

Graphical Displays For Distributions

- Quantitative Variables

Dot Plots

- Show
 - Shape
 - Center
 - Spread
- Work best when
 - Relatively small number of values
 - Want to see (approximately) individual values
 - Want to see shape
 - Have one group or small number of groups to compare

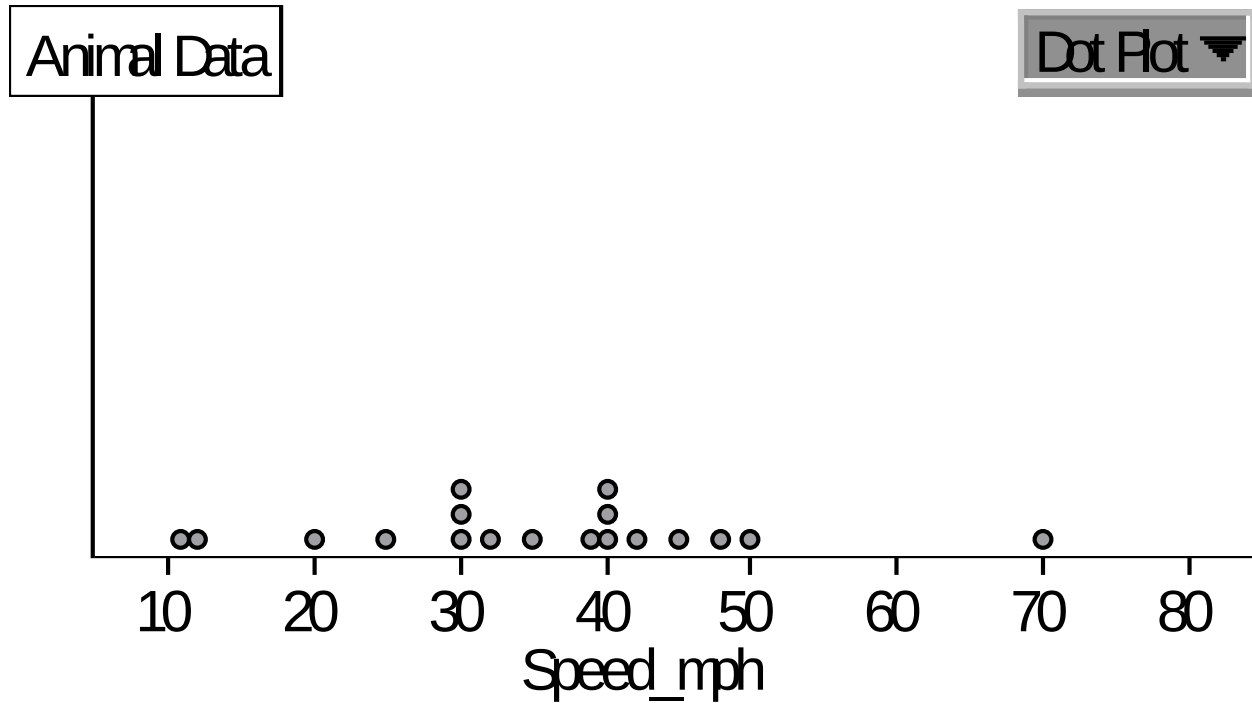
Dot Plots

- **One Axis**
 - **Horizontal Axis**
 - ◉ Label the horizontal axis with the name of the variable
 - **Scale**
 - ◉ Tick marks with numerical labels
 - ▣ Equally spaced
 - ▣ Common measure – for example, by 2's, by 5's, ... whatever is most appropriate for the data
 - **No Vertical Axis**
 - ◉ *Caution:* Some software uses boxes or side lines to define the limitations of the graph.

Dot Plots

- **Data Values**
 - Graphed as dots
 - Line up dots over the appropriate location on the horizontal axis
 - Dots may pile up, one over the other, for repeated data values
- **Problems**
 - If scale is not sufficiently fine, you may not be able to graph the data values
 - If data values too varied, for example between 1 and 200, inclusive, with values such as 156, 157, 158, it may be difficult to graph these points by hand; graph created using software may have distortion
 - Solution: Use a different type of graph!!!

Dot Plot



Histogram

- Shows groups of cases as rectangles or bars
- *A dot plot with bars*
- Height of bar corresponds to the number of cases
- Shows
 - Center, spread, shape
 - Frequency or Relative Frequency on vertical axis

Histogram

- **Work best**
 - With a large number of values to plot
 - Do not need to see individual values exactly
 - Want to see general shape
 - One distribution or small number of distributions to examine
 - When using a computer or calculator

Histogram

- Two Types

- Frequency Histogram
- Relative Frequency Histogram
- Two Axes

- ◉ Vertical axis

- ▣ Frequency or Relative Frequency, in percent

- ▣ Scale

- ◆ Tick marks with numerical labels

- ◆ Equally spaced

- ◆ Common measure – for example by 2's, by 5's, by 10%'s ... whatever is most appropriate

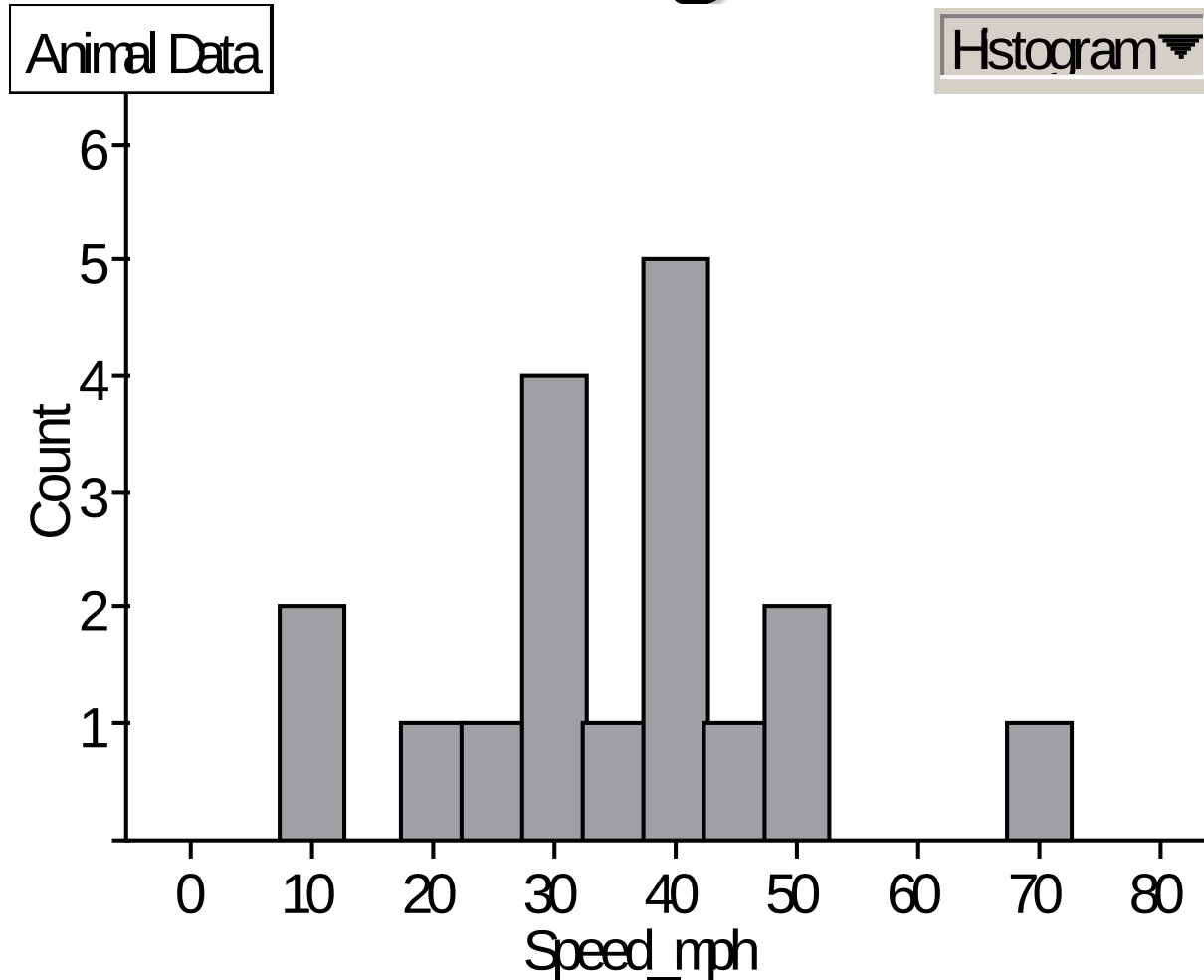
Histogram

- Two Types
 - Frequency Histogram
 - Relative Frequency Histogram
 - Two Axes
 - ◉ Horizontal Axis
 - ▣ The variable that you are analyzing
 - ◆ Axis Label: the name of the variable
 - ▣ Scale
 - ◆ Tick marks with numerical labels
 - ◆ Equally spaced
 - ◆ Common measure that is most appropriate for the *data*

Histogram

- Two Types
 - Frequency Histogram
 - Relative Frequency Histogram
 - Bars
 - ◉ Same width
 - ▣ Width determined by the class width selected when creating the frequency distribution or relative frequency distribution for the quantitative variable

Histogram



Animal Data

[illegible]

Dot Plots and Histograms

- Preserve
 - Shape
 - Center
 - Spread
- DO NOT retain exact values

Stem Plots or Stem and Leaf Plots

- Show
 - Shape
 - Center
 - Spread
- Show exact values for two-digit data are displayed within the graph

Stem Plots or Stem and Leaf Plots

- Show
 - Shape
 - Center
 - Spread
- Show exact values for two-digit data are displayed within the graph

Stem and leaf plots are also known as stemplots.

Making a Stem and Leaf Plot

- What do we do?

Making a Stem and Leaf Plot

- Suppose we examine the following data: 55, 65, 66, 69, 71, 73, 79, 81, 83, 84, 84, 85, 86, 88, 89, 90, and 94

Making a Stem and Leaf Plot

- Stems - the numbers on the left of the bar
- Leaves - the number on the right of the bar

Making a Stem and Leaf Plot

- Stems - the numbers on the left of the bar - the ten's digit
- Leaves - the number on the right of the bar - the one's digit

Making a Stem and Leaf Plot

- For our data,
55, 65, 66, 69, 71, 73, 79, 81,
83, 84, 84, 85, 86, 88, 89, 90,
and 94
- The stems for are 5, 6, 7, 8, and 9 since the data start in the 50's and end in the 90's

Making a Stem and Leaf Plot

- For our data,
55, 65, 66, 69, 71, 73, 79, 81,
83, 84, 84, 85, 86, 88, 89, 90,
and 94
- The stems for are 5, 6, 7, 8, and
9 since the data start in the 50's
and end in the 90's
i.e. the ten's digits are 5, 6, 7, 8,
and 9 for this data

Making a Stem and Leaf Plot

5

6

7

8

9



Making a Stem and Leaf Plot

5

6

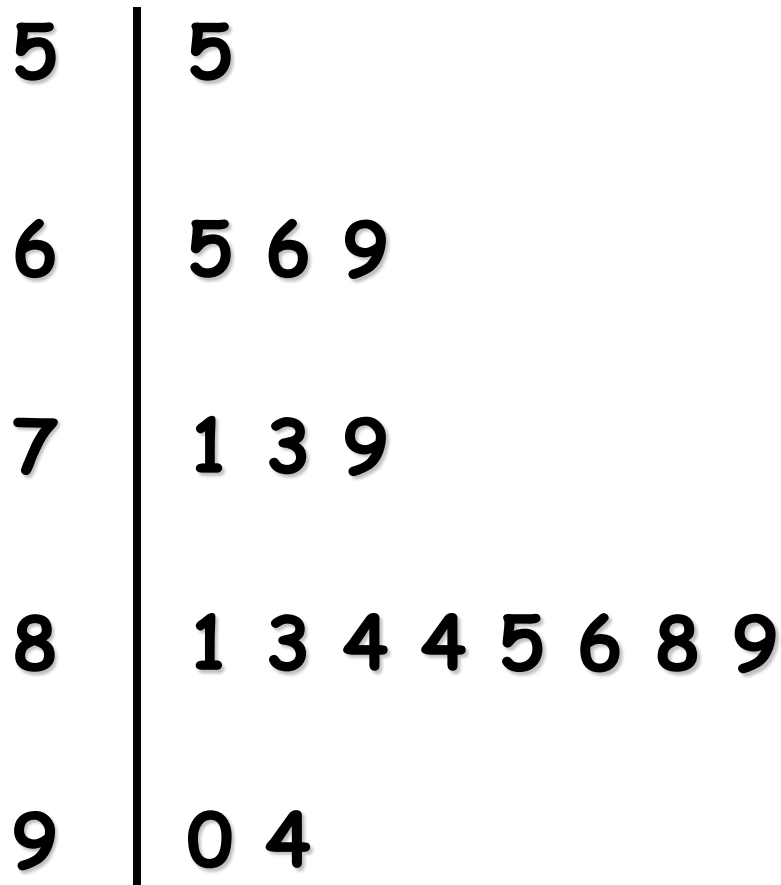
7

8

9

Now, we record the
leaves, the one's digit
for each value

Making a Stem and Leaf Plot

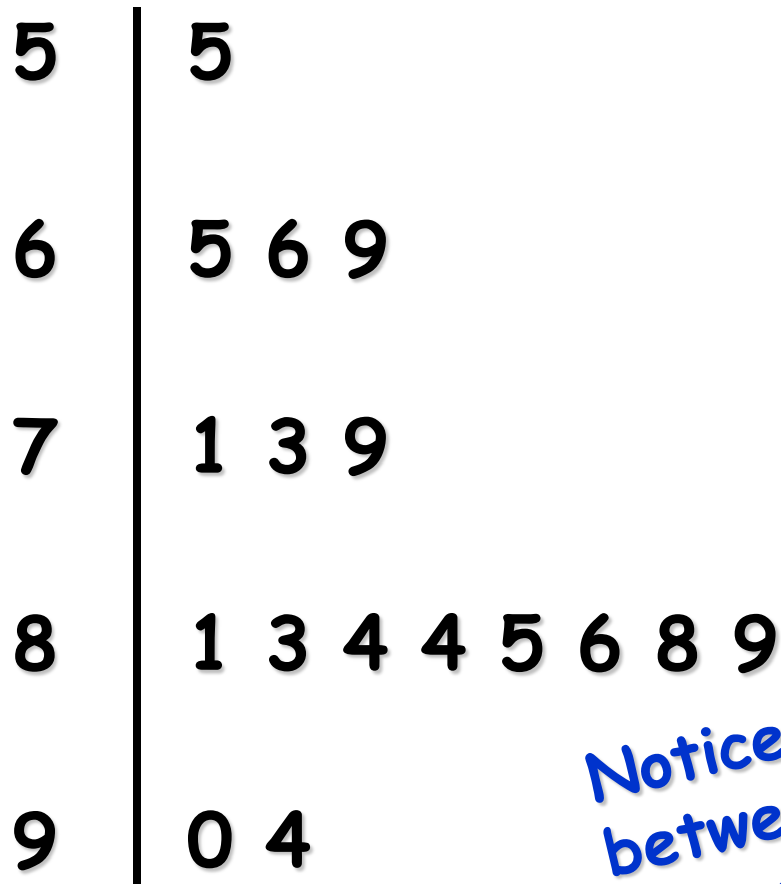


Making a Stem and Leaf Plot

5	5
6	5 6 9
7	1 3 9
8	1 3 4 4 5 6 8 9
9	0 4

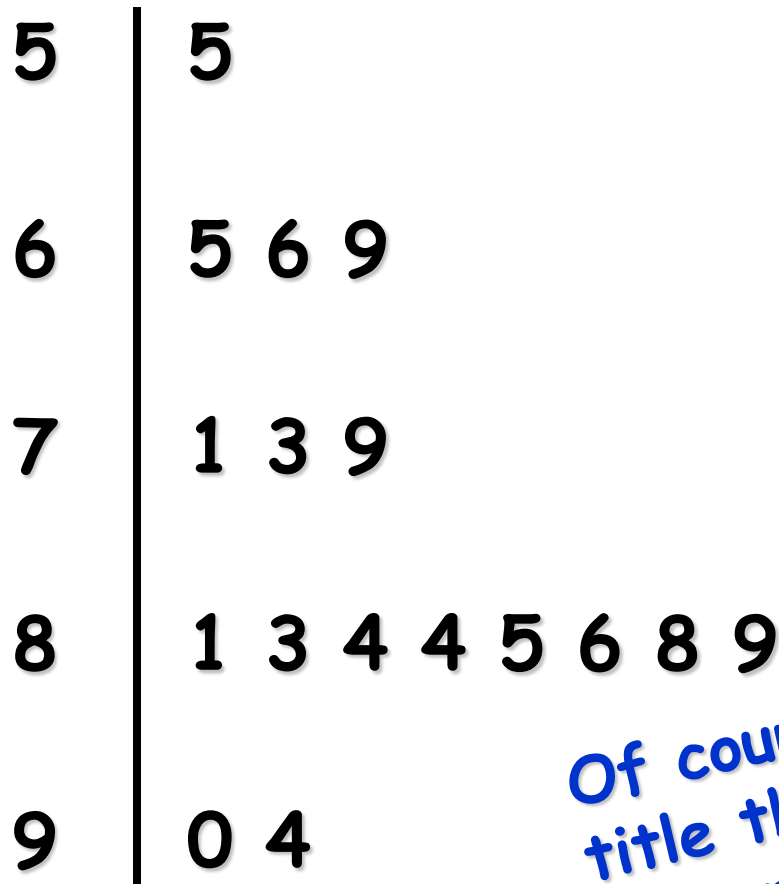
Notice that the one's digits, the leaves, are in numerical order coming out of the stems.

Making a Stem and Leaf Plot



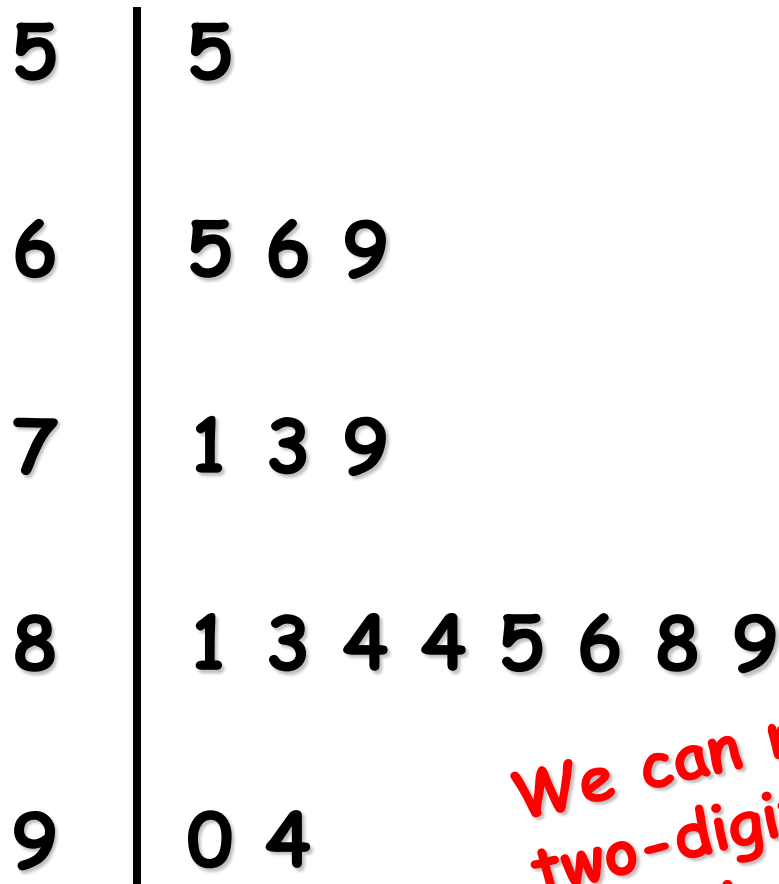
Notice that there are no marks between the leaves and that the leaves are aligned.

Making a Stem and Leaf Plot



Of course, we need an appropriate title that includes the variable, the units, and the context.

Making a Stem and Leaf Plot



We can make stem and leaf plots for two-digit data that involves a decimal point in a similar manner.

Making a Stem and Leaf Plot

5	5
6	5 6 9
7	1 3 9
8	1 3 4 4 5 6 8 9
9	0 4

For such data, the whole numbers would be the stems and the tenths digits would be the leaves.

Stem and Leaf Plots with Split Stems

- Split the stems so that the original stem becomes two stems

Stem and Leaf Plots with Split Stems

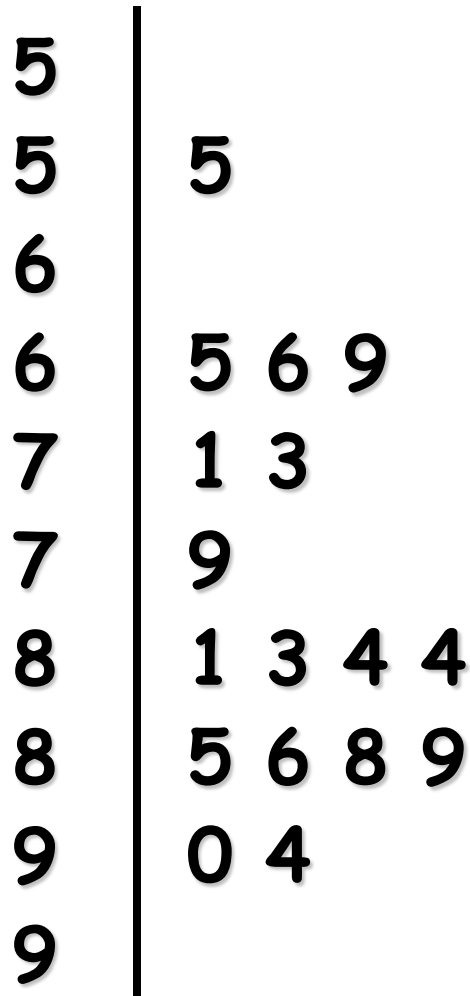
- Split the stems so that the original stem becomes two stems
 - One for the digits 0, 1, 2, 3, 4 - placed on first line of the stem
 - One for digits 5, 6, 7, 8, 9 - placed on second line of the stem

Stem and Leaf Plots with Split Stems

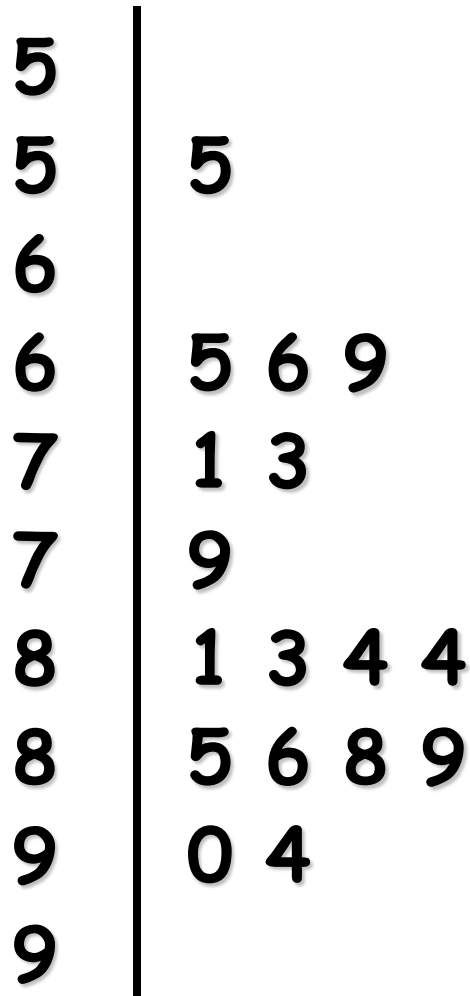
- Split the stems so that the original stem becomes two stems
 - One for the digits 0, 1, 2, 3, 4 - placed on first line of the stem
 - One for digits 5, 6, 7, 8, 9 - placed on second line of the stem

Think of what you do when you round numbers.

Stem and Leaf Plots with Split Stems

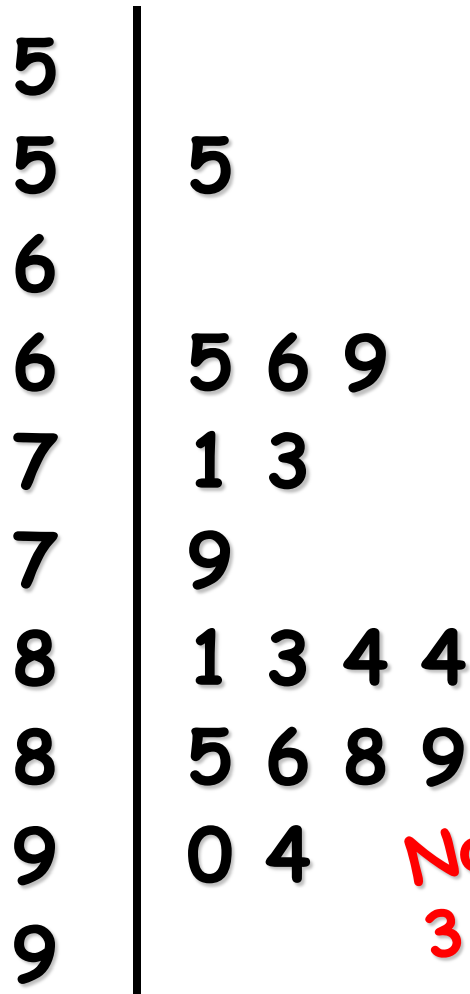


Stem and Leaf Plots with Split Stems



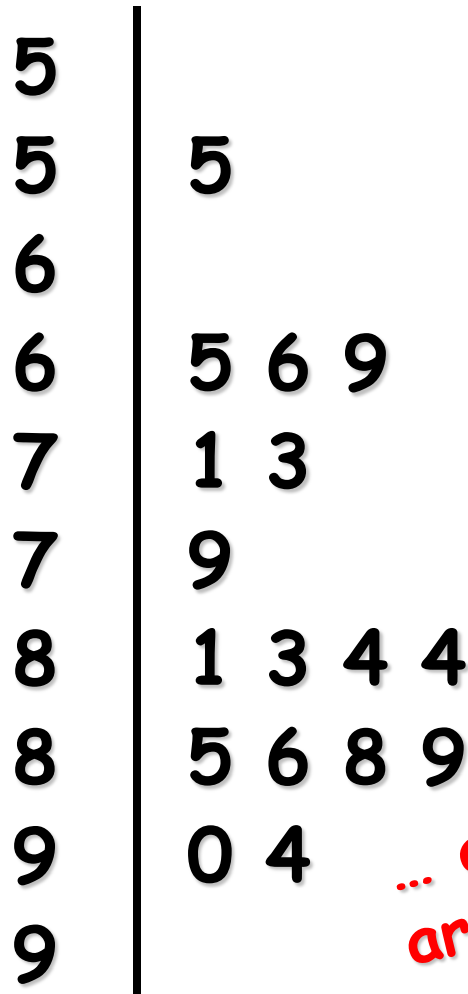
Variation with the double stems
represented using numbers

Stem and Leaf Plots with Split Stems



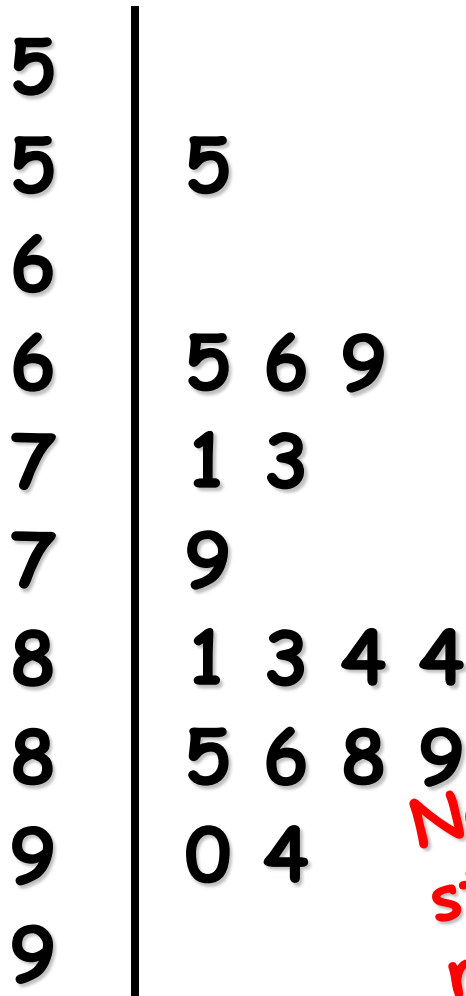
Notice that any leaves that are 0, 1, 2, 3, or 4 are on the first line for the stem

Stem and Leaf Plots with Split Stems



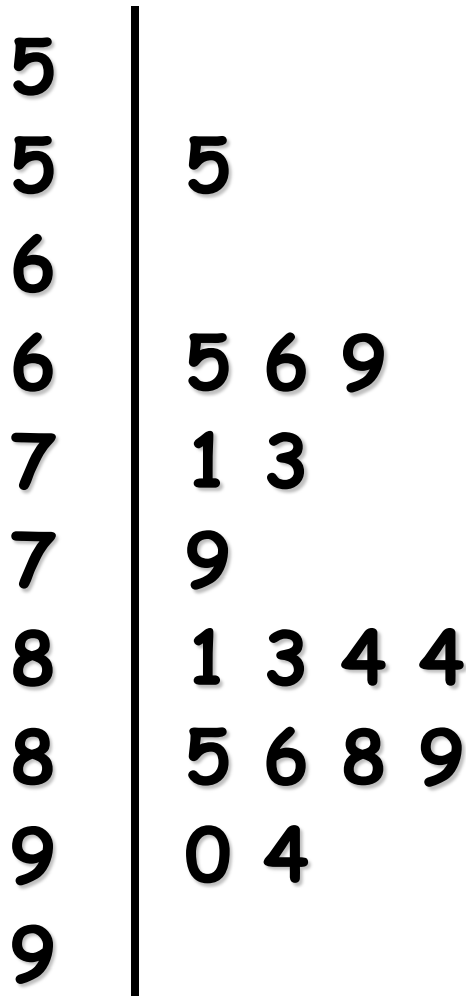
... and any leaves that are 5, 6, 7, 8, or 9
are on the second line for the stem.

Stem and Leaf Plots with Split Stems



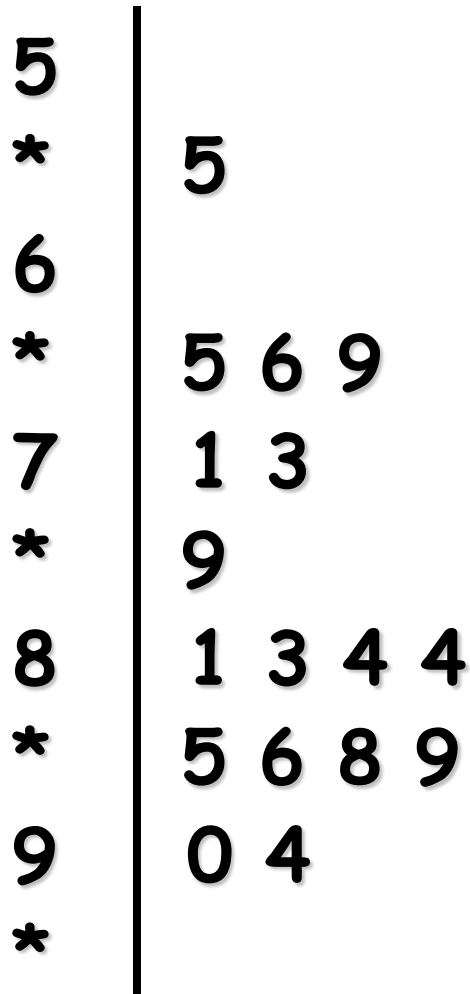
Notice that there are two rows for each stem even when one of these rows does not have any leaves.

Stem and Leaf Plots with Split Stems

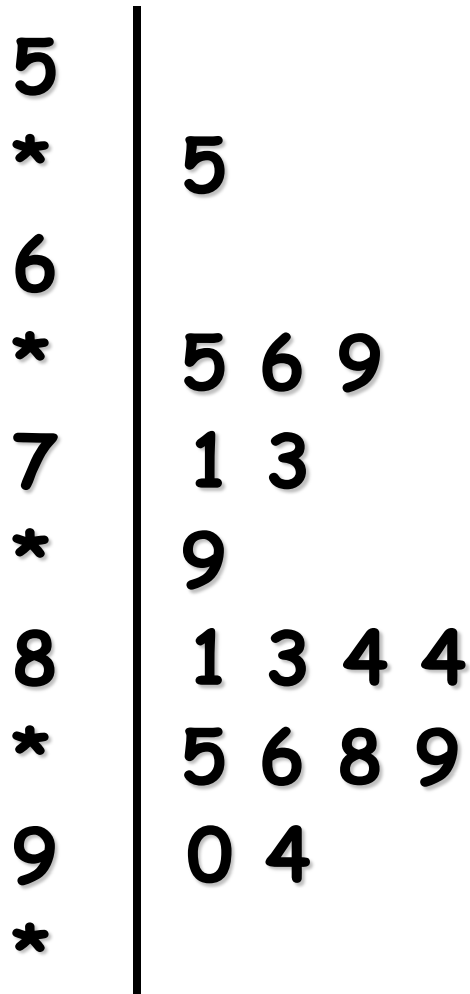


Of course, we need an appropriate title that includes the variable, the units, and the context.

Stem and Leaf Plots with Split Stems

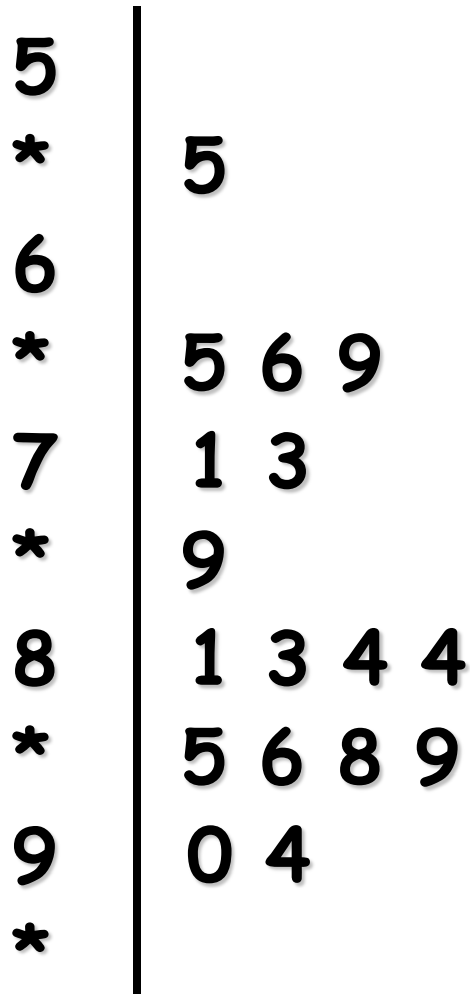


Stem and Leaf Plots with Split Stems



Variation with the second stem
represented using a star

Stem and Leaf Plots with Split Stems



Of course, we need an appropriate title that includes the variable, the units, and the context.

Stem Plots to Compare Distributions

- Back-to-back stem and leaf plots
 - Two distributionsor
 - One distribution that can be separated into two groups using a qualitative variable

Stem Plots to Compare Distributions

- Back-to-back stem and leaf plots
 - Two distributionsor
 - One distribution that can be separated into two groups using a qualitative variable
 - ◉ *Example:* predators and nonpredators for the sample of mammals

Stem Plots to Compare Distributions

- Back-to-back stem and leaf plots
 - For the predators and nonpredators
 - ◉ The leaves for the predators are put on one side (here, the left-hand side)
 - ◉ The leaves for the nonpredators are put on the other side (here, the right-hand side)

Stem Plots to Compare Distributions

- Back-to-back stem and leaf plots
 - For the predators and nonpredators
 - ◉ The leaves for the predators are put on one side (here, the left-hand side)
 - ◉ The leaves for the nonpredators are put on the other side (here, the right-hand side)
 - This allows us to compare the average speed for the predators and nonpredators in the sample.

Back-to-Back Stem and Leaf Plots for the Average Speed, in miles/hour, for the Predators and Nonpredators in a Sample of Mammals

<u>Predator</u>		<u>Nonpredator</u>
	1	12
	2	05
900	3	025
2	4	00058
0	5	
	6	
0	7	

Back-to-Back Stem and Leaf Plots for the Average Speed, in miles/hour, for the Predators and Nonpredators in a Sample of Mammals

<u>Predator</u>		<u>Nonpredator</u>
	1	12
	2	05
900	3	025
2	4	00058
0	5	
	6	
0	7	

Notice that the leaves on both sides of the stems are in numerical order coming out of the stems ...

Back-to-Back Stem and Leaf Plots for the Average Speed, in miles/hour, for the Predators and Nonpredators in a Sample of Mammals

<u>Predator</u>		<u>Nonpredator</u>
	1	12
	2	05
900	3	025
2	4	00058
0	5	
	6	
0	7	

... and there are no marks between the leaves on either side.

Back-to-Back Stem and Leaf Plots for the Average Speed, in miles/hour, for the Predators and Nonpredators in a Sample of Mammals

<u>Predator</u>		<u>Nonpredator</u>
	1	12
	2	05
900	3	025
2	4	00058
0	5	
	6	
0	7	

Notice that the title includes the variable, the units, and the context.

Back-to-Back Stem and Leaf Plots for the Average Speed, in miles/hour, for the Predators and Nonpredators in a Sample of Mammals

<u>Predator</u>		<u>Nonpredator</u>
	1	12
	2	05
900	3	025
2	4	00058
0	5	
	6	
0	7	

We can create back-to-back stem and leaf plots with split stems as well.

Graphical Displays

- Helpful for “viewing” data
- Helpful for estimating the center and spread of data
- Display the shape of distributions
- Useful for classifying distribution

Graphical Displays

- Helpful for “viewing” data
- Helpful for estimating the center and spread of data
- Display the shape of distributions
- Useful for classifying distribution

How do we determine the actual center and spread for the data?