

Learn How to Calculate a Five Number Summary in SPSS: A Step-by-Step Guide

Authored by
Mohammed looti

November 8, 2025

RECOMMENDED CITATION

Mohammed looti (2025). *Learn How to Calculate a Five Number Summary in SPSS: A Step-by-Step Guide*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=12770>

The **Five Number Summary** (FNS) stands as a foundational concept in descriptive statistics, designed to succinctly characterize the underlying shape and spread of a **dataset**. This powerful statistical tool distills complex distributions into just five key numerical values, offering analysts an immediate and clear snapshot of the data's central tendency, variability, and overall range. By relying on measures resistant to extreme outliers, the FNS facilitates rapid initial analysis, making large or complex distributions easily comprehensible before moving on to more intricate inferential tests. It is defined by these five essential measurements:

The **Minimum** observation, which establishes the lower boundary.

The **First Quartile** (Q1), representing the 25th percentile.

The **Median** (Q2), which is the true center of the distribution.

The **Third Quartile** (Q3), representing the 75th percentile.

The **Maximum** observation, which establishes the upper boundary.

The significance of the **five number summary** extends far beyond simple tabular reporting. These five values are the exact mathematical prerequisites for constructing **box plots** (or box-and-whisker plots), which are indispensable graphical tools in data visualization for comparing distributions across different groups or variables. By isolating these specific percentiles--the minimum, maximum, median, and quartiles--researchers can quickly assess the symmetry or skewness inherent in the distribution, allowing for efficient identification of potential issues or notable characteristics without the need to scrutinize every single data point.

This comprehensive guide is designed to empower researchers and students by detailing the exact contribution of each component within the FNS. Furthermore, we will provide a rigorous, step-by-step methodology for deriving this crucial summary efficiently and accurately for any given **dataset** using the industry-standard software, **SPSS** (Statistical Package for the Social Sciences). Understanding this process is fundamental to mastering exploratory data analysis.

Deconstructing the Five Number Summary Components

The **five number summary** provides a fundamentally robust alternative to measures like the mean and standard deviation, particularly when dealing with non-normal or skewed data distributions. It is specifically designed to describe the central tendency and the overall variability of observations in a way that is highly resistant to the influence of extreme scores, or outliers. A correct interpretation hinges entirely on understanding the distinct role that each of the five statistics plays in painting a complete picture of the data's characteristics.

The **Minimum** and **Maximum** values serve as the bookends of the distribution, establishing the absolute limits of the observed data range. They are crucial for defining the full extent of variability, quantifying the span from the lowest recorded observation to the highest. However, it is essential to remember that these two boundary values are the most susceptible to distortion by outliers; a

single, unusually high or low score can drastically inflate or deflate the perceived total range, often masking the true spread of the majority of the data points.

The core structure of the FNS is built upon the central measures: the [median](#) and the [quartiles](#). The [median](#) (often denoted as Q2) is the quintessential measure of central tendency for this summary. When the [dataset](#) is ordered from smallest to largest, the median is the value that perfectly splits the data into two equal halves, ensuring that precisely 50% of the observations fall below it and 50% fall above it. Because its calculation is based purely on rank position rather than magnitude, the median remains highly stable and is virtually unaffected by extreme values, offering a more reliable measure of the distribution's true center compared to the mean in skewed contexts.

Analyzing the Quartiles: Q1, Q3, and the Interquartile Range

The first and third [quartiles](#) complete the [five number summary](#), acting as internal markers that divide the 100% distribution into four distinct, equal quarters, with each containing 25% of the total observations. This partitioning process allows for a granular understanding of how the data is spread around the median and across the overall range. These quartile values are often referred to as the 25th and 75th percentiles, respectively, offering insight into the concentration of data within the central half of the distribution.

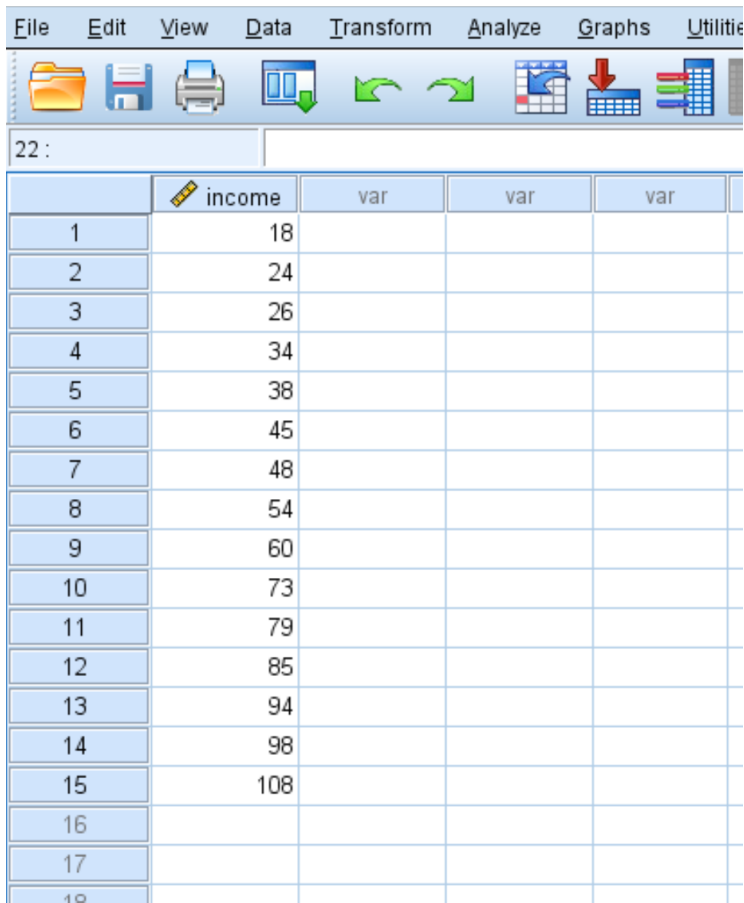
Specifically, the **First Quartile (Q1)** is the value marking the boundary below which the lowest 25% of the data points reside. Conversely, the **Third Quartile (Q3)** defines the point below which 75% of the data falls, meaning the top 25% of observations are located above it. The critical statistical measure derived directly from these two values is the [Interquartile Range \(IQR\)](#), calculated simply as the difference between Q3 and Q1 ($IQR = Q3 - Q1$). The IQR is indispensable as a measure of statistical dispersion.

The [Interquartile Range](#) quantifies the spread of the middle 50% of the [dataset](#). This specific focus on the central portion makes the IQR an exceptionally robust indicator of variability, as it systematically excludes the influence of the extreme 25% tails on either end--where most outliers reside. Consequently, a small IQR suggests that the central data points are tightly clustered around the median, indicating low variability in the core data. A large IQR, conversely, signals significant dispersion within the middle half of the distribution. Analyzing the relative distance between Q1, the median, and Q3 is highly effective for detecting subtle skewness or asymmetry in the distribution shape.

Step-by-Step Guide: Calculating the FNS in SPSS

To demonstrate the practical application of these concepts, we will walk through the process of calculating the [five number summary](#) using [SPSS](#) Statistics. For this example, we assume a researcher has gathered a sample [dataset](#) representing the annual income (measured in

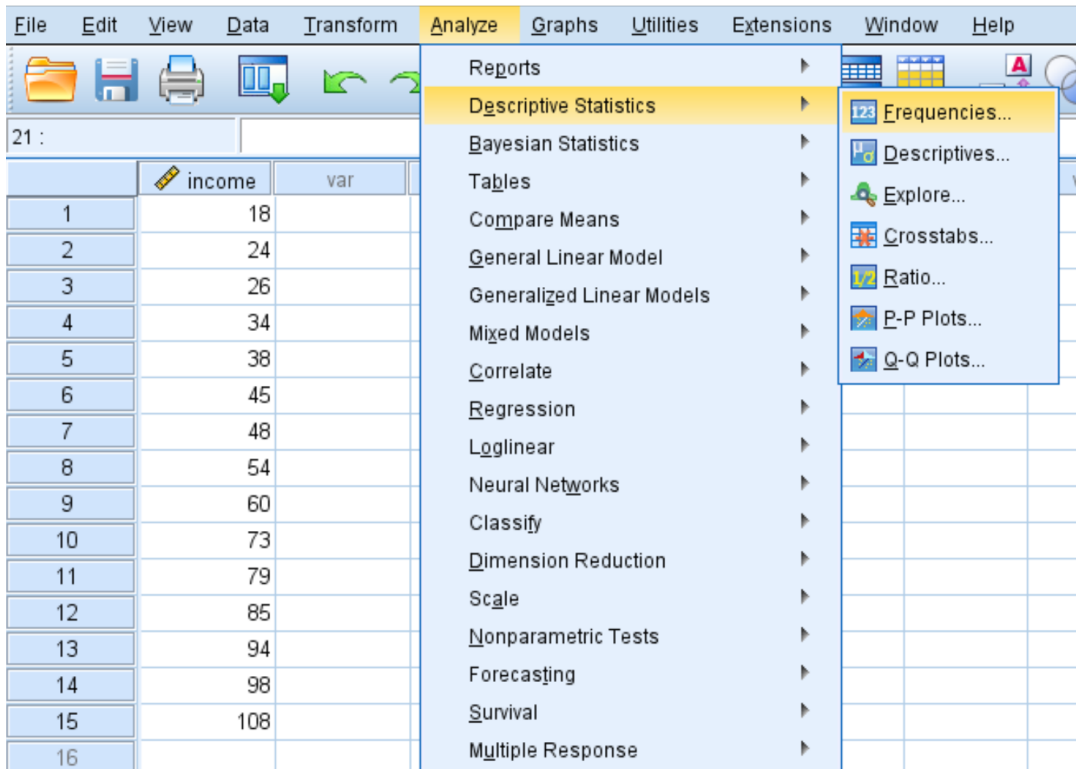
thousands of currency units) reported by 15 distinct individuals. Our objective is to leverage the powerful analytical capabilities of [SPSS](#) to rapidly and accurately derive the five crucial statistics for this income variable.



The screenshot shows the SPSS software interface. The menu bar includes File, Edit, View, Data, Transform, Analyze, Graphs, and Utilities. Below the menu bar is a toolbar with icons for file operations (Folder, Save, Print, Copy, Paste) and data manipulation (Undo, Redo, Find, Replace, Sort, Filter). The main window displays a data editor with a table containing 15 rows of data. The first column is labeled 'income' and the second column is labeled 'var'. The data values are: 18, 24, 26, 34, 38, 45, 48, 54, 60, 73, 79, 85, 94, 98, 108.

	income	var	var	var
1	18			
2	24			
3	26			
4	34			
5	38			
6	45			
7	48			
8	54			
9	60			
10	73			
11	79			
12	85			
13	94			
14	98			
15	108			
16				
17				
18				

The most efficient and user-friendly method for generating all five required statistics simultaneously in [SPSS](#) is through the utilization of the Frequencies procedure. This procedure is specifically designed for quick generation of descriptive statistics and initial distribution information. To initiate the calculation, navigate to the main menu bar at the top of the program interface. Click the **Analyze** tab, hover over **Descriptive Statistics**, and then select the **Frequencies** option from the cascading menu. This action will open the primary dialog box necessary for configuring the analysis.



Upon opening the Frequencies dialog box, the next immediate step involves identifying and specifying the variable of interest. Locate the variable, in this case **income**, within the variable list displayed on the left-hand side. Use the central arrow button to transfer this variable into the box labeled **Variable(s)**. This crucial step instructs **SPSS** precisely which column of data should be processed for the calculation of the descriptive measures. Once the variable is selected, you must proceed to define the specific statistics required for the analysis.

Configuring SPSS for the Five Number Summary

After successfully moving the target variable into the analysis box, the most critical configuration step is defining the exact statistics **SPSS** must calculate. To access the selection options, click the dedicated **Statistics** button located within the Frequencies dialog box. A new configuration window will immediately appear, presenting an extensive array of statistical measures covering central tendency, dispersion, and distribution characteristics. Researchers must exercise precision here to ensure only the necessary components of the FNS are requested.

To correctly instruct the software to calculate the complete **five number summary**, ensure that you check the boxes corresponding to all five elements. It is important to note the software efficiency: selecting the single option **Quartiles** automatically calculates both the First Quartile (Q1) and the Third Quartile (Q3). Therefore, to obtain the full set of five values, you need only check the following four selections under the 'Percentile Values' and 'Dispersion' sections:

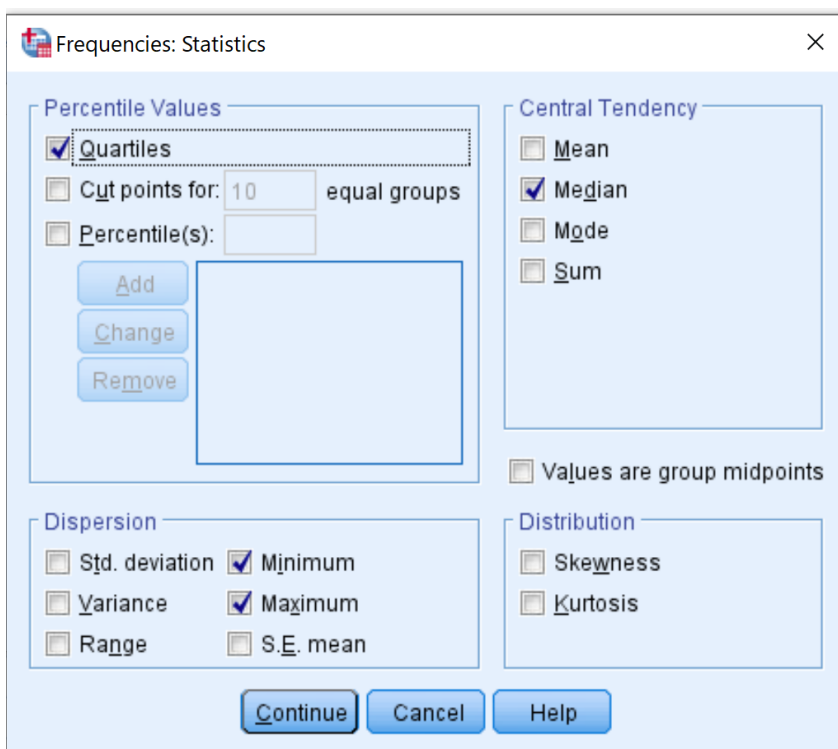
Check **Quartiles** (This selection generates Q1 and Q3).

Check **Minimum** (Defines the lower range boundary).

Check **Maximum** (Defines the upper range boundary).

Check **Median** (Provides the 50th percentile, Q2).

Once these four essential boxes are selected, confirm your choice by clicking the **Continue** button to close the Statistics window, which returns you to the main Frequencies dialog box. Before executing the analysis, it is generally recommended practice for descriptive summaries to uncheck the option labeled "Display frequency tables." Unless the detailed distribution table is specifically required, disabling this option results in a cleaner and more focused output viewer, presenting only the requested descriptive statistics.



Interpreting the SPSS Output Results

The final step in the process is executing the analysis. Click **OK** in the main Frequencies dialog box. **SPSS** will immediately run the procedure and generate the results within the dedicated Output Viewer window. The output is presented in a highly organized, tabular format, cleanly displaying the computed values for the **five number summary** based on the input income **dataset**. This table serves as the authoritative source for the derived descriptive statistics.

income		
N	Valid	15
	Missing	0
Median		54.00
Minimum		18
Maximum		108
Percentiles	25	34.00
	50	54.00
	75	85.00

By reviewing the generated table, we can formally articulate the structure of the income distribution. The output confirms the calculations for all five components, providing the minimum, maximum, [median](#), and the two [quartiles](#). For our sample data, the resulting summary is:

The **Minimum** observed income is: **18** (thousands).

The **First Quartile** (Q1) is: **34.00** (thousands).

The [Median](#) income (Q2) is: **54.00** (thousands).

The **Third Quartile** (Q3) is: **85.00** (thousands).

The **Maximum** observed income is: **108** (thousands).

Through the analysis of these five values, several critical characteristics of the distribution become apparent. We know the total income range spans from 18k to 108k. More importantly, the central 50% of the individuals report incomes falling strictly between 34k and 85k, defining the Interquartile Range (IQR). A closer examination of the relationship between the median and the quartiles reveals that the distance from the median (54k) to Q3 (85k), which is 31k, is noticeably larger than the distance from the median to Q1 (34k), which is 20k. This disparity strongly suggests a slight positive skew (or right skew) in the income distribution, indicating that the higher income values are more spread out and pull the upper tail further than the lower values.

Conclusion and Next Steps

Calculating the [five number summary](#) using [SPSS](#) represents a fundamental and highly effective step in any quantitative data analysis workflow. This procedure yields immediate, highly resistant insights into the dispersion, range, and central tendency of a variable, which is critical for initial data exploration, validation, and screening for anomalies. By mastering this simple, yet powerful technique, researchers gain the ability to quickly visualize the core shape of their data distribution, identify the potential presence of extreme values (outliers), and confidently move forward to more advanced inferential statistical procedures.

The FNS is an indispensable tool for summarizing numerical data, providing a foundation for graphical representation via box plots and robust metrics for comparison. For those analysts seeking to cross-validate their **SPSS** calculations or explore parallel descriptive methods within other software environments, the following resources are recommended:

[Five Number Summary Calculator](#)

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