

Calculating a Five-Number Summary in Excel: A Comprehensive Tutorial

Authored by
Mohammed loot

November 8, 2025

RECOMMENDED CITATION

Mohammed loot (2025). *Calculating a Five-Number Summary in Excel: A Comprehensive Tutorial*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=13807>

Understanding the Foundation: What is a Five-Number Summary?

The [Five-Number Summary](#) (FNS) is a fundamental concept in descriptive statistics, offering a succinct, yet powerful way to characterize the distribution of any given [dataset](#). Unlike simple measures like the mean or standard deviation, the FNS relies entirely on positional statistics, making it highly robust against outliers. This set of five specific values provides a complete picture of the data's center, spread, and overall range, regardless of whether the distribution is skewed or symmetric.

When we calculate the FNS, we are essentially dividing the data into four equal segments, each containing 25% of the observations. This partition allows analysts to quickly grasp how data points are clustered and where the majority of the variance lies. The five crucial components that make up this summary are:

The **Minimum** value

The **First Quartile** (Q1)

The **Median** (Q2)

The **Third Quartile** (Q3)

The **Maximum** value

The FNS is invaluable because it provides a comprehensive overview of the data's characteristics without requiring complex calculations. It addresses three primary aspects of the distribution simultaneously. First, it pinpoints the central tendency using the [Median](#) (Q2), which is the middle value of the entire dataset. Second, it quantifies the dispersion or spread of the central 50% of the data by utilizing the First and Third [Quartiles](#). Finally, it defines the overall limits of the data through the Minimum and Maximum values, thus establishing the range. Understanding these five points allows for sophisticated comparisons between different datasets and informs further statistical modeling.

The Core Components: Minimum, Maximum, and Quartiles Explained

The Minimum and Maximum values are the simplest components, representing the smallest and largest observations, respectively. These two values define the total range of the data, indicating the distance between the lowest and highest data points. While essential for defining boundaries, the true utility of the [Five-Number Summary](#) emerges from the three measures of central position: the quartiles.

The quartiles are critical because they divide the dataset into four segments. The [Median](#) (Q2) splits the data exactly in half (50th percentile). The First Quartile (Q1) marks the 25th percentile, meaning 25% of the data falls below this value. Conversely, the Third Quartile (Q3) marks the 75th percentile, meaning 75% of the data falls below this value, and 25% falls above it. The distance

between Q3 and Q1 is known as the Interquartile Range (IQR), which measures the spread of the middle half of the data, offering a more robust measure of dispersion than the standard range, as the IQR is unaffected by extreme outliers.

Interpreting the spacing between these five numbers reveals the symmetry and potential skewness of the distribution. If the distance between the Minimum and Q1 is significantly larger than the distance between Q3 and the Maximum, it suggests the data may be skewed to the left. Conversely, if the section between Q3 and the Maximum is wider, the data may be skewed to the right. Furthermore, the position of the [Median](#) within the IQR (the box in a boxplot) indicates where the central bulk of the observations is concentrated. Analyzing these relationships is the first step in sophisticated data diagnostics.

Calculating the Summary in Microsoft Excel

Microsoft Excel provides straightforward functions for deriving all five required values, making the calculation of the [Five-Number Summary](#) highly accessible, even for large datasets. The process requires only two essential steps: organizing the raw data and applying the appropriate statistical functions. We recommend entering all data points into a single, contiguous column to ensure the formulas reference the correct range consistently.

Step 1: Enter the data values in one column.

Begin by inputting your raw observations into a vertical range, such as column B, starting from cell B2. For this example, we use a small sample dataset to illustrate the functions. This initial organization is crucial for the subsequent calculations, as all formulas will reference this specific range (e.g., B2:B17).

	A	B	C	D	E
1	Data values				
2	4				
3	6				
4	6				
5	7				
6	8				
7	12				
8	15				
9	17				
10	20				
11	21				
12	21				
13	23				
14	24				
15	27				
16	28				
17					
18					
19					

Step 2: Find the five number summary using dedicated Excel functions.

Once the data is prepared, we utilize specific Excel functions to calculate each component. The Minimum is found using the `MIN()` function, and the Maximum using the `MAX()` function. For the quartiles, Excel provides the robust `QUARTILE.INC()` function, which takes two arguments: the data array (our column B range) and the quartile number (1 for Q1, 2 for the [Median](#), and 3 for Q3). The following image illustrates the implementation of these formulas:

	A	B	C	D	E
1	Data values		Min	4	=MIN(A2:A16)
2	4		1st Quartile	7.5	=QUARTILE(A2:A16, 1)
3	6		Median	17	=MEDIAN(A2:A16)
4	6		3rd Quartile	22	=QUARTILE(A2:A16, 3)
5	7		Max	28	=MAX(A2:A16)
6	8				
7	12				
8	15				
9	17				
10	20				
11	21				
12	21				
13	23				
14	24				
15	27				
16	28				
17					
18					
19					

Based on the calculations displayed in the image above, which reference the data in column B, the comprehensive [Five-Number Summary](#) for this specific data is derived as follows, providing a clear statistical profile:

Minimum: 4 (Calculated using =MIN(B2:B17))

1st Quartile: 7.5 (Calculated using =QUARTILE.INC(B2:B17, 1))

Median: 17 (Calculated using =QUARTILE.INC(B2:B17, 2) or =MEDIAN(B2:B17))

3rd Quartile: 22 (Calculated using =QUARTILE.INC(B2:B17, 3))

Maximum: 33 (Calculated using =MAX(B2:B17))

Technical Deep Dive into Excel's QUARTILE Functions

While calculating quartiles in theory seems simple, different statistical software packages use slightly varying methods for interpolation, especially when dealing with smaller or odd-sized datasets. Microsoft Excel, recognizing these methodological differences, offers two primary functions for calculating [Quartiles](#), which are essential for precision in professional statistical analysis.

The historical function, `QUARTILE()`, is maintained for backward compatibility but is often superseded by the newer, more explicit functions: `QUARTILE.INC()` and `QUARTILE.EXC()`. Understanding the subtle difference between "inclusive" and "exclusive" methods is paramount for ensuring the accuracy of your descriptive statistics.

The `QUARTILE.INC()` function, which stands for inclusive, calculates percentiles using a "greater than or equal to" rule. This method is the more common and traditional approach, ensuring that the 0th percentile (Minimum) and 100th percentile (Maximum) are included in the calculation range. This is the default behavior used by the older `QUARTILE()` function and is generally preferred when calculating the traditional [Five-Number Summary](#).

Conversely, the `QUARTILE.EXC()` function, which is exclusive, calculates percentiles using a strict "greater than" rule. This method excludes the 0th and 100th percentiles from the calculation, resulting in slightly different Q1 and Q3 values, particularly in smaller datasets. Analysts typically use the inclusive method unless specific, advanced statistical requirements dictate the use of the exclusive method. Always be aware of which function you are deploying, as it directly impacts the resulting quartiles.

Technical Note: Understanding Quartile Methodologies

There are technically two quartile functions in Excel that determine how the percentile position is located:

QUARTILE.INC() - Calculates percentiles using "greater than or equal to" in an "inclusive" manner, including the minimum and maximum values (0% and 100%). This is the standard method for FNS calculation.

QUARTILE.EXC() - Calculates percentiles using "greater than" in an "exclusive" manner, excluding the 0% and 100% boundary points.

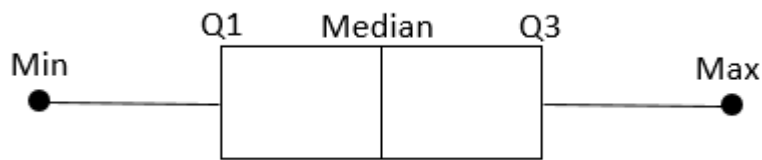
The default function **QUARTILE()** uses the **QUARTILE.INC()** method, which is recommended for consistency with standard statistical interpretation of the [Quartiles](#).

Visualizing Data Distribution with the Box-and-Whisker Plot

While the numeric summary is crucial, visualizing the Five-Number Summary significantly enhances understanding. The most common graphical tool for this purpose is the [Box plot](#), often called a box-and-whisker plot. This chart provides an immediate, visual representation of the data's central tendency, variability, and symmetry, as it is constructed entirely from the five values we calculated.

A [Box plot](#) consists of a central rectangular box and two lines, or "whiskers," extending from the

sides of the box. The box itself spans the Interquartile Range (IQR), representing the middle 50% of the data. The line inside the box marks the [Median](#), and the whiskers extend outward to the Minimum and Maximum values (or to points defined by an outlier rule).



Creating a [Box plot](#) in Excel is surprisingly simple using its built-in charting tools. This feature, introduced in more recent versions of Excel, automates the complex calculation and rendering process, allowing users to rapidly generate high-quality visualizations directly from their raw data.

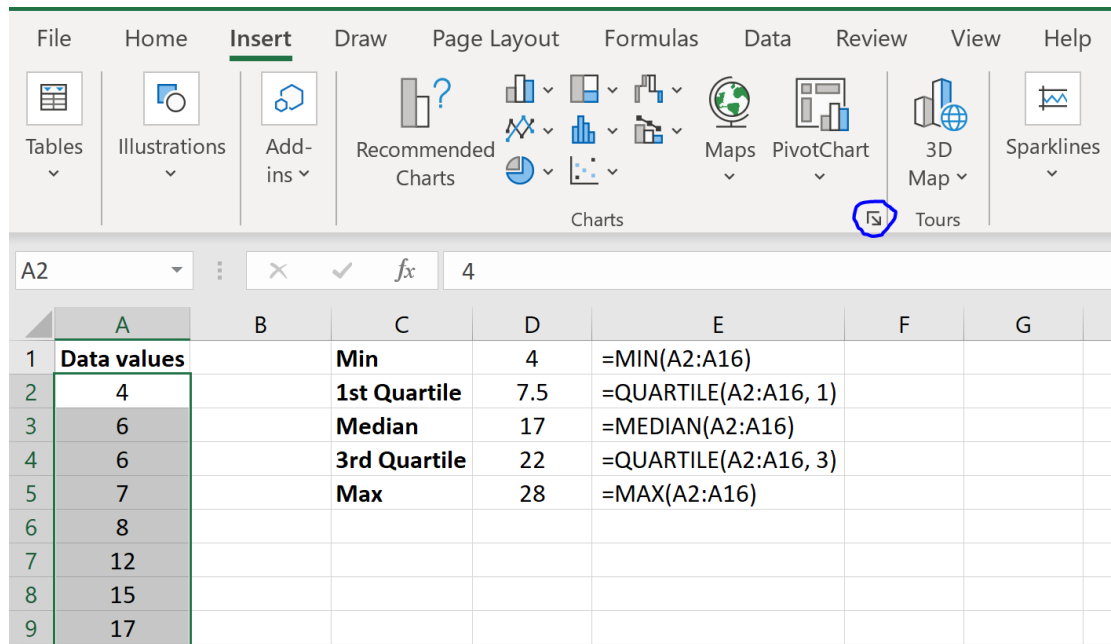
Step 1: Highlight the data values.

Select the entire column containing your raw numerical data (e.g., B2:B17 from our previous example). Ensure that you only select the numerical entries, excluding any column headers, unless you intend for the header to become the chart title.

	A	B	C	D	E
1	Data values		Min	4	=MIN(A2:A16)
2	4		1st Quartile	7.5	=QUARTILE(A2:A16, 1)
3	6		Median	17	=MEDIAN(A2:A16)
4	6		3rd Quartile	22	=QUARTILE(A2:A16, 3)
5	7		Max	28	=MAX(A2:A16)
6	8				
7	12				
8	15				
9	17				
10	20				
11	21				
12	21				
13	23				
14	24				
15	27				
16	28				
17					
18					
19					

Step 2: Access the chart insertion dialogue.

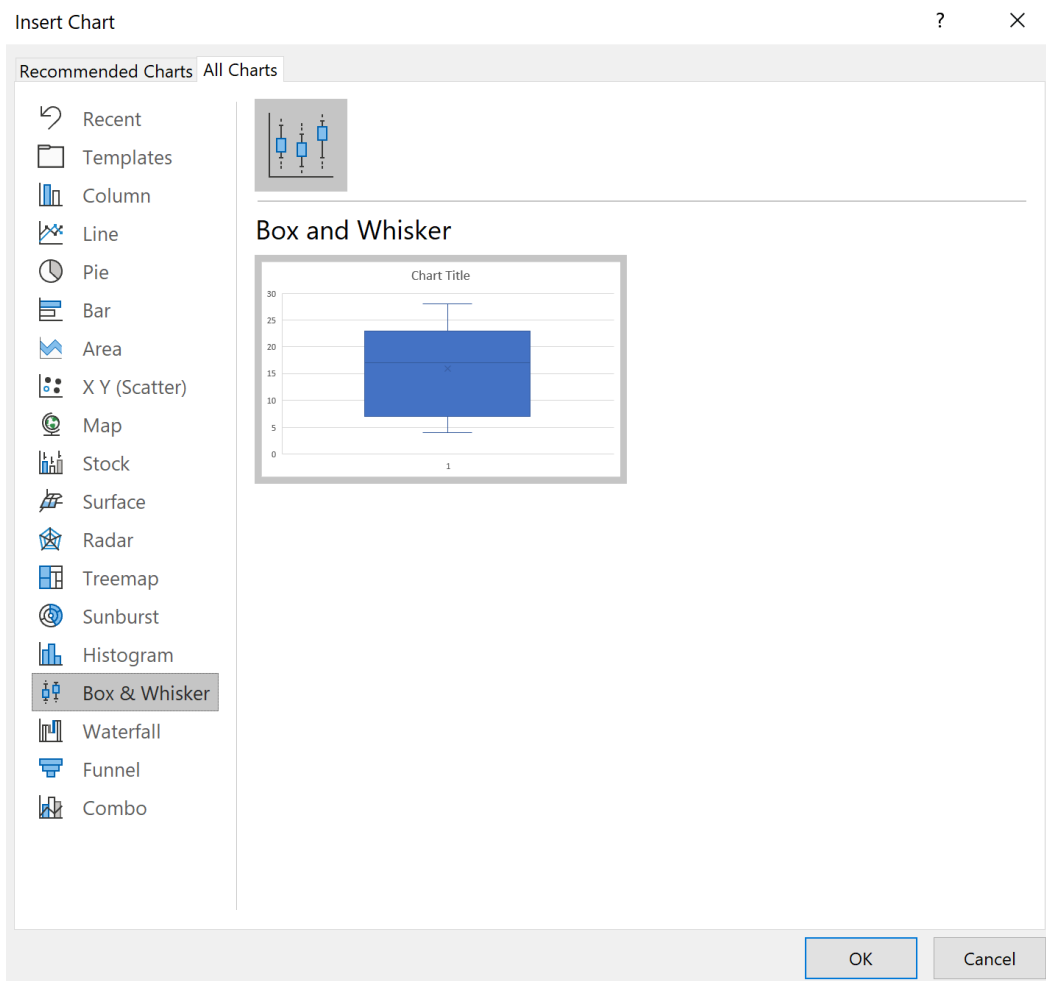
Navigate to the *Insert* tab on the Excel ribbon. Within the *Charts* group, look for the icon that resembles a small box or histogram, which often contains the statistical charts menu. Click the tiny arrow in the bottom right corner of the Charts group to open the "See All Charts" dialogue box, which provides access to advanced chart types.



	A	B	C	D	E	F	G
1	Data values		Min	4	=MIN(A2:A16)		
2	4		1st Quartile	7.5	=QUARTILE(A2:A16, 1)		
3	6		Median	17	=MEDIAN(A2:A16)		
4	6		3rd Quartile	22	=QUARTILE(A2:A16, 3)		
5	7		Max	28	=MAX(A2:A16)		
6	8						
7	12						
8	15						
9	17						

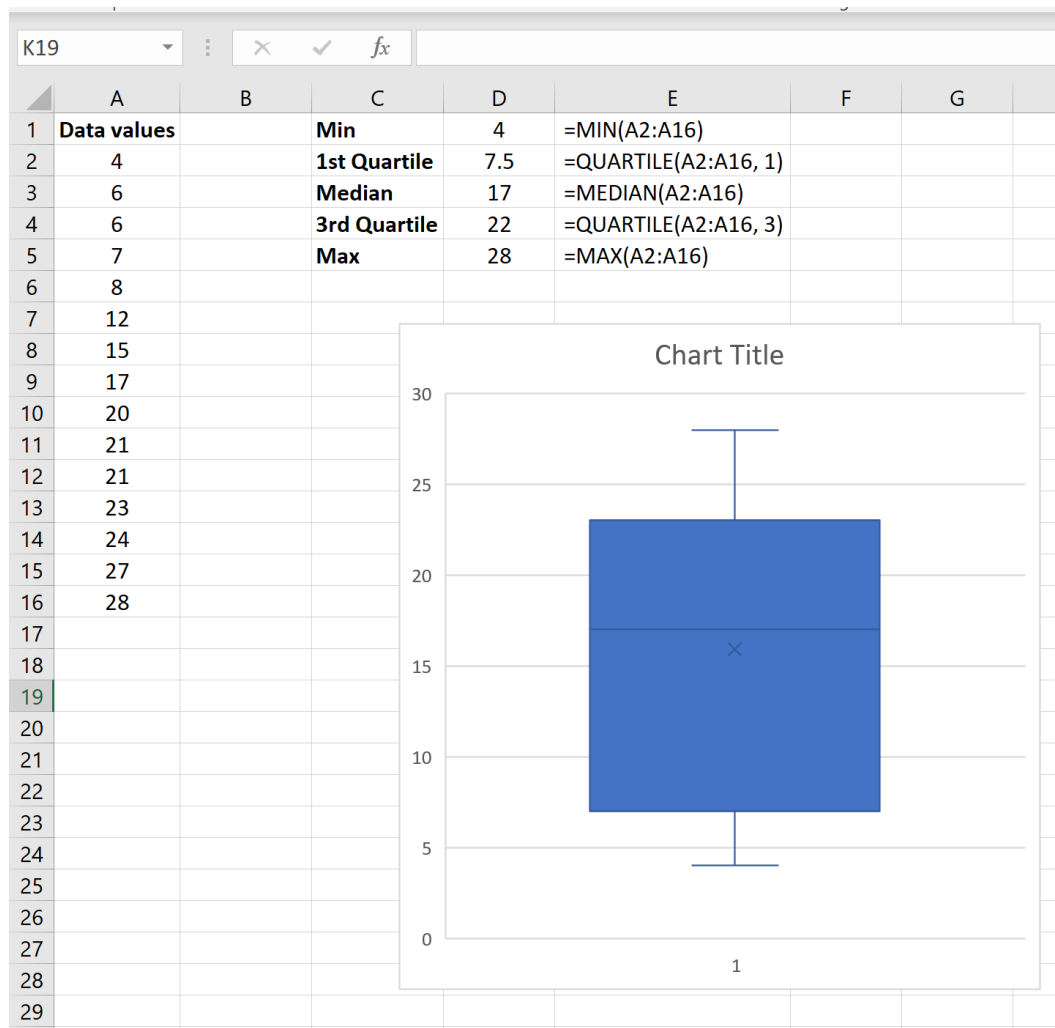
Step 3: Select "Box & Whisker" and finalize the chart.

In the "Insert Chart" dialogue box, click the "All Charts" tab on the left, and then scroll down to find and select the "Box & Whisker" option. Clicking OK will automatically generate the plot, using the selected data range to calculate the necessary [Quartiles](#) and identifying any potential outliers based on standard statistical rules (typically 1.5 times the IQR above Q3 or below Q1).



Interpreting the Excel Boxplot: Reading the Data Visually

Once the [Box plot](#) is generated, it serves as an intuitive visual key to the [Five-Number Summary](#). Every element of the plot corresponds directly to one of the five calculated values, plus an important additional measure: the mean. This visualization instantly reveals the skewness and concentration of the data that might be obscured in a simple list of numbers.



The interpretation of the Excel-generated [Box plot](#) is straightforward and maps precisely to the calculated FNS values:

The **Top Whisker** represents the **Maximum** value (the largest data point that is not considered an outlier).

The **Top Edge of the Box** represents the **Third Quartile (Q3)**, or the 75th percentile.

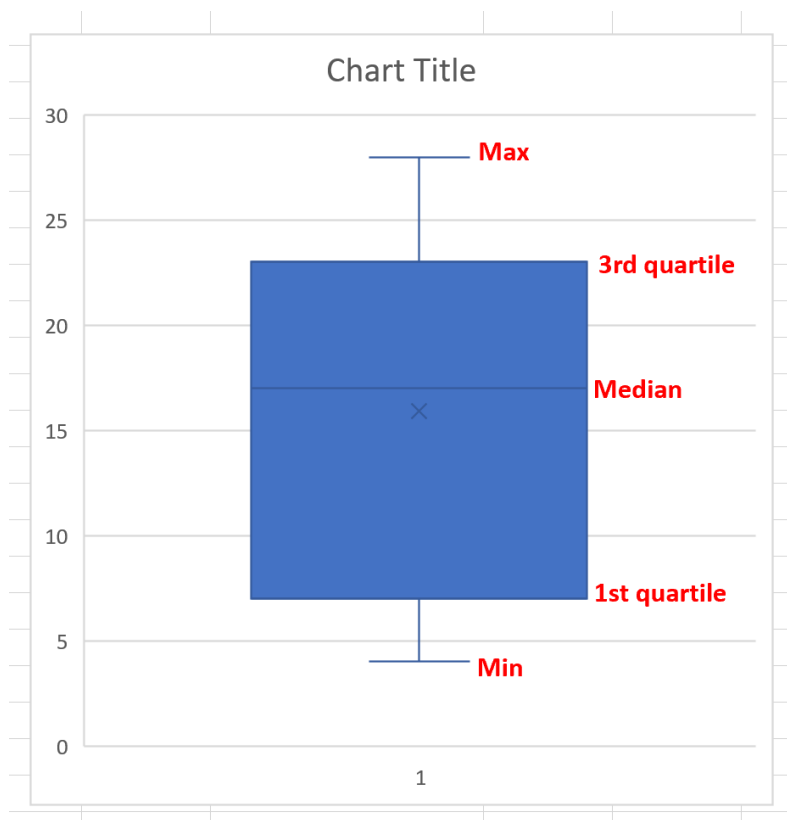
The **Line inside the Box** represents the **Median (Q2)**, or the 50th percentile.

A **Tiny 'X'** often appears inside the box, representing the arithmetic **Average (Mean)** of the data.

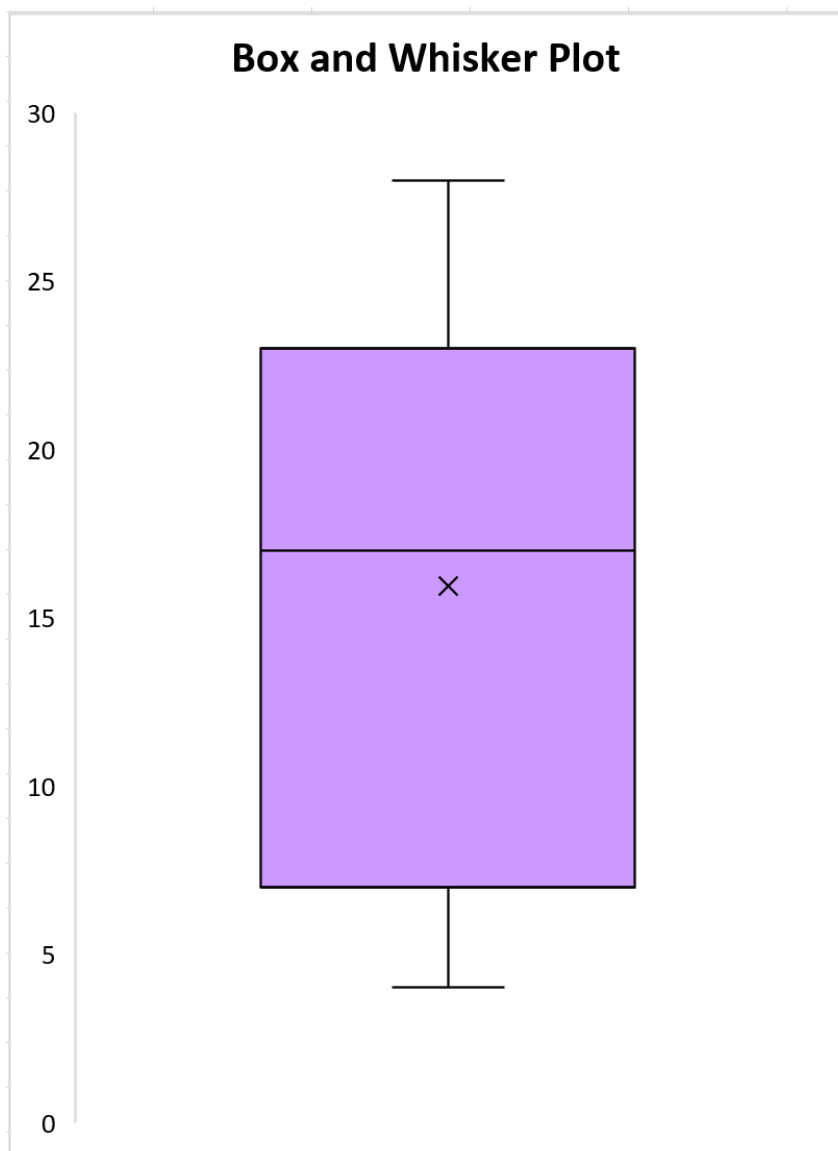
Comparing the position of the mean ('x') to the median (line) helps confirm skewness.

The **Bottom Edge of the Box** represents the **First Quartile (Q1)**, or the 25th percentile.

The **Bottom Whisker** represents the **Minimum** value (the smallest data point that is not considered an outlier).



The final step in visualization involves customizing the aesthetic elements of the plot. Excel allows users to easily modify the background color, chart title, axis labels, and font styles to ensure the graphic is clear, professional, and consistent with presentation standards. A well-formatted box plot effectively communicates the statistical characteristics of the [dataset](#) at a glance, transforming raw numbers into meaningful visual insights, which is the ultimate goal of descriptive analysis.



Additional Resources for Data Analysis

For those seeking to expand their knowledge of the Five-Number Summary beyond Microsoft Excel, the methodology is transferable across many statistical platforms and spreadsheet applications. The underlying principles of identifying the minimum, maximum, median, and [quartiles](#) remain universal.

The following links provide guides for implementing this powerful statistical summary in other popular data analysis environments:

[How to Calculate a Five Number Summary in Google Sheets](#)

[How to Calculate a Five Number Summary in SPSS](#)