

How to Calculate Percentiles in Excel: Understand the Difference Between PERCENTILE.INC and PERCENTILE.EXC

Authored by
stats writer

December 4, 2025

RECOMMENDED CITATION

stats writer (2025). *How to Calculate Percentiles in Excel: Understand the Difference Between PERCENTILE.INC and PERCENTILE.EXC*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=105214>

While calculating statistical measures in Microsoft Excel, users often encounter multiple functions designed for the same purpose, leading to confusion regarding which method is appropriate for their data analysis needs. This is particularly true for calculating the percentile of a dataset. Specifically, the distinction between the **PERCENTILE.EXC** and **PERCENTILE.INC** functions is a common point of inquiry for analysts and statisticians.

The fundamental difference between **PERCENTILE.EXC** and **PERCENTILE.INC** hinges on their handling of the data boundaries, specifically the 0 and 1 (or 0% and 100%) extremes of the percentile range. **PERCENTILE.INC** operates using an inclusive method, guaranteeing that the result will fall between the smallest and largest values within the provided data array. Conversely, **PERCENTILE.EXC** employs an exclusive method, which means it excludes the extreme endpoints, offering a slightly different calculation that may be preferred in specific statistical contexts where boundary values are considered outliers or are not relevant to the distribution study.

Understanding these subtle differences is critical because, although the resulting values from both functions are usually very close, they represent two different widely accepted methodologies for calculating quantiles. This comprehensive guide will dissect the underlying statistical methodologies, illustrate their application through practical examples, and provide clear recommendations for when to deploy the inclusive versus the exclusive percentile calculation in your data analysis workflow.

Understanding the Core Concept: What is a Percentile?

To fully appreciate the functional distinction between Excel's percentile functions, we must first establish a firm understanding of what a percentile represents. In descriptive statistics, the n th percentile of a distribution is the value at or below which n percent of the observations may be found. Essentially, percentiles divide a sorted dataset into 100 equal parts. When calculating the percentile, the first mandatory step is sorting the raw data values from the lowest magnitude to the highest magnitude.

Consider a simple example: if a student scores in the 85th percentile on a standardized test, it signifies that their score is greater than or equal to 85% of all other scores recorded. Conversely, only 15% of scores are higher than theirs. This measure is fundamentally important in fields like finance, healthcare, and performance metrics, as it provides relative standing rather than absolute measurement. Quantiles, such as percentiles (100 divisions), quartiles (4 divisions), and deciles (10 divisions), are essential tools for visualizing the shape and dispersion of data distributions, particularly in identifying central tendencies and potential skewness.

A critical statistical component of percentile calculation is **interpolation**, especially when the calculated index position does not correspond precisely to an actual data point in the array. Since

datasets often contain a finite number of discrete observations, the exact position representing, say, the 20.5th element in a 100-element set must be estimated. Interpolation is the process used to estimate a value that lies between two known data points. Excel's implementation of **PERCENTILE.EXC** and **PERCENTILE.INC** rely on slightly different interpolation techniques, which ultimately determines the final percentile value returned and is the source of their divergence.

The Evolution of Percentile Functions in Excel

Microsoft Excel has historically provided a variety of methods for statistical calculation, and the method for calculating percentiles has undergone refinements to align with contemporary statistical standards. Initially, Excel used a single function, PERCENTILE, which is now classified as a legacy function. Due to ambiguities in quantile calculation methods--statisticians generally agree on nine different methods--Excel introduced the suffixed functions to provide greater clarity and adherence to specific standards: **PERCENTILE.INC** and **PERCENTILE.EXC**.

The existence of these three functions provides backward compatibility while adhering to modern statistical conventions. The three functions available for percentile calculation are:

PERCENTILE.INC: This modern function calculates the kth percentile of a dataset using an inclusive method, which means it **includes** the bounds of 0 and 1 (or 0% and 100%) in the potential range of k.

PERCENTILE.EXC: This modern function calculates the kth percentile using an exclusive method, **excluding** the bounds of 0 and 1 from the percentile range.

PERCENTILE: This is the legacy function. It remains in Excel primarily for compatibility with older worksheets. Crucially, it returns the exact same result as the **PERCENTILE.INC** function, using the inclusive method by default. Users are generally advised to migrate to the dedicated **PERCENTILE.INC** or **PERCENTILE.EXC** functions for new analyses to improve clarity.

The use of separate, explicitly named functions (INC for inclusive, EXC for exclusive) ensures that the user is intentionally selecting the desired statistical method, reducing potential errors in communication and analysis replication. Although the legacy **PERCENTILE** function is still functional, relying on **PERCENTILE.INC** provides a cleaner, more understandable syntax consistent with other modern statistical functions in the Excel library.

Defining PERCENTILE.INC: The Inclusive Method

The **PERCENTILE.INC** function corresponds to the specific method (often labeled R-7) for calculating quantiles, which is widely recognized and frequently used in statistical software packages. The "INC" designation signifies that the percentile calculation is **inclusive** of the boundaries. When using **PERCENTILE.INC(array, k)**, the value of k (the percentile coefficient) is permitted to range from 0 to 1, inclusive. If k=0, the function returns the minimum value in the

array. If $k=1$, it returns the maximum value. This inclusivity ensures the calculated percentile value always falls within the observed range of the input data.

The index calculation for the inclusive method follows a specific formula. If N is the number of data points, the calculation uses the formula: $\text{Index} = k * (N - 1) + 1$. This formula ensures that when $k=0$, the index is 1 (the first element, or minimum), and when $k=1$, the index is N (the last element, or maximum). This method inherently treats the dataset as including all possible values between the minimum and maximum, providing a percentile calculation that is highly robust and representative of the entire data spread.

Because **PERCENTILE.INC** is inclusive of the end-points, it is often the default choice in many standard statistical applications and is aligned with the method used by many professional statistical textbooks. It provides a result that is intuitive: 0% corresponds exactly to the minimum observed value, and 100% corresponds exactly to the maximum observed value. This methodology is preferred in situations where the full range of observed data is statistically relevant, such as analyzing financial returns, medical trial data, or demographic distributions.

Defining PERCENTILE.EXC: The Exclusive Method

The **PERCENTILE.EXC** function implements a different quantile calculation method (often labeled R-6). The "EXC" designation indicates that the calculation is **exclusive** of the extreme boundaries, meaning the k value must be strictly greater than 0 and strictly less than 1 ($0 < k < 1$). This mathematical constraint ensures that **PERCENTILE.EXC** can never return the absolute minimum or maximum values of the dataset, even if k is set extremely close to 0 or 1.

The index calculation for the exclusive method differs significantly from the inclusive method. It uses the formula: $\text{Index} = k * (N + 1)$. By adding 1 to N before multiplication, the range of possible indices is effectively expanded, leading to a calculated index that always falls between the first and last data points when k is between 0 and 1. This method implicitly assumes that the observed dataset is a sample drawn from a larger population, and the percentile estimation should reflect this broader population distribution, rather than simply defining the boundaries based on the sample extremes.

Statistically, the exclusive method is particularly relevant when you are trying to estimate where a specific percentile falls within the underlying population distribution, and you wish to avoid the bias introduced by the absolute minimum and maximum values of the specific sample you possess. Because the function excludes 0 and 1, the calculated value will mathematically never be the minimum or maximum observed value, even if the user attempts to find the 1st or 99th percentile, leading to a result that might be slightly lower than the corresponding **PERCENTILE.INC** result for a small k , and slightly higher for a large k .

Step-by-Step Example: Calculating Percentiles in Practice

To highlight the practical outcome of using these two distinct methodologies, we will apply both functions to a simple numerical dataset within Excel. Suppose we are analyzing the following array of scores, ranging from 1 to 10:

Suppose we have the following dataset in Excel:

	A	B	C	D	E	F	G	H
1	Data							
2	2							
3	4							
4	5							
5	6							
6	6							
7	7							
8	9							
9	12							
10	14							
11	15							
12	18							
13	19							
14	22							
15	24							
16	26							
17	28							
18								
19								
20								
21								
22								
23								
24								

We are interested in calculating the 20th percentile ($k = 0.2$) for this set of numbers. We will use all three available Excel functions--**PERCENTILE.INC**, **PERCENTILE.EXC**, and the legacy **PERCENTILE**--to demonstrate how the choice of function influences the final output. The raw data is located in cells A2 through A11.

The following screenshot illustrates the formulas used and the resulting calculations for the 20th percentile:

	A	B	C	D	E	F
1	Data			20th percentile	Formula	
2	2		PERCENTILE	6	=PERCENTILE(A2:A17, 0.2)	
3	4		PERCENTILE.INC	6	=PERCENTILE.INC(A2:A17, 0.2)	
4	5		PERCENTILE.EXC	5.4	=PERCENTILE.EXC(A2:A17, 0.2)	
5	6					
6	6					
7	7					
8	9					
9	12					
10	14					
11	15					
12	18					
13	19					
14	22					
15	24					
16	26					
17	28					
18						
19						
20						
21						
22						
23						
24						
25						
26						

As clearly demonstrated by the results, the functions yield two distinct values. Using the **PERCENTILE.INC** function (or the identical legacy **PERCENTILE** function), we calculate the 20th percentile to be 6.

Using the **PERCENTILE.EXC** function we calculate the 20th percentile to be 5.4.

Interpreting the Results: Why the Values Differ

The difference between 6 (INC) and 5.4 (EXC) is directly attributable to the specific index calculation and subsequent interpolation method employed by each function. In our example dataset, N = 10 (ten data points), and we are seeking the 20th percentile (k = 0.2).

The disparity arises because the underlying formulas calculate different positional indices within the sorted array. Using the **PERCENTILE.INC** method, the index position calculation favors the inclusive range, resulting in an index of 2.8 for the 20th percentile (Index = 0.2 * (10 - 1) + 1 = 2.8). This position requires interpolation between the second data point and the third data point. The resulting interpolated value, as shown in the Excel output, is 6.

Conversely, for the **PERCENTILE.EXC** method, the index calculation uses the formula $k*(N+1)$, yielding an index of 2.2 (Index = $0.2 * (10 + 1) = 2.2$). This index is closer to the second data point than the 2.8 index calculated by the inclusive method. Since the index (2.2) is lower than the index (2.8), the resulting interpolated value (5.4) is also lower than the inclusive result (6). This comparison clearly illustrates that the exclusive method pushes the percentile estimate slightly inward, away from the boundary extremes, leading to the observed difference in the final calculated percentile value.

Practical Recommendations: Choosing the Right Function

In determining whether to use **PERCENTILE.INC** or **PERCENTILE.EXC**, the key consideration is whether the data is best treated as a full population or a small sample, and whether the extreme values (0% and 100%) should be achievable percentile results. For the vast majority of standard business intelligence, financial modeling, and academic analyses, the **PERCENTILE.INC** function is generally recommended and is considered the statistical standard.

The preference for the inclusive method stems from its ability to handle the full range of observed data. When analyzing a closed set of measurements--for example, the distribution of heights in a measured group, or the income range of a specific neighborhood--it is logical that the 0th percentile should correspond to the minimum observed height, and the 100th percentile to the maximum. Because **PERCENTILE.INC** aligns with this intuitive understanding and corresponds to the most common statistical definition (Method R-7), it is the function that aligns most closely with external statistical packages.

It is worth noting that the standard quantile methods used in popular programming environments, such as the [R programming language](#) and the [Python programming language](#) (specifically the NumPy library's 'linear' interpolation), default to methodologies that match or closely resemble the **PERCENTILE.INC** approach. If you are preparing data in Excel that will later be processed or compared against results from these platforms, using **PERCENTILE.INC** ensures better consistency and comparability across tools.

When Should PERCENTILE.EXC Be Used?

While **PERCENTILE.INC** is the default recommendation, there are niche scenarios where the **PERCENTILE.EXC** function may be statistically appropriate. The exclusive method is sometimes employed when analyzing theoretical distributions or when the sample size N is very small. The exclusive method (R-6) is often found in older statistical literature and some specific disciplines that prefer the index calculation based on $N+1$.

Since **PERCENTILE.EXC** requires k to be strictly between 0 and 1, it is implicitly designed for situations where the minimum and maximum observed values in the sample are not considered

true percentile points but rather boundaries of a finite sample. If the goal is to estimate the percentile rank for the population from which the sample was drawn, the exclusive method provides a definition that can be statistically slightly less biased for small samples, as it smooths the estimation by ensuring the extreme values are not returned.

However, users must be highly cautious when deploying **PERCENTILE.EXC**. If a user intends to find the 50th percentile (the median), both functions will typically yield the same result or a result extremely close to it, especially with large datasets. The divergence is most pronounced at the tails (near 0% and 100%). Unless a specific statistical methodology or institutional standard dictates the use of the exclusive method, adherence to the inclusive **PERCENTILE.INC** function simplifies communication and aligns with broader industry practices.

The Role of the Legacy PERCENTILE Function

The original PERCENTILE function in Excel has been superseded by the dedicated **PERCENTILE.INC** and **PERCENTILE.EXC** functions. As noted previously, the legacy **PERCENTILE** function is mathematically identical to **PERCENTILE.INC**. This means that if you open an older spreadsheet that utilizes the original function, the results will be consistent with the modern inclusive method.

While Microsoft maintains the legacy function for backward compatibility, new statistical models and templates should avoid its use. Relying on **PERCENTILE.INC** ensures that the intent of the calculation--to include the boundaries of the data distribution--is explicitly clear to any user or auditor examining the spreadsheet formulas. This practice enhances the transparency and maintainability of complex data analysis sheets, aligning with best practices for documentation and formula auditing.

Ultimately, the inclusion of the legacy **PERCENTILE** function acts as a bridge during the transition to standardized methods. For modern analysis, users should choose between the explicitly defined inclusive (INC) or exclusive (EXC) functions based on the statistical interpretation required for their specific dataset, completely bypassing the legacy function to ensure clarity.

Conclusion: Consistency and Context

The choice between **PERCENTILE.INC** and **PERCENTILE.EXC**, while sometimes yielding small numerical differences, represents a choice between two distinct statistical standards for interpolation. The inclusive method (INC) aligns with the common definition of quantiles, ensuring that 0% and 100% of the distribution correspond directly to the minimum and maximum observed values, respectively. The exclusive method (EXC) is designed to estimate percentiles for a population by assuming the sample boundaries are not the true extremes.

In almost all standard data analysis scenarios where a clear understanding of the sample's distribution is required, utilizing **PERCENTILE.INC** (or the legacy **PERCENTILE**) is the statistically sound and industry-consistent choice. The differences between the values calculated by **PERCENTILE.INC** and **PERCENTILE.EXC** will usually be minor, particularly in large datasets where the impact of the boundary index calculation diminishes substantially. However, precision in statistical reporting demands awareness of this methodological difference.

Regardless of which function is chosen, consistent application across a single analysis project is paramount. Analysts must document their choice to ensure reproducibility and accurate interpretation of results, especially when dealing with smaller datasets where the results of **PERCENTILE.INC** and **PERCENTILE.EXC** can diverge more noticeably. The key takeaway is that both functions are valid; the correct one depends entirely on the statistical context and the definition of the percentile required by the specific analysis.

No matter which function you use to calculate percentiles, the difference between the values calculated by **PERCENTILE.INC** and **PERCENTILE.EXC** will be very similar in most cases.

In some cases, it's even possible that the two functions will return the same values depending on the sequence of numbers in the dataset.