

How to Calculate Skewness in Excel: A Step-by-Step Guide

Authored by
stats writer

March 6, 2026

RECOMMENDED CITATION

stats writer (2026). *How to Calculate Skewness in Excel: A Step-by-Step Guide*.
PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=134287>

Understanding the Fundamentals of Skewness in Data Analysis

In the expansive field of **statistics**, the concept of **skewness** serves as a critical diagnostic tool for understanding the underlying structure of a dataset. It is a measure that quantifies the degree of **asymmetry** observed in a **probability distribution**, deviating from the perfectly balanced bell curve known as a **normal distribution**. By analyzing skewness, data scientists and analysts can determine whether data points are concentrated more heavily on one side of the **mean** than the other, providing insights that a simple average or median might fail to capture.

When working within **Microsoft Excel**, calculating this metric becomes an efficient process, allowing users to move beyond basic descriptive statistics into deeper analytical territory. The importance of identifying skewness lies in its ability to inform the selection of appropriate statistical models and tests. For instance, many parametric tests assume a **normal distribution**, and a high degree of skewness might suggest that non-parametric alternatives are more suitable for the analysis at hand. This foundational understanding is essential for anyone looking to perform rigorous data validation and interpretation.

Furthermore, skewness provides a visual and mathematical representation of the "tails" of a distribution. These tails indicate the presence of **outliers** or extreme values that can significantly pull the mean away from the center of the data. In financial modeling, risk assessment, or quality control, recognizing these imbalances is paramount. **Skewness** allows an analyst to predict the likelihood of extreme events, making it a cornerstone of modern **data analysis** workflows.

The Characteristics and Implications of Negative Skewness

A distribution is described as having a **negative skew**, or being **left-skewed**, when the left tail of the distribution is longer or fatter than the right tail. In such cases, the mass of the distribution is concentrated on the right side of the figure, suggesting that the majority of data points are relatively high, with a few exceptionally low values pulling the **mean** downward. Mathematically, in a negatively skewed distribution, the mean is typically less than the **median**, which in turn may be less than the **mode**.

Real-world examples of negative skewness are frequently found in phenomena where there is a natural upper limit or a high concentration of high-performing subjects. For instance, the age of retirement in many developed nations often exhibits a negative skew; most individuals retire in their 60s or 70s, but a small number of people retire much earlier due to various circumstances, creating a long tail toward the younger ages. Similarly, in an academic setting, a very easy exam might result in a negatively skewed distribution of scores, as most students score near the maximum, while only a few students score significantly lower.

Understanding negative skewness is vital for **decision-making** because it highlights the risks

associated with the lower end of the spectrum. Analysts must be aware that while the bulk of the data appears positive, the presence of a negative tail indicates a potential for infrequent but significant "downside" events. In **finance**, a negatively skewed return distribution implies that an investor might experience frequent small gains but occasional large losses, which is a critical factor in **risk management**.

Exploring Positive Skewness and the Right-Hand Tail

Conversely, **positive skewness**, also known as a **right-skewed** distribution, occurs when the tail on the right side of the distribution is longer or thicker than the left side. This indicates that the data values are clustered at the lower end of the scale, with several extreme **outliers** extending toward much higher values. In a positively skewed dataset, the **mean** is generally greater than the **median**, as the high-value outliers disproportionately influence the average.

Common examples of positive skewness include household income distributions and real estate prices. In most economies, the majority of the population earns a moderate income, but a small percentage of ultra-high-income earners creates a long tail that stretches to the right. Similarly, in the housing market, most homes may be priced within a certain range, but luxury estates priced in the tens of millions create a positively skewed distribution. This asymmetry is a fundamental characteristic of many economic and social **variables**.

For researchers and business analysts, recognizing **skewness** that leans to the right is essential for setting realistic expectations. When the mean is inflated by a few high-value outliers, relying solely on the average can be misleading. In these scenarios, the **median** often serves as a more accurate representation of the "typical" value within the dataset. By identifying positive skewness early in the **data analysis** process, one can avoid making erroneous conclusions based on skewed averages.

Defining the SKEW Function in Microsoft Excel

To facilitate the calculation of this complex metric, **Microsoft Excel** provides a built-in function known as **SKEW**. This function is designed to return the **skewness** of a distribution based on a sample of data. The syntax for the function is straightforward: **=SKEW(number1, , ...)**, where the arguments can be individual numbers, cell references, or ranges of data. By automating the calculation, Excel allows users to focus on the interpretation of the results rather than the intricacies of the underlying arithmetic.

The **SKEW** function specifically measures the degree to which a sample distribution departs from a symmetrical distribution. It is important to note that Excel uses the adjusted **Fisher-Pearson standardized moment coefficient** to arrive at this value. This specific method is preferred for sample data as it provides an unbiased estimate of the population skewness. For users dealing

with large datasets, the ability to simply highlight an **array** of values and receive an instantaneous result is a significant productivity advantage.

When applying the **SKEW** function, it is best practice to ensure that the data range is clean and free of non-numeric characters that might interfere with the calculation. While Excel is robust enough to ignore empty cells or logical values in many cases, maintaining a well-structured **spreadsheet** ensures the accuracy of the statistical output. As we will see in the subsequent sections, the numerical value returned by this function provides a clear indication of the direction and intensity of the data's asymmetry.

Mathematical Foundations: Decoding the Skewness Formula

While **Excel** handles the heavy lifting, understanding the mathematical formula behind the **SKEW** function can provide deeper insight into what the result actually represents. The formula used by Excel for sample skewness is expressed as:

$$\text{Skewness} = \frac{\sum (x_i - \bar{x})^3}{n \cdot s^3}$$

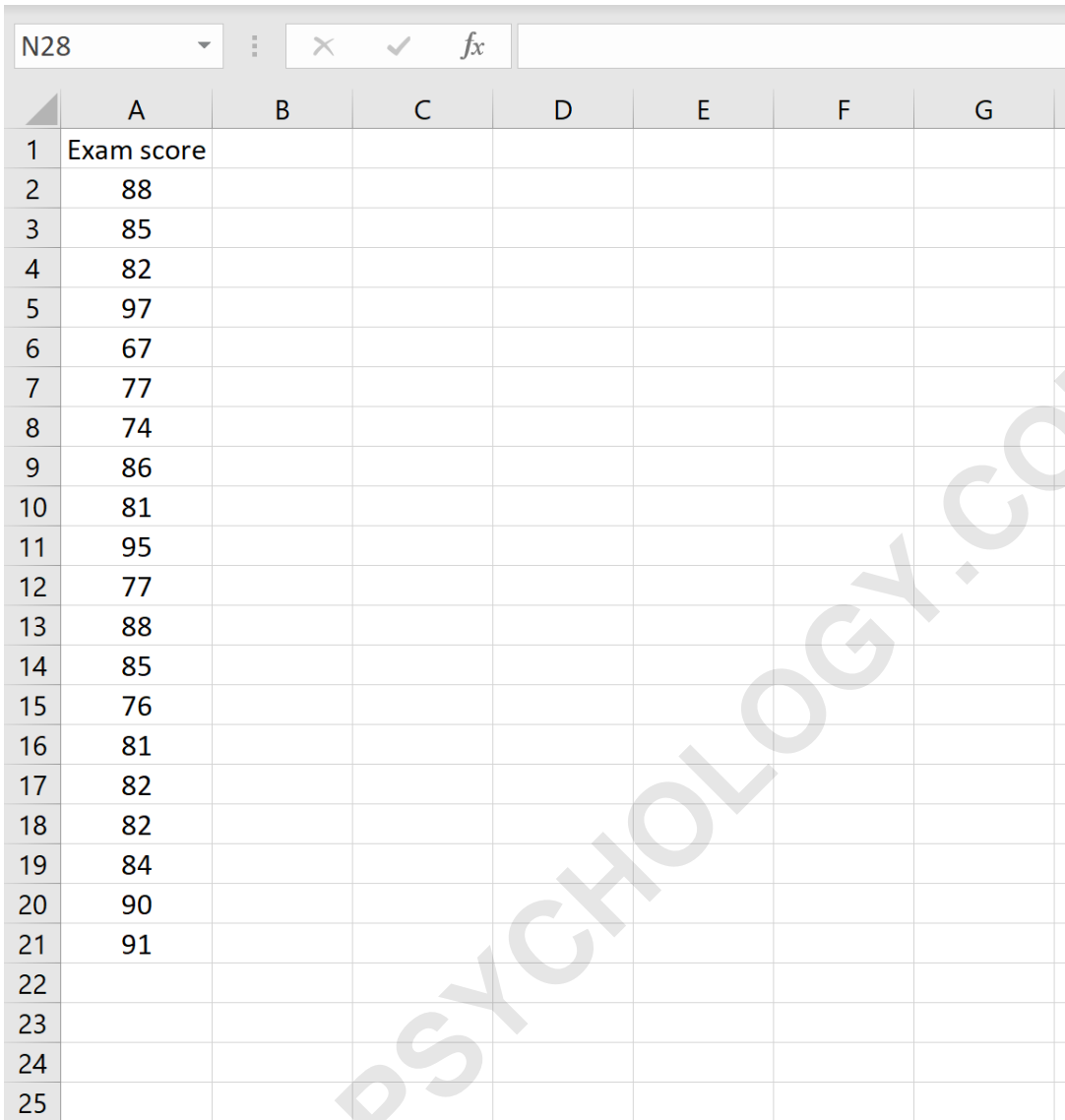
In this equation, **n** represents the **sample size**, or the total number of data points in the set. The symbol Σ (sigma) denotes the summation of the values that follow it. The term **x_i** refers to each individual value within the dataset, while **\bar{x}** (with a bar, though often written as mean) represents the **arithmetic mean** of the sample. Finally, **s** represents the **standard deviation** of the sample.

The formula essentially calculates the third **moment** of the distribution. By cubing the difference between each data point and the mean, the formula preserves the sign of the difference (unlike the squaring used in variance calculations). This means that values far above the mean result in large positive numbers, while values far below the mean result in large negative numbers. The summation of these cubed values, when normalized by the standard deviation and adjusted for the sample size, yields the final **skewness** value.

This complexity underscores why the **SKEW** function in Excel is so valuable. Manually calculating the cubed deviations for every point in a dataset of hundreds or thousands of entries would be highly susceptible to human error. By leveraging Excel's **algorithm**, analysts can ensure precision and consistency in their statistical reporting, regardless of the size of the **dataset**.

A Step-by-Step Tutorial: Calculating Skewness in Excel

To begin the process of calculating **skewness**, you must first have your data organized within an Excel worksheet. Typically, data is arranged in a single column or row for clarity. Once your data is entered, select an empty cell where you wish the final skewness value to appear. This cell will act as the output for your statistical analysis.



	A	B	C	D	E	F	G
1	Exam score						
2	88						
3	85						
4	82						
5	97						
6	67						
7	77						
8	74						
9	86						
10	81						
11	95						
12	77						
13	88						
14	85						
15	76						
16	81						
17	82						
18	82						
19	84						
20	90						
21	91						
22							
23							
24							
25							

Next, initiate the function by typing **=SKEW(** into the selected cell. You can then use your mouse to highlight the **array** of values you wish to analyze. Alternatively, you can manually type the range, such as **A2:A21**, if you are familiar with the cell coordinates. After closing the parenthesis and pressing the Enter key, Excel will process the values using the formula discussed previously and display the skewness coefficient.

	A	B	C	D	E	F
1	Exam score		Skewness	Formula used		
2	88		-0.1849	=SKEW(A2:A21)		
3	85					
4	82					
5	97					
6	67					
7	77					
8	74					
9	86					
10	81					
11	95					
12	77					
13	88					
14	85					
15	76					
16	81					
17	82					
18	82					
19	84					
20	90					
21	91					
22						
23						
24						

In the provided example, the function **=SKEW(A2:A21)** yields a result of **-0.1849**. This negative value confirms that the distribution is slightly **left-skewed**, meaning the tail extends toward the lower values on the left side of the **mean**. While a value of -0.1849 indicates a relatively mild skew, it nonetheless provides a more nuanced understanding of the data's shape than a simple average would allow.

Troubleshooting and Technical Requirements for SKEW

While the **SKEW** function is a powerful tool, it does have specific technical requirements that must be met to avoid errors. One of the most common issues users encounter is the **#DIV/0!** error. This error typically occurs under two specific conditions related to the mathematical constraints of the

formula. First, if the **sample size** (n) is fewer than three data points, the formula cannot be completed because the denominator in the adjustment factor would become zero.

The second scenario that triggers a **#DIV/0!** error is when the **standard deviation** of the sample is zero. A standard deviation of zero implies that all data points in the set are identical. In such a case, there is no variation or spread to measure, and the **skewness** cannot be calculated because the formula involves dividing by the standard deviation. Ensuring your dataset contains at least three distinct values is the best way to avoid these technical pitfalls.

To summarize these technical constraints:

The dataset must contain at least **three data points** to satisfy the degrees of freedom required by the formula.

There must be some **variability** in the data; if all values are the same, the standard deviation is zero, making the calculation impossible.

Ensure that the data range does not contain **text strings** or **error values**, as these can cause the function to return incorrect results or additional errors.

Utilizing External Calculators for Validation

For those who wish to verify their **Excel** results or require additional metrics, using a **Skewness & Kurtosis Calculator** can be an excellent alternative. These online tools often provide a comprehensive look at the shape of your data by calculating both the skewness and the **kurtosis** simultaneously. Kurtosis is another statistical measure that describes the "tailedness" of the distribution, or how peaked or flat the distribution is relative to a normal distribution.

Using these calculators is generally a simple process of inputting the raw data into a designated text area and clicking a button to execute the calculations. This can be particularly useful for double-checking complex **algorithms** or for those who may not have immediate access to spreadsheet software. As shown in the comparison below, the values generated by these authoritative calculators should align perfectly with the results produced by Excel's internal **SKEW** function.

Skewness and Kurtosis Calculator

Skewness is a measure of the asymmetry of a dataset or distribution. This value can be positive or negative. A negative skew typically indicates that the *tail* is on the left side of the distribution. A positive value typically indicates that the tail is on the right.

Kurtosis is simply a measure of the "tailedness" of a dataset or distribution. The kurtosis formula used by this calculator is identical to the formula used in Excel, which finds what is known as *excess kurtosis*.

To find the skewness and kurtosis of a dataset, simply enter the comma-separated values in the box below, then click the "Calculate" button.

88, 85, 82, 97, 67, 77, 74, 86, 81, 95, 77, 88, 85, 76,
81, 82, 82, 84, 90, 91

CALCULATE

Skewness: -0.18490

By cross-referencing your Excel output with an external **statistical** calculator, you can gain higher confidence in your analysis. Whether you are preparing a report for academic submission or making high-stakes business decisions, ensuring the accuracy of your **skewness** and **kurtosis** measurements is a hallmark of professional data stewardship. Through the combination of Excel's powerful functions and external validation tools, you can achieve a sophisticated and accurate profile of any dataset.