

How to Create a Stem-and-Leaf Plot in SPSS: A Step-by-Step Guide

Authored by
stats writer

March 14, 2026

RECOMMENDED CITATION

stats writer (2026). *How to Create a Stem-and-Leaf Plot in SPSS: A Step-by-Step Guide*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=135671>

How can I create a Stem-and-Leaf plot in SPSS?

The **Stem-and-leaf plot** is a fundamental tool in the field of **exploratory data analysis**, originally popularized by the mathematician **John Tukey**. Unlike a standard **histogram**, which groups data into bins and loses individual values, this specific type of visualization preserves the raw data points while simultaneously providing a visual summary of the **frequency distribution**. By splitting each numerical value into a "stem" (the leading digits) and a "leaf" (the trailing digit), researchers can quickly identify the shape, center, and spread of a dataset without sacrificing the precision of the original observations.

Generating such a plot manually can be tedious, especially as the size of the dataset grows, which is why utilizing a powerful statistical software package like **SPSS** is highly advantageous. Within the **SPSS** environment, the process is streamlined through automated dialogs that handle the sorting and scaling of data. This allows for a more efficient workflow when performing **descriptive statistics**, enabling the analyst to focus on interpreting the underlying patterns rather than the mechanics of the drawing itself.

In the following sections, we will delve into a comprehensive, step-by-step guide on how to leverage the specialized features of **SPSS** to construct a clear and informative **Stem-and-leaf plot**. Whether you are analyzing clinical trial results, academic test scores, or sports performance metrics, mastering this visualization technique will significantly enhance your ability to communicate complex **quantitative data** to your audience.

The Concept and Utility of Stem-and-Leaf Displays

A **Stem-and-leaf plot** serves as a unique hybrid between a table and a graphical representation, offering a high-resolution view of **discrete or continuous variables**. The primary utility of this plot lies in its ability to reveal the **distribution** of the data while maintaining the integrity of every individual score. This is particularly useful for identifying **outliers** or unusual clusters that might be obscured in more generalized charts like boxplots or bar graphs.

The "stem" typically represents the higher-order place values, such as tens or hundreds, while the "leaf" represents the final significant digit. For instance, in a dataset consisting of exam scores ranging from 50 to 100, the tens digits (5, 6, 7, 8, 9, 10) would form the vertical stem, and the units digits would be listed horizontally as leaves. This structure allows the observer to see exactly how many data points fall into each interval, effectively functioning as a sideways histogram that uses the actual numbers as the "bars."

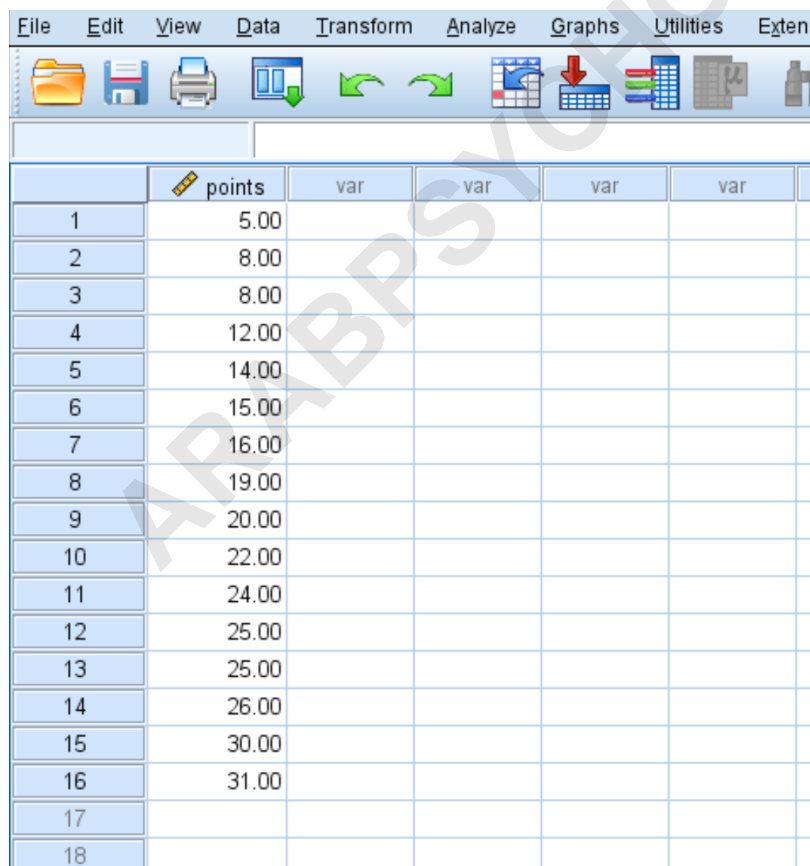
Furthermore, this visualization is an excellent choice for smaller datasets where every data point is crucial for the final analysis. It facilitates a quick assessment of **skewness** and **kurtosis**, providing

immediate feedback on whether the data follows a **normal distribution**. By examining the length of the rows of leaves, an analyst can instantly determine the **mode** of the dataset, as the longest row represents the most frequent value range.

Preparing Your Dataset for Analysis in SPSS

Before initiating the plotting process, it is essential to ensure that your data is correctly formatted within the **SPSS** Data Editor. Each observation should occupy a single row, and the variable you intend to plot must be defined with a **numeric measurement level**. Accurate data entry is the cornerstone of reliable **data analysis**, and **SPSS** provides a robust interface for organizing these variables effectively.

Consider the following example dataset, which tracks the average points per game for a group of 16 basketball players. This raw information provides the foundation for our visualization. You should verify that there are no missing values or formatting errors in the column, as these can disrupt the calculation of the stem intervals. In the **SPSS** variable view, ensure the "Type" is set to "Numeric" and the "Measure" is set to "Scale" to allow the software to perform the necessary mathematical operations.



The screenshot displays the SPSS Data Editor interface. The menu bar includes File, Edit, View, Data, Transform, Analyze, Graphs, Utilities, and Extensions. The toolbar contains icons for file operations (Open, Save, Print, Copy, Paste), navigation (Home, Back, Forward), and data management (Import, Export, Import Wizard, Export Wizard, Import Data Source, Export Data Source, Import Data Source Wizard, Export Data Source Wizard). The data grid shows 18 rows and 6 columns. The first column contains row numbers 1 through 18. The second column, labeled 'points', contains the following values: 5.00, 8.00, 8.00, 12.00, 14.00, 15.00, 16.00, 19.00, 20.00, 22.00, 24.00, 25.00, 25.00, 26.00, 30.00, 31.00. The remaining four columns are labeled 'var'.

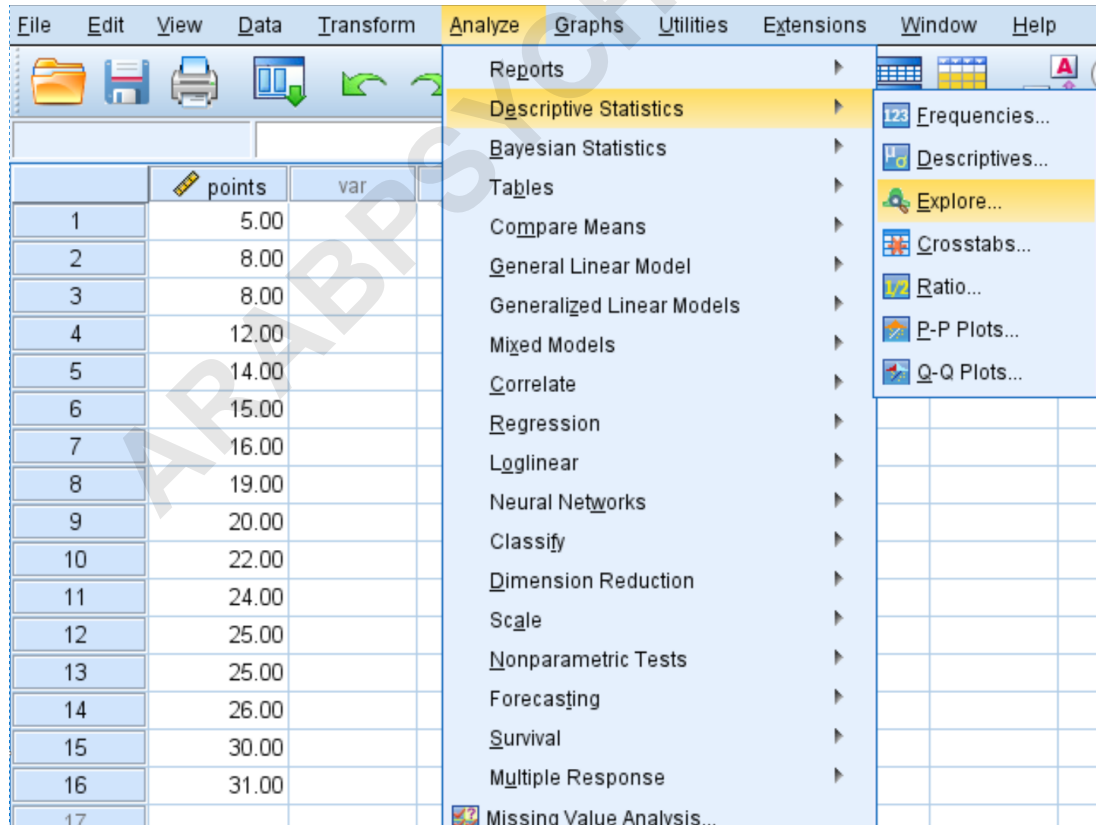
	points	var	var	var	var
1	5.00				
2	8.00				
3	8.00				
4	12.00				
5	14.00				
6	15.00				
7	16.00				
8	19.00				
9	20.00				
10	22.00				
11	24.00				
12	25.00				
13	25.00				
14	26.00				
15	30.00				
16	31.00				
17					
18					

Once the data is entered, it is often helpful to perform a preliminary check using the "Descriptive Statistics" menu to understand the range of your values. Knowing the **minimum** and **maximum** values will help you anticipate how the stems will be organized. For the basketball dataset shown above, the values range from single digits to the low thirties, suggesting that our stems will likely represent the tens place (0, 1, 2, and 3).

Navigating the SPSS Menu System

To begin the creation of the plot, you must navigate through the **SPSS** hierarchical menu system. The software organizes its statistical functions by their analytical purpose. For visual summaries and detailed data inspection, the "Explore" command is the most appropriate choice. This command is housed under the "Analyze" tab, which is the primary gateway for all **statistical** procedures in the software.

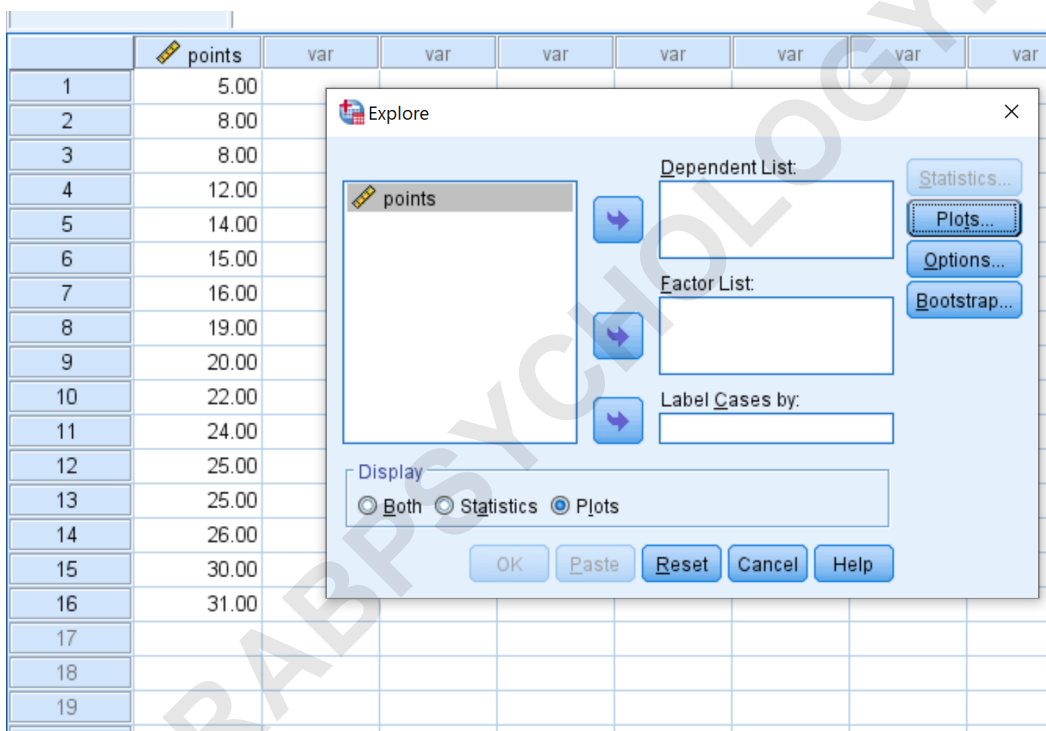
First, click on the **Analyze** tab located in the top menu bar. From the resulting dropdown menu, hover your cursor over **Descriptive Statistics** to reveal a secondary list of options. Within this submenu, select **Explore**. This specific path is chosen because the "Explore" function is designed to provide a comprehensive look at the data, offering not just the **Stem-and-leaf plot**, but also boxplots and various summary statistics in a single output window.



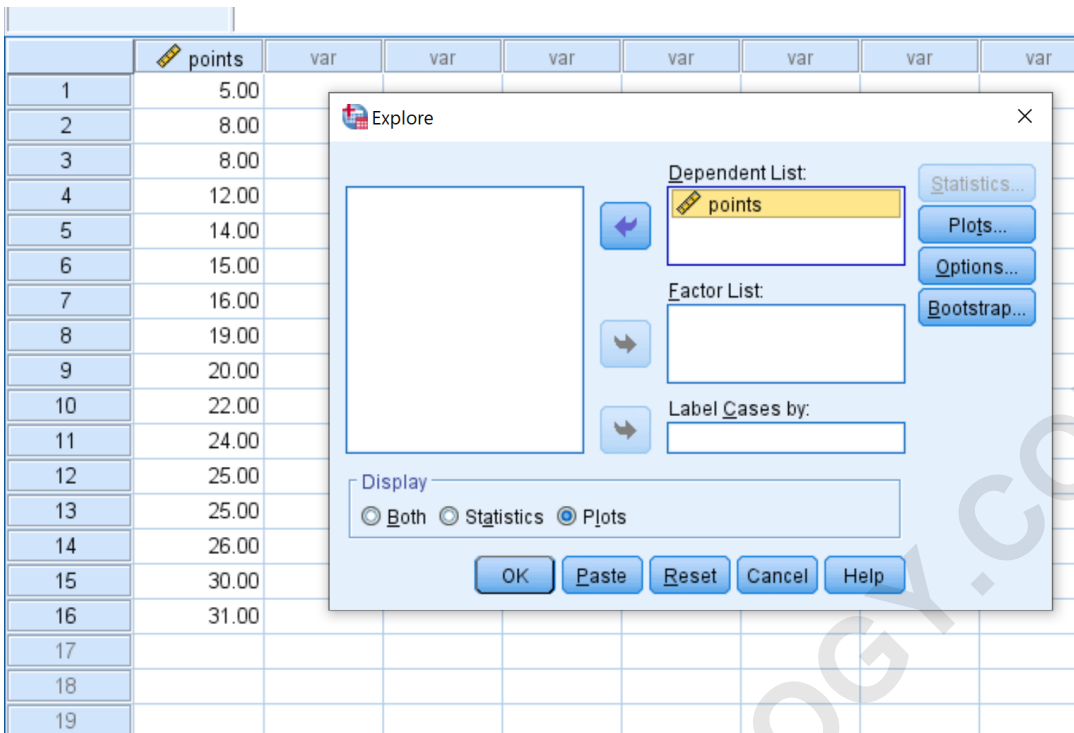
Using the "Explore" dialog is generally preferred over the "Legacy Dialogs" for this task because it offers more sophisticated handling of **dependent variables**. It allows you to split the analysis by factor levels if you wish to compare different groups, such as comparing the point distributions of two different teams. This flexibility makes it a staple for researchers performing initial **data cleaning** and verification.

Configuring the Explore Dialog Box

Once the **Explore** dialog box appears on your screen, you will see a list of your available variables in the left-hand pane. This interface is where you define the parameters of your analysis. You must select the variable you wish to visualize--in this case, "points"--and move it into the section titled **Dependent List**. This tells **SPSS** which **dataset** column it needs to process for the plot.



After moving the variable, look toward the bottom of the dialog box under the **Display** section. Here, you have three choices: "Both," "Statistics," or "Plots." While "Both" is often the default and provides a wealth of information, if your primary goal is the visualization, you can select **Plots**. This ensures that the output focuses on the graphical elements, making the results easier to digest for a quick **data visualization** check.



Before clicking the final execution button, you may also click on the "Plots..." button on the right side of the dialog box to verify that the **Stem-and-leaf** option is checked. In most modern versions of **SPSS**, this is selected by default when you enter the Explore menu. This level of customization allows you to toggle other visual aids, such as **normality plots with tests**, which can provide a mathematical confirmation of the distribution shape you see in the stem-and-leaf output.

Generating and Reviewing the Output

With the variables selected and the options correctly set, click the **OK** button. **SPSS** will process the request and open the "Output Viewer" window. This window contains the results of your analysis, usually starting with a Case Processing Summary followed by the requested plots. The **Stem-and-leaf plot** will appear as a text-based graphic, organized into distinct columns labeled "Frequency," "Stem," and "Leaf."

points

points Stem-and-Leaf Plot

Frequency	Stem &	Leaf
.00	0 .	
3.00	0 .	588
2.00	1 .	24
3.00	1 .	569
3.00	2 .	024
3.00	2 .	556
2.00	3 .	01

Stem width: 10.00
Each leaf: 1 case(s)

The output is remarkably intuitive once you understand the logic of the split. The **Stem** column serves as the anchor, representing the leading digit(s). The **Leaf** column contains the trailing digits, with each individual digit representing a single observation from your dataset. If multiple players scored the same number of points, or points within the same stem range, their individual "leaves" will be listed sequentially in the same row, sorted in ascending order.

This automated output also includes a "Stem width" and "Each leaf" indicator at the bottom. The **Stem width** tells you the mathematical scale--for instance, if the stem represents tens, the width is 10. The **Each leaf** indicator tells you how many cases each digit represents (usually 1). Understanding these scales is vital for **interpreting** the magnitude of the data accurately, especially when dealing with very large or very small numbers.

Detailed Interpretation of the Stem and Leaf Columns

To interpret the plot effectively, let's look at specific rows from our basketball example. The **Stem** column identifies the group, and the **Leaf** column identifies the specific value. In the first row of our output, the stem is 0 and the leaf is 5. Because the stem represents the tens place, this indicates a player who averaged 0 tens and 5 units--simply 5 points per game.

points Stem-and-Leaf Plot

Frequency	Stem &	Leaf
.00	0 .	
3.00	0 .	588
2.00	1 .	24
3.00	1 .	569
3.00	2 .	024
3.00	2 .	556
2.00	3 .	01

Stem width: 10.00
Each leaf: 1 case(s)

As we move down the plot, the stems increase. A stem of 1 with leaves of 2, 5, and 8 would represent three different data points: 12, 15, and 18. This layout allows you to see the **range** of the data at a glance. You can see where most players fall; if the "1" and "2" stems have the most leaves, you know the majority of the team averages between 10 and 29 points per game.

Finally, the last row in our example shows a stem of 3 and a leaf of 1. This represents a high-performing player who averages 31 points per game. By looking at the distance between this row and the others, you can visually assess the **dispersion** of the scores. If there were many empty stems between the 2 and the 3, it would suggest that the 31-point score is an outlier significantly removed from the rest of the team's performance.

points Stem-and-Leaf Plot

Frequency	Stem &	Leaf
.00	0 .	
3.00	0 .	588
2.00	1 .	24
3.00	1 .	569
3.00	2 .	024
3.00	2 .	556
2.00	3 .	01

Stem width: 10.00
Each leaf: 1 case(s)

Advantages of Stem-and-Leaf Plots over Histograms

While both **Stem-and-leaf plots** and **histograms** show the shape of a distribution, the former offers distinct advantages for detailed analysis. A histogram provides a "big picture" view by aggregating data into bins, which is excellent for very large datasets where individual values are less important. However, once the histogram is created, you can no longer see the original data points without referring back to the source table. The **Stem-and-leaf plot**, conversely, is a **lossless** visualization.

Another benefit is the ease with which one can calculate the **median** and **quartiles** directly from the plot. Since the leaves are already sorted in ascending order, you can simply count through the leaves to find the middle value. This makes the plot an invaluable tool for students and researchers who need to verify their **statistical parameters** manually or explain their findings to others without complex software output.

In **SPSS**, the **Stem-and-leaf plot** also automatically handles the scaling of stems. If your data spans a massive range, **SPSS** might use a stem width of 100 or 1000, or it might "split" stems (e.g., having two rows for the "1" stem, one for leaves 0-4 and another for 5-9) to maintain a readable shape. This intelligent **algorithm** ensures that the plot remains informative regardless of whether you are looking at small test scores or large financial figures.

Summary and Best Practices for Data Visualization

In conclusion, creating a **Stem-and-leaf plot** in **SPSS** is a straightforward yet powerful way to gain deep insights into your numerical data. By following the **Analyze > Descriptive Statistics > Explore** workflow, you can generate a professional visualization that facilitates better data understanding and more accurate **statistical inference**. This plot is an essential component of any researcher's toolkit, bridging the gap between raw data and abstract graphical summaries.

When using these plots in your reports, it is a best practice to always include the "n" (sample size) and to ensure the stem width is clearly labeled. While **SPSS** does this automatically, the analyst should double-check these values to ensure they haven't been misinterpreted. Clear communication of **metadata** is just as important as the data itself when it comes to reproducible research and professional presentations.

Ultimately, the choice of visualization should depend on your audience and your specific analytical goals. For exploratory phases, the **Stem-and-leaf plot** is often superior to almost any other chart because of its transparency. By mastering the **SPSS** tools described in this guide, you are well-equipped to conduct thorough and insightful analyses of any **quantitative data** you encounter in your professional or academic career.