = 264. This functional disparity is why the mean is often the preferred metric in accounting and finance where summary totals are critical for reporting and budget planning.

Conclusion: When to Choose the Median vs. the Mean

Both the mean and the median are powerful metrics, and the choice between them hinges entirely on the distribution characteristics of the data and the specific analytical objective.

If the dataset is relatively symmetrical and lacks extreme outliers, the mean is generally preferred because it uses all data points and allows for powerful mathematical derivations, such as summation.

However, when dealing with highly skewed distributions (like income or house prices) or when the analyst needs a highly robust measure that is immune to the influence of extreme values, the median is undoubtedly the superior and more representative measure of central tendency.

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