

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!!!

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

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

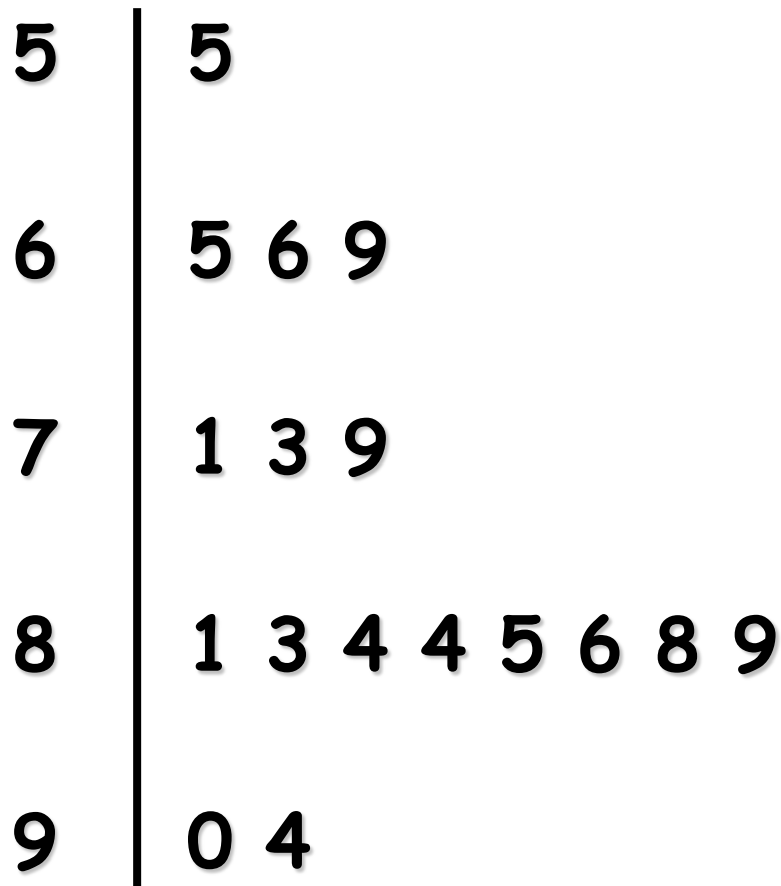
7

8

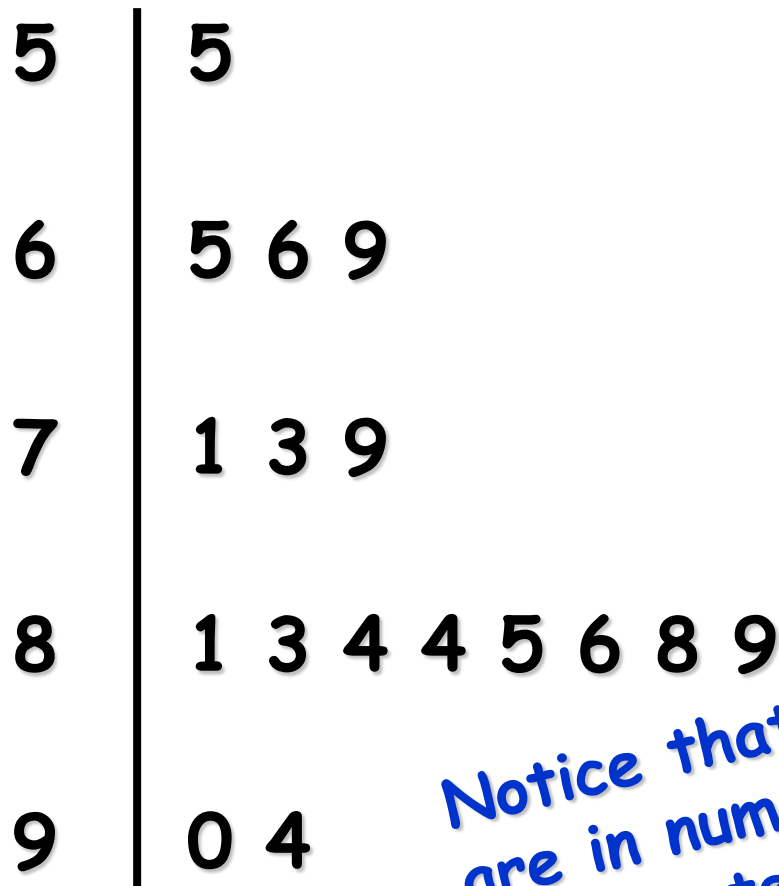
9

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

Making a Stem and Leaf Plot

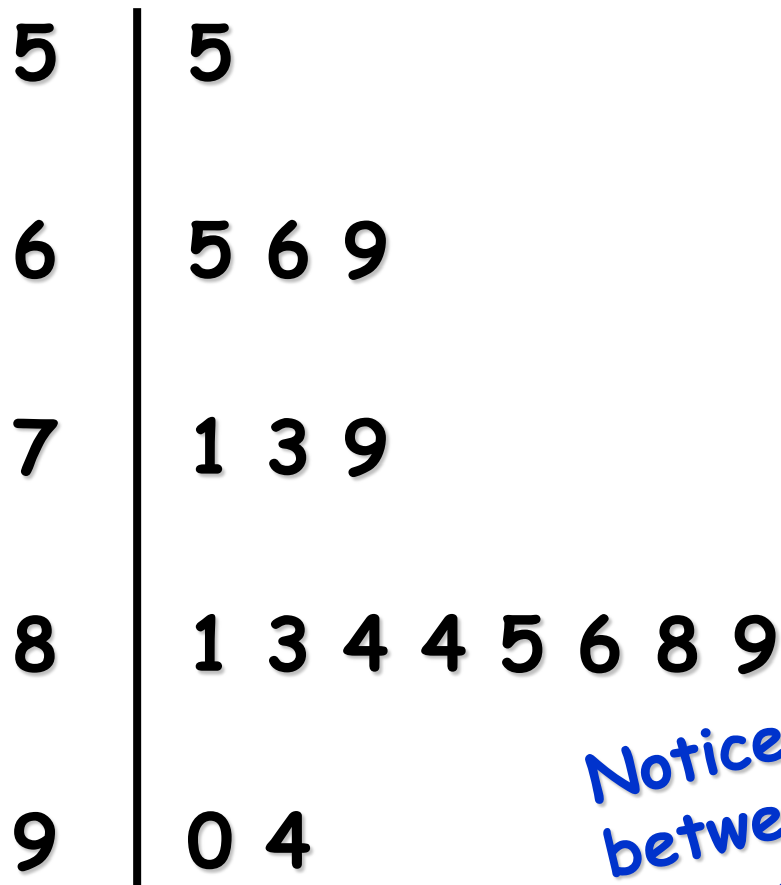


Making a Stem and Leaf Plot



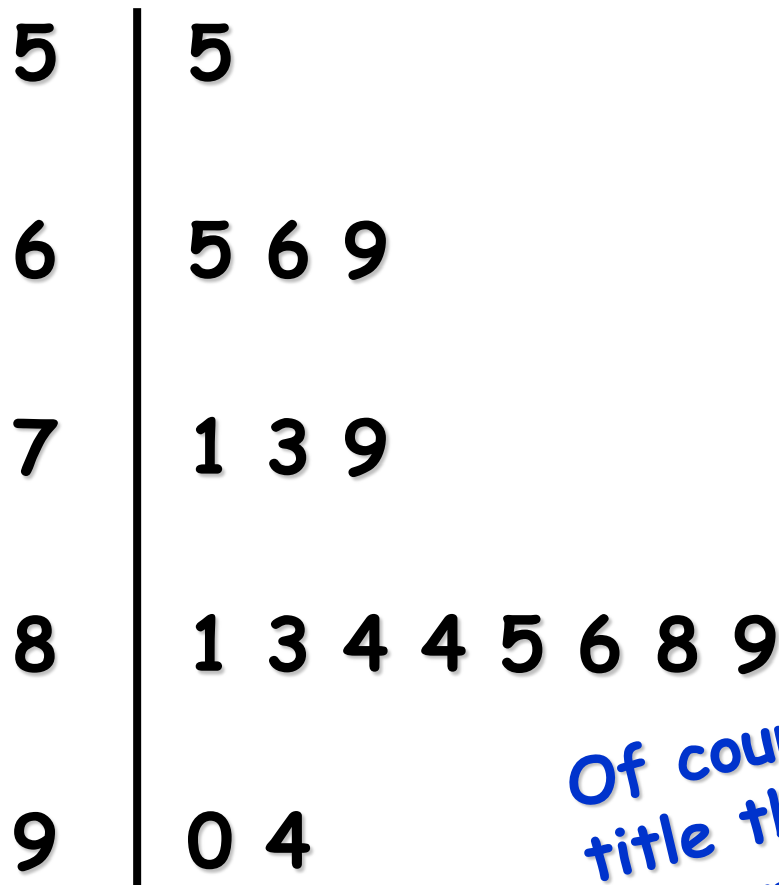
Notice that the one's digits, the leaves, are in numerical order coming out of the stems.

Making a Stem and Leaf Plot



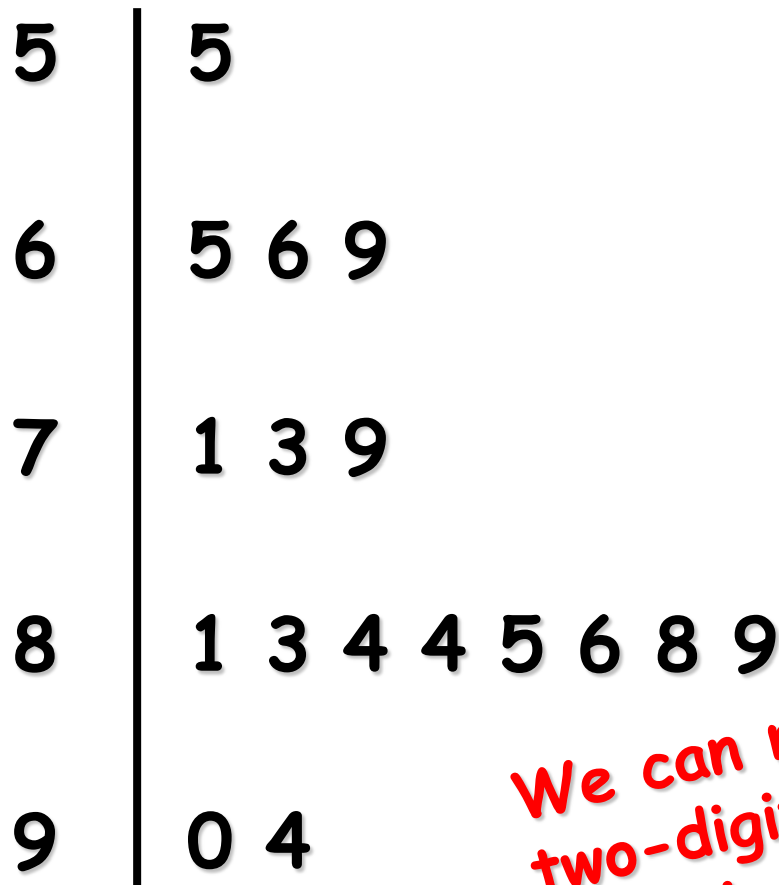
Notice that there are no marks between the leaves and that the leaves are aligned.

Making a Stem and Leaf Plot



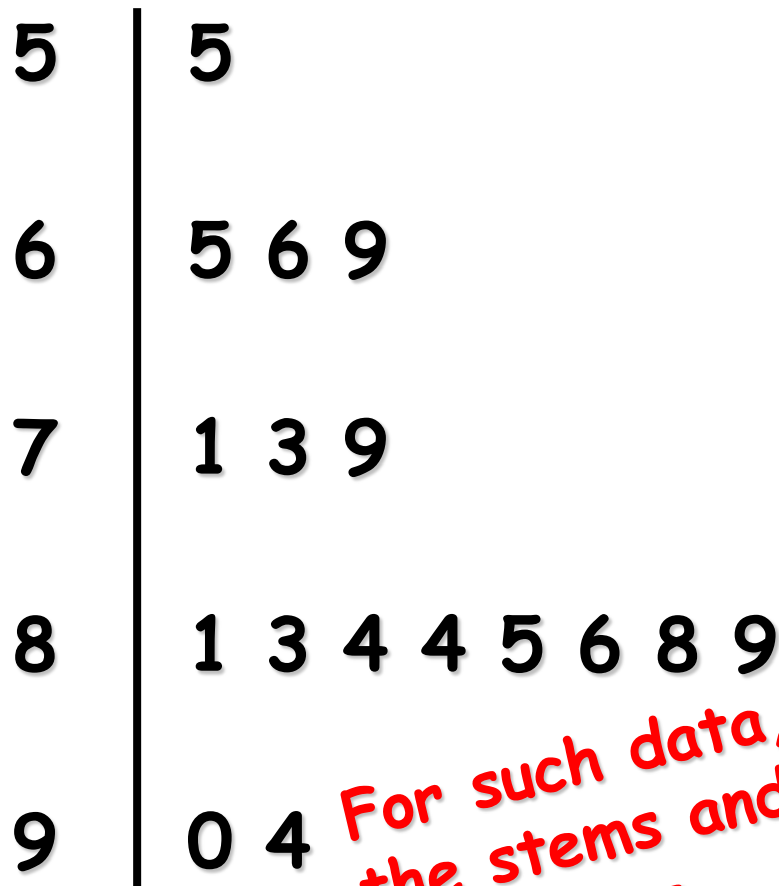
Of course, we need an appropriate title that includes the variable, the units, and the context.

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



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

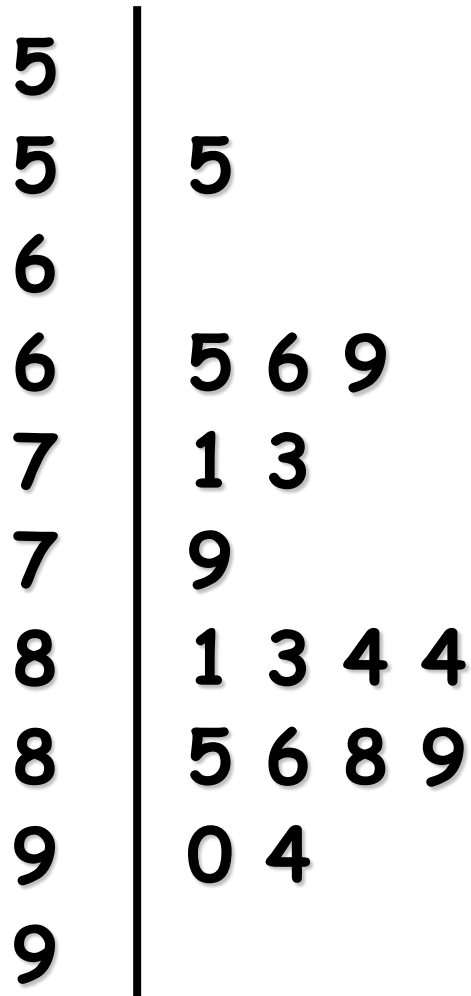
- Split the stems so that the original stem becomes two stems
 - One for the digits 0, 1, 2, 3, 4 - placed on first line of the stem
 - One for digits 5, 6, 7, 8, 9 - placed on second line of the stem

Stem and Leaf Plots with Split Stems

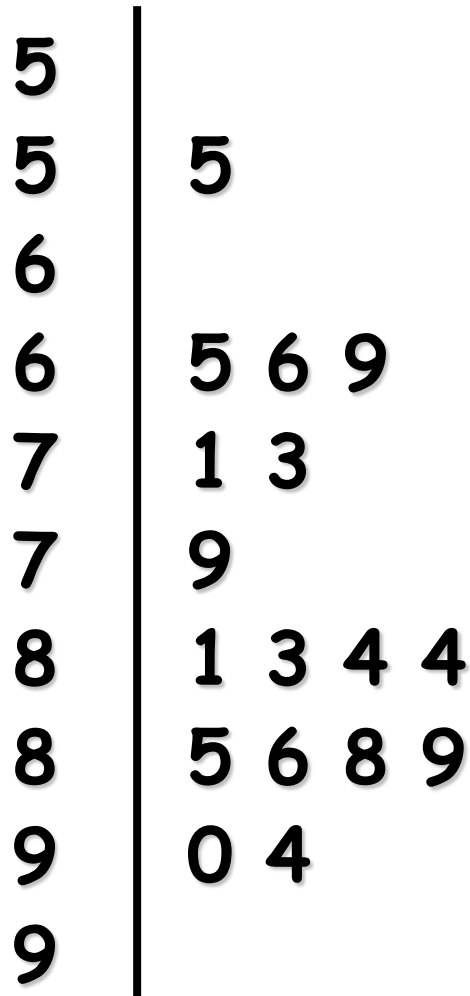
- Split the stems so that the original stem becomes two stems
 - One for the digits 0, 1, 2, 3, 4 - placed on first line of the stem
 - One for digits 5, 6, 7, 8, 9 - placed on second line of the stem

Think of what you do when you round numbers.

Stem and Leaf Plots with Split Stems

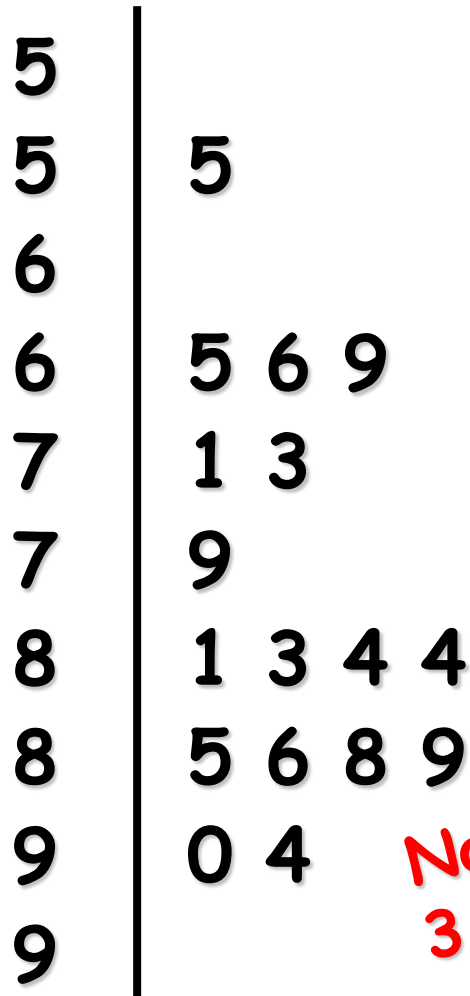


Stem and Leaf Plots with Split Stems



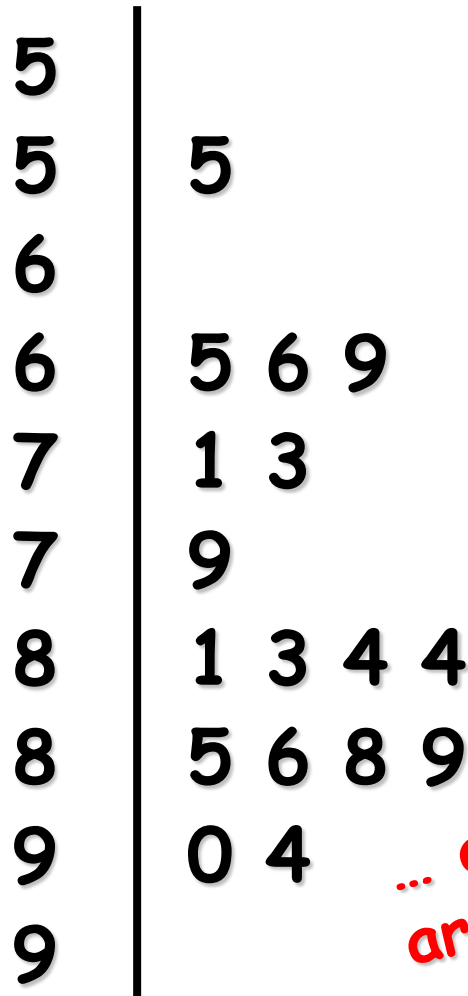
Variation with the double stems represented using numbers

Stem and Leaf Plots with Split Stems



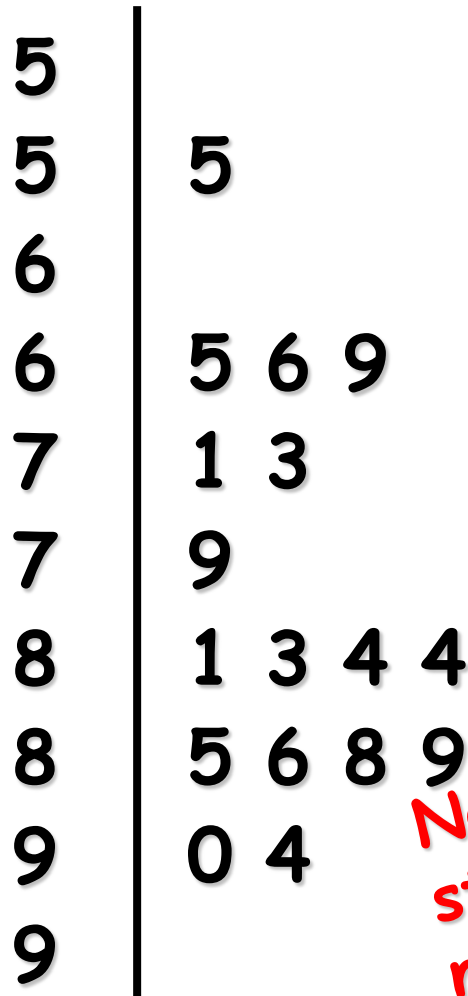
Notice that any leaves that are 0, 1, 2, 3, or 4 are on the first line for the stem

Stem and Leaf Plots with Split Stems



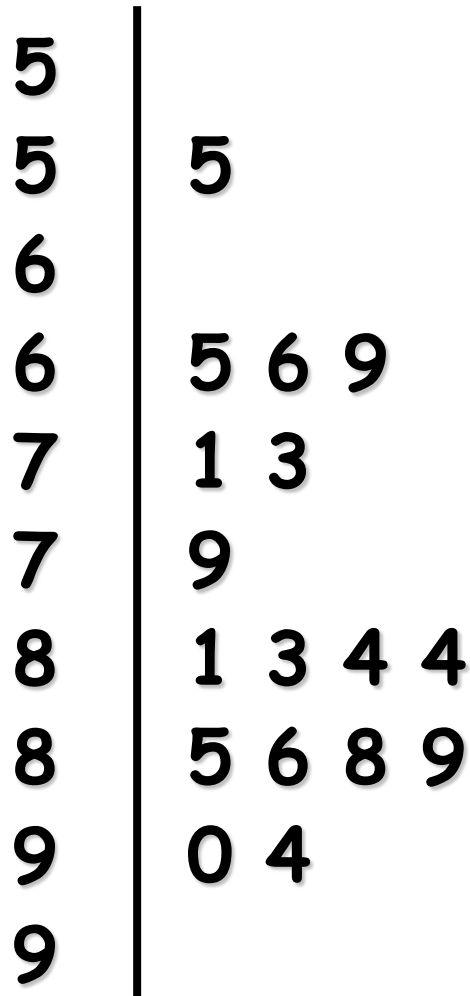
... and any leaves that are 5, 6, 7, 8, or 9 are on the second line for the stem.

Stem and Leaf Plots with Split Stems



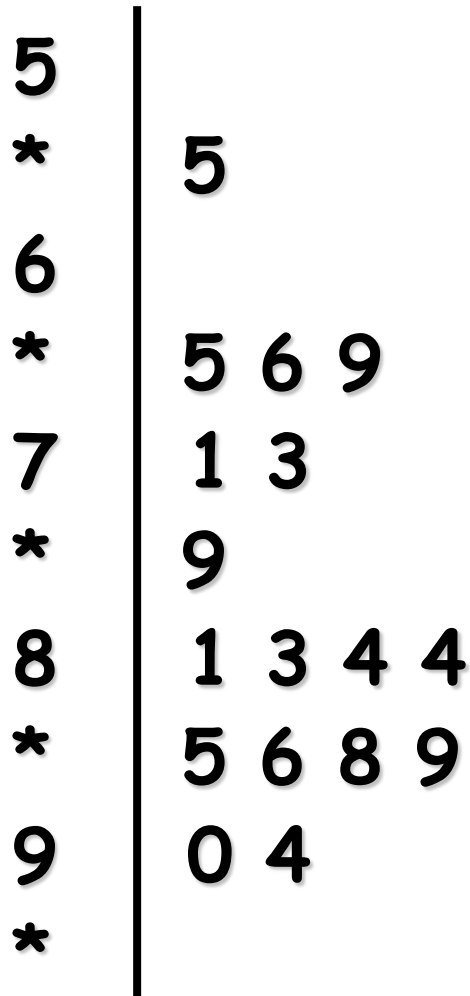
Notice that there are two rows for each stem even when one of these rows does not have any leaves.

Stem and Leaf Plots with Split Stems

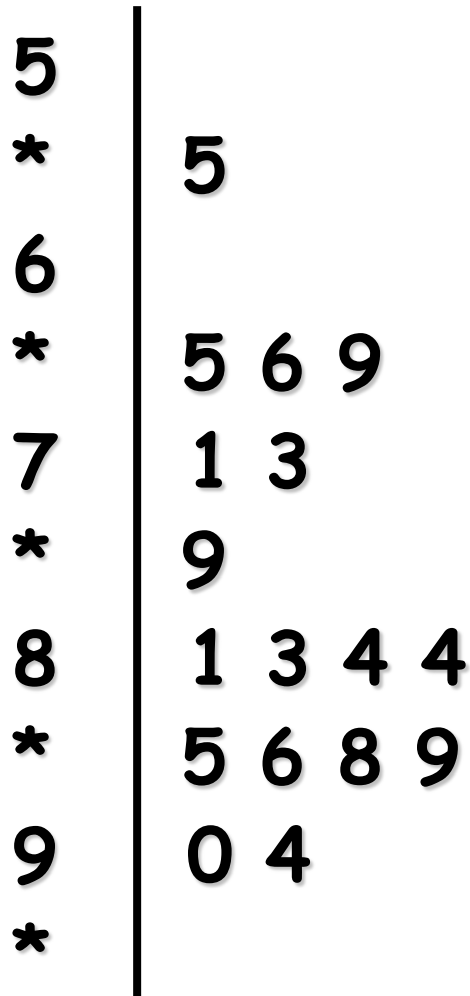


Of course, we need an appropriate title that includes the variable, the units, and the context.

Stem and Leaf Plots with Split Stems

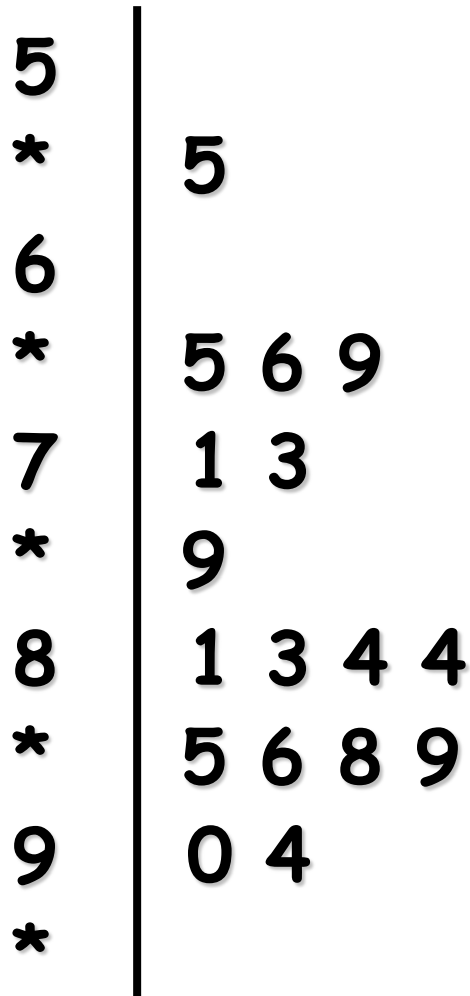


Stem and Leaf Plots with Split Stems



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.

Mammals being studied by the Zoological Society of San Diego

Mammal	Gestation Period, in days	Average Life Span, in years	Maximum Life Span, in years	Average Speed, in miles per hour	Wild (1 if Wild, 0 if not Wild)	Predator (1 if predator, 0 if not Predator)	Wild	Predator
Baboon	187	20	45	*	1	0	yes	no
Grizzly bear	225	25	50	30	1	1	yes	yes
Beaver	105	5	50	*	1	0	yes	no
Bison	285	15	40	*	1	0	yes	no
Camel	406	12	50	*	1	0	yes	no
Cat	63	12	28	30	0	1	no	yes
Cheetah	*	*	14	70	1	1	yes	yes
Chimpanzee	230	20	53	*	1	0	yes	no
Chipmunk	31	6	8	*	1	0	yes	no
Cow	284	15	30	*	0	0	no	no
Deer	201	8	20	30	1	0	yes	no
Dog	61	12	20	39	0	1	no	yes
Donkey	365	12	47	40	0	0	no	no
Elephant	660	35	70	25	1	0	yes	no
Elk	250	15	27	45	1	0	yes	no
Fox	52	7	14	42	1	1	yes	yes
Giraffe	425	10	34	32	1	0	yes	no
Goat	151	8	18	*	0	0	no	no
Gorilla	258	20	54	*	1	0	yes	no
Guinea pig	68	4	8	*	0	0	no	no
Hippopotamus	238	41	54	20	1	0	yes	no
Horse	330	20	50	48	0	0	no	no
Kangaroo	36	7	24	40	1	0	yes	no
Leopard	98	12	23	*	1	1	yes	yes
Lion	100	15	30	50	1	1	yes	yes
Monkey	166	15	37	*	1	0	yes	no
Moose	240	12	27	*	1	0	yes	no
Mouse	21	3	4	*	1	0	yes	no
Opossum	13	1	5	*	1	1	yes	yes
Pig	112	10	27	11	0	0	no	no
Puma	90	12	20	*	1	1	yes	yes
Rabbit	31	5	13	35	0	0	no	no
Rhinoceros	450	15	45	*	1	0	yes	no
Sea lion	350	12	30	*	1	1	yes	yes
Sheep	154	12	20	*	0	0	no	no
Squirrel	44	10	23	12	1	0	yes	no
Tiger	105	16	26	*	1	1	yes	yes
Wolf	63	5	13	*	1	1	yes	yes
Zebra	365	15	50	40	1	0	yes	no

* Data not available

Back-to-Back Stem and Leaf Plots for the Average Speed, in miles/hour, for the Mammals, Classified as Predator or Non-Predator, being Studied by the Zoological Society of San Diego

<u>Predator</u>		<u>Non-Predator</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 Mammals, Classified as Predator or Non-Predator, being Studied by the Zoological Society of San Diego

<u>Predator</u>		<u>Non-Predator</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 Mammals, Classified as Predator or Non-Predator, being Studied by the Zoological Society of San Diego

<u>Predator</u>		<u>Non-Predator</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 Mammals, Classified as Predator or Non-Predator, being Studied by the Zoological Society of San Diego

<u>Predator</u>		<u>Non-Predator</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 Mammals, Classified as Predator or Non-Predator, being Studied by the Zoological Society of San Diego

<u>Predator</u>		<u>Non-Predator</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	

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?