## Sophia.

## Stem-and-Leaf Plots

## by Sophia

## : = WHAT'S COVERED

This tutorial will cover the topic of stem-and-leaf plots. Our discussion breaks down as follows:

1. Stem-and-Leaf Plots
2. Variations

2a. Splitting Stems
2b. Rounding
2c. Two-digit leaves.
2d. Back-to-Back
3. Advantages of Stem-and-Leaf Plots

## 1. Stem-and-Leaf Plots

While these graphs, known as stem-and-leaf plots, have a funny name, they actually serve a useful purpose and are very versatile.

Many quantitative data sets can be displayed in stem-and-leaf plots, such as the one below:
Percent of College Students Enrolled in Public Colleges by State

| 95 | 95 | 87 | 76 | 70 | 81 | 43 | 90 | 84 | 55 |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| 84 | 81 | 84 | 83 | 75 | 76 | 85 | 89 | 84 | 81 |
| 56 | 80 | 52 | 80 | 63 | 89 | 73 | 56 | 62 | 72 |
| 91 | 96 | 82 | 80 | 73 | 86 | 81 | 87 | 55 | 82 |
| 74 | 79 | 89 | 92 | 87 | 85 | 77 | 75 | 88 | 82 |

This data set represents the 50 states in the United States, and these numbers are the percent of college students in each state that are enrolled in public colleges. For example, in one state, $95 \%$ of its college students are in public schools, whereas in another state, only $52 \%$ are enrolled in public colleges.

To create a stem-and-leaf plot, follow these steps:

## 解 STEP BY STEP

Step 1: Decide on a natural classification. Here, 10 seems like an obvious choice. Those are going to be our bins. You should also choose your bins based on some digit. In this case, go by the tens digit. Note, if these numbers were in the hundreds, you might decide to go by the hundreds digit, though you could still go by the tens if desired.

Step 2: Next, create "stems." These are going to be the stems based on the bins you selected.

| 9 | $\quad$ |
| :--- | :--- |
| 8 |  |
| 7 |  |
| 6 |  |
| 5 |  |
| 4 |  |

The 9 means that this is going to be a state with 90 percent, or a percent in the 90 's, of their students at public school; the 8 represents states where the percent is in the 80 's, and so forth. Write them in order, least to greatest or greatest to least (the direction doesn't really matter), to the left of a vertical line, as shown above.

Step 3: List the values by their ones digit ascending away from those stems. Those are considered the leaves. When it is completed, it will look like the one shown below. Notice, for example, that $43 \%$ of the students in one state were at a public school.

```
9012556
8000 0 1 1 1 1 2 2 2 2 3 4 4 4 4 5 5 6 7 7 7 8 9 9 9
7002 3 3 4 5 5 6 6 7 9
6 3
525566
4 3
```


## $\square$ HINT

It is important to note that if a value appears more than once, you will list it more than once. For example, 80 appears three times. Also, note that those numbers are ascending away from the stem.

Step 4: Create a key. There's one more important feature of a stem-and-leaf plot. You need to be able to convey to someone who's looking at it, exactly what they're looking at. For example, in our graph above, the "6 bar 2" means that there's a state that has $62 \%$ of its students going to public colleges. Therefore, make this clear to the reader by saying, in a key, " 4 bar 3 means $43 \%$." This tells the reader how these numbers should be interpreted.

```
901 2 5 5 6
8000401414crllllllllllllllllllll
7 0 2 3 3 4 5 5 6 6 7 9
6 2 3
525566
4
```


## - TERM TO KNOW

## Stem-and-Leaf Plot/Stemplot

A distribution of quantitative data that shows natural numerical breaks in the data as categories called "stems" and individual values as "leaves."

## 2. Variations

There is more than one way to display data on a stem-and-leaf plot. Other variations include:

- Splitting the stems
- Rounding the data
- Using two-digit leaves
- Showing two groups with a back-to-back plot


## 2a. Splitting Stems

In our previous graph, the 80's had more than any other grouping. In fact, they had more than twice as much as any other single grouping, which looked a little strange. Is there anything we can do about that?

Suppose that you decided that tens were too wide of a bin. Instead, you could break it down by fives, and then write two 8 's--a low 8 and a high 8, or 85 to 89 for the high, and 80 to 84 for the low. Note, though, that if you're going to split one bucket, you need to split them all.

| 9 | 5 | 5 | 6 |
| :--- | :--- | :--- | :--- |
| 9 | 0 | 1 | 2 |


| 8 | 5 | 5 | 6 |  | 7 | 7 | 7 | 8 | 9 | 9 | 9 |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 8 | 0 | 0 | 0 |  | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 3 | 4 | 4 | 4 | 4 |
| 7 | 5 | 5 | 6 |  | 6 | 7 | 9 |  |  |  |  |  |  |  |  |  |
| 7 | 0 | 2 | 3 |  | 3 | 4 |  |  |  |  |  |  |  |  |  |  |
| 6 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 6 | 2 | 3 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 5 | 5 | 5 | 6 |  | 6 |  |  |  |  |  |  |  |  |  |  |  |
| 5 | 2 |  |  |  |  |  |  |  |  |  |  | Key: 4 \| 3 means 43\% |  |  |  |  |
| 4 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 4 | 3 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

If you split the stems into lows and highs, the graph will look like the one above. Because this separates the stems so that no one stem has so much more data than any other, this is a more of an appropriate visual than the first one.

## 2b. Rounding

Take a look at this set of high school GPAs for this group of students. Make a stem-and-leaf plot of these GPAs.

| Student | GPA |
| :--- | :--- |
| Amy | 2.95 |
| Blake | 3.55 |
| Holly | 3.75 |
| Isaiah | 1.94 |
| Jenny | 2.23 |
| Jesse | 3.41 |
| Jim | 1.96 |
| Johnathan | 2.25 |
| Katherine | 2.56 |
| Kelly | 2.89 |
| Ryan | 3.24 |
| Sherry | 3.61 |
| Teri | 4.00 |
| Todd | 2.78 |

In this option, we can round the GPAs, which is a legitimate thing to do. Take Jim, for example. His GPA is 1.96, which would round to 2.0 . Isaiah's GPA of 1.94 would round down to 1.9 , Amy's GPA of 2.95 would round to 3.0 .


The graphs says " 2 bar 0 means the GPA rounds to 2.0." That refers to Jim, whose GPA rounds to 2.0.

## 2c. Two-digit leaves.

In the same GPA example as above, we can leave the numbers as is and not round. This is also a completely legitimate way to represent this data as long as you visually separate these numbers.

| 4 | 00 |  |  |  |  |  |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| 3 | 12 | 24 | 41 | 55 | 61 | 75 |
| 2 | 23 | 25 | 56 | 78 | 89 | 95 |
| 1 | 94 | 96 |  |  |  |  |

Key: $2 \mid 23$ means GPA is 2.23

The graph's key says, " 2 bar 23 means a GPA of 2.23 ." For example, Tyler's GPA of 3.12 would be represented by the stem 3 and leaf 12.

## 2d. Back-to-Back

Again, using the same GPA data from above, suppose you are interested in the differences between girls' GPAs, like Amy, Holly, Jenny, Katherine, etc., and the boys' GPAs.

You could compare those by putting one group of leaves to the right of the stem and another group of leaves to the left of the stem. This is known as a back-to-back stem-and-leaf plot, and it would look like this:

| Girls |  |  |  | Boys |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | 0 | 4 |  |  |  |  |  |
| 8 | 6 | 0 | 3 | 1 | 2 | 4 | 6 |  |
| 9 | 6 | 2 | 2 | 0 | 3 | 8 |  | Key: $\|3\| 1$ means GPA rounds to 3.1 |

This graph rounds the numbers again, saying that " 3 bar 1 means the GPA rounds to 3.1." Here, the girls' GPAs are on the left. The boys' GPAs are on the right. This allows you to compare the distributions of boys' GPAs to girls' GPAs, illustrating that the girls' GPAs are typically a little bit higher.

## - TERM TO KNOW

## Back-to-Back Stem-and-Leaf Plot

Two stem-and-leaf plots on the same set of stems. This allows us to compare the distributions of two different categories.

## 3. Advantages of Stem-and-Leaf Plots

Why use a stem-and-leaf plot instead of other graphical displays, like histograms or dot plots? Stem-and-leaf plots have a couple of advantages:

- They are like a dot plot in that all the data points can be seen.
- They work over a larger range and all the data points can be seen, making them better than a dot plot in some circumstances.

The drawback to stem-and-leaf plots is that they are difficult to create if the data set is too big. For example, the graph at the beginning had 50 data values, and it might be challenging to see all the data values at once. Therefore, a stem-and-leaf plot might not be the most useful display for that data set.

## ( SUMMARY

Stem-and-leaf plots are very useful displays of quantitative data. There are very versatile, and there are many ways you can use them. To make these plots, you start by creating bins from natural numerical breaks so that the reader can identify the numbers, followed by making a key to convey how the numbers should be interpreted. To make the plot clearer, you can split stems, round the values, create leaves with double digits, or compare across categories using a back to back plot.

## Good luck!

Source: THIS TUTORIAL WAS AUTHORED BY JONATHAN OSTERS FOR SOPHIA LEARNING. PLEASE SEE OUR TERMS OF USE.

TERMS TO KNOW

## Back-to-Back Stem-and-Leaf Plot

Two stem-and-leaf plots on the same set of stems. This allows us to compare the distributions of two different categories.

## Stem-and-Leaf Plot/Stemplot

A distribution of quantitative data that shows natural numerical breaks in the data as categories called "stems" and individual values as "leaves."

