

# Understanding Stem-and-Leaf Plots: A Guide to Calculating Mean, Median, and Mode

Authored by  
**Mohammed loot**

November 4, 2025

## RECOMMENDED CITATION

Mohammed loot (2025). *Understanding Stem-and-Leaf Plots: A Guide to Calculating Mean, Median, and Mode*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=9921>

Data visualization is fundamental to statistical analysis, providing clarity and insight into raw numbers. Among the various tools available, the [stem-and-leaf plot](#) stands out as a unique and effective method for displaying the distribution of a [dataset](#) while retaining all original data points.

Unlike histograms, which group data into bins and lose the individual values, a stem-and-leaf plot displays data by splitting up each value in the [dataset](#) into a *stem* and a *leaf*. This structure makes it exceptionally useful for calculating key measures of [central tendency](#) directly from the plot, without needing to recreate a separate list of numbers.

Understanding how to interpret this plot is essential. The stem typically represents the larger place value digits (e.g., tens or hundreds), while the leaf represents the smallest place value digit (e.g., ones). The resulting visual display resembles a bar chart rotated 90 degrees, offering a quick view of data symmetry and density.

Points Scored	
Stem	Leaf
1	1, 2, 2, 3
2	2, 3, 6, 6
3	0, 1, 2
4	2, 5, 6
5	1, 2
6	1

This comprehensive tutorial explains precisely how to calculate the three primary measures of [central tendency](#)—the [mean](#), [median](#), and [mode](#)—directly from a given stem-and-leaf plot. Mastering these calculations is crucial for summarizing the characteristics of any statistical distribution.

## Deconstructing the Stem-and-Leaf Plot Structure

Before proceeding with calculations, it is necessary to fully grasp how the plot encodes information. Every row in the plot starts with a stem value, followed by one or more leaf values. The key, which is usually provided alongside the plot, defines the precise relationship between the stem and the leaf, indicating the decimal place or unit measurement used.

For example, if a stem is '4' and a leaf is '5', and the key states "4|5 = 45 units," then the data point represented is 45. If the key stated "4|5 = 4.5 units," the data point would be 4.5. This structural

clarity is why the [stem-and-leaf plot](#) remains a powerful tool for exploratory data analysis, particularly with smaller [datasets](#).

The primary advantage of this structure is that it automatically sorts the data. Since the stems are typically listed in ascending numerical order and the leaves within each stem are also ordered, the analyst immediately has a ranked list of data points. This feature significantly simplifies the process of finding the [median](#) and range.

## Case Study: Plant Height Dataset

To illustrate the calculation methods, let us consider a specific example. Suppose we have the following [stem-and-leaf plot](#) which represents the measured height (in centimeters) of 19 different plants. The key indicates that the stem represents the tens digit and the leaf represents the ones digit.

Plant Height (inches)	
Stem	Leaf
1	1, 2, 3
2	6, 6
3	0, 1, 2
4	2, 5, 6
5	1, 2
6	1
7	8
8	2, 2
9	3, 4

The first step in any analysis is to extract the complete list of individual data points from the plot. By combining the stem and leaf values sequentially, we obtain the raw data set. This ensures that every calculation performed subsequently is based on the correct input values.

From the plot above, we derive the following 19 observations:

Stem 1: 11, 12, 13

Stem 2: 26, 26

Stem 3: 30, 31, 31

Stem 4: 42, 45, 46

Stem 5: 51, 52

Stem 6: 61

Stem 7: 78

Stem 8: 82, 82

Stem 9: 93, 94

The complete ordered [dataset](#) is: 11, 12, 13, 26, 26, 30, 31, 31, 42, 45, 46, 51, 52, 61, 78, 82, 82, 93, 94. We note that the total [sample size](#) (N) is 19.

## Calculating the Mean (Average Value)

The [mean](#), often referred to as the arithmetic average, is the sum of all values in a [dataset](#) divided by the total number of values. It is the most commonly used measure of [central tendency](#), representing the typical value within the distribution.

To find the [mean](#) ( $\bar{x}$ ) of this particular dataset, we must first sum all 19 individual height values and then divide that sum by the total [sample size](#), which is  $N = 19$ . The formula for the mean is:  $\bar{x} = (\sum x_i) / N$ .

The summation of all values is performed as follows:

Sum ( $\sum x_i$ ) = (11 + 12 + 13 + 26 + 26 + 30 + 31 + 31 + 42 + 45 + 46 + 51 + 52 + 61 + 78 + 82 + 82 + 93 + 94).

The calculated sum is 906.

Now we apply the division:

Mean =  $906 / 19 = 47.6842...$

Rounding to two decimal places, the [mean](#) height of the plants turns out to be **47.68** cm. This statistic is sensitive to outliers and provides a balanced point for the entire distribution.

## Determining the Median (Middle Value)

The [median](#) represents the exact middle value in a ranked [dataset](#). Unlike the [mean](#), the median is robust against extreme outliers, making it a valuable measure when data distribution is skewed. Because the stem-and-leaf plot inherently orders the data, finding the median is straightforward.

Since our [sample size](#) N is 19 (an odd number), the position of the median value can be determined using the formula: Position =  $(N + 1) / 2$ . In this case, Position =  $(19 + 1) / 2 = 10$ . We

are looking for the 10th value in the ordered list.

We review the ordered list of heights derived from the plot and count to the 10th position:

11, 12, 13, 26, 26, 30, 31, 31, 42, **45**, 46, 51, 52, 61, 78, 82, 82, 93, 94

The 10th value is 45. Therefore, the [median](#) height is **45** cm. This means that half of the measured plants are shorter than 45 cm, and half are taller than 45 cm. If the sample size were even, the median would be the average of the two middle values.

## Identifying the Mode (Most Frequent Value)

The [mode](#) is the value or values that appear most frequently in a [dataset](#). A dataset can have one mode (unimodal), two modes (bimodal), or multiple modes (multimodal). If all values occur only once, the dataset has no mode. Identifying the mode is particularly straightforward with a stem-and-leaf plot, as repeated leaves within a single stem, or across multiple stems, are clearly visible.

We systematically examine the ordered list to count the frequency of each observation:

11, 12, 13 (Occur once)

**26** (Occurs twice)

30 (Occurs once)

31 (Occurs twice)

42, 45, 46 (Occur once)

51, 52 (Occur once)

61 (Occurs once)

78 (Occurs once)

**82** (Occurs twice)

93, 94 (Occur once)

By inspecting the frequencies, we observe that the values 26, 31, and 82 all occur exactly twice, which is more frequent than any other individual value.

Because there are three values tied for the highest frequency, the dataset is classified as trimodal, and the modes are **26**, **31**, and **82**. The [mode](#) offers insights into typical occurrences or "peaks" in the data distribution, which is useful when analyzing qualitative data or discrete counts.

## Summary of Measures and Contextual Application

The analysis of the plant height data using the three measures of [central tendency](#) provides a complete statistical profile:

[Mean](#): 47.68 cm

[Median](#): 45 cm

[Mode](#): 26, 31, 82 cm

The close proximity of the [mean](#) (47.68) and the [median](#) (45) suggests that the data distribution is relatively symmetric, although the presence of multiple modes indicates distinct clustering of plant heights. The difference between these two measures (2.68 cm) is small, confirming that there are no extreme outliers pulling the mean significantly away from the center.

When presenting these findings, it is vital to choose the measure that best reflects the purpose of the study. If the goal is to calculate the total biomass (which depends on the magnitude of all values), the [mean](#) is superior. If the goal is to find the typical height unaffected by extremely tall or short plants, the [median](#) is the preferred metric.

## Further Resources for Statistical Analysis

To deepen your understanding of data visualization and statistical metrics, consider exploring related topics such as box plots, histograms, and measures of dispersion like variance and standard deviation. These tools work in tandem with the [stem-and-leaf plot](#) to provide a holistic view of data characteristics.

We recommend the following resources for continued learning:

Official documentation on exploratory data analysis techniques.

Academic journals focusing on quantitative research methodologies.

Tutorials detailing the construction and interpretation of advanced data plots.

Mastering these fundamental concepts provides a strong foundation for any discipline requiring rigorous data interpretation and reporting.