

# **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**

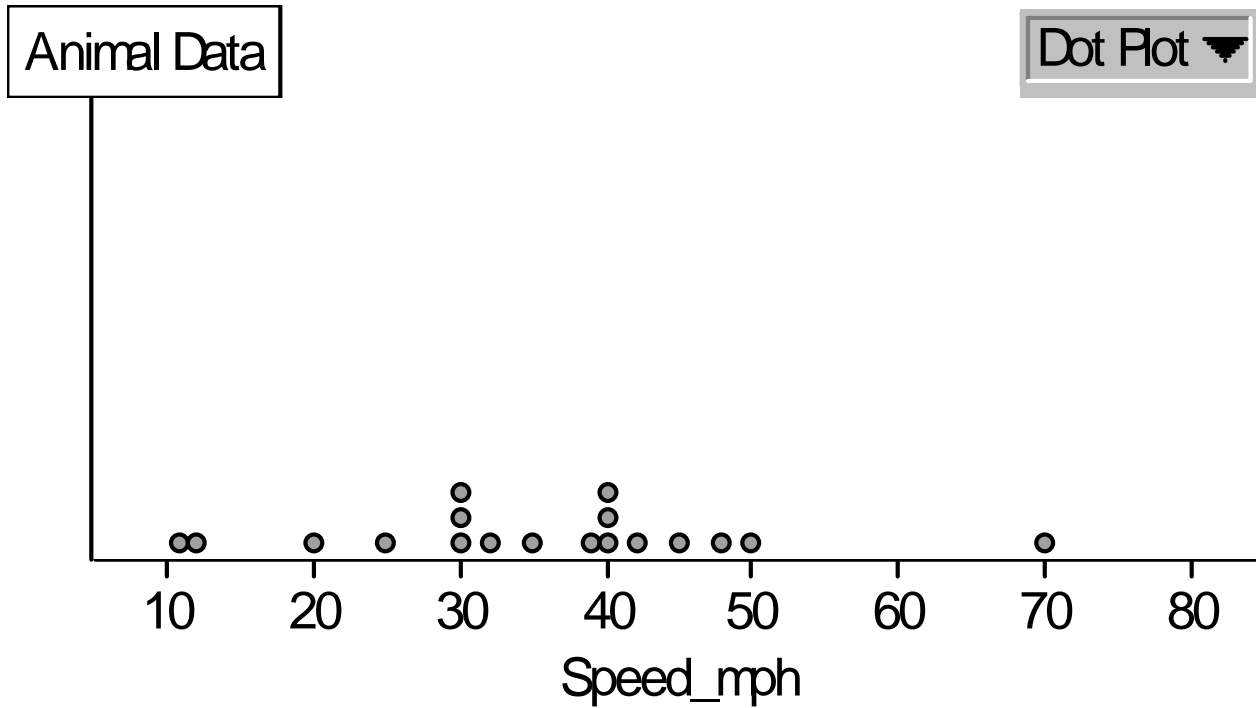
# Dot Plots

- **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

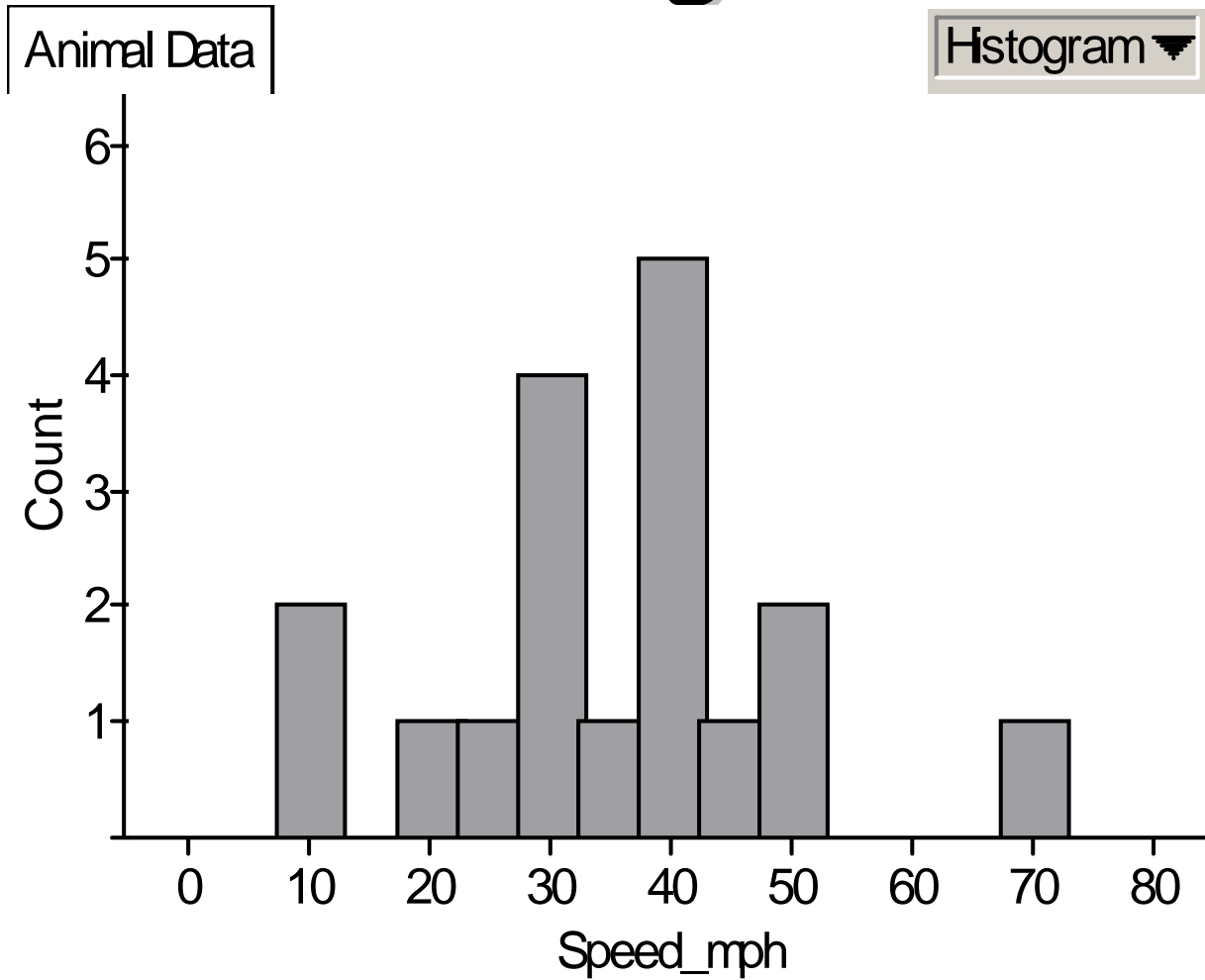
# Histogram

- 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





# Dot Plots and Histograms

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

# Stemplots or Stem-and-Leaf Plot

- **Show**
  - **Shape**
  - **Center**
  - **Spread**
- **Show exact values**

# Making a Stem-and-Leaf Plot (Stemplot)

- What do we do?

# Making a Stem-and-Leaf Plot (Stemplot)

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

# Making a Stem-and-Leaf Plot (Stemplot)

- 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 (Stemplot)

- Suppose we examine the following data: 55, 65, 66, 69, 71, 73, 79, 81, 83, 84, 84, 85, 86, 88, 89, 90, and 94
- The stems for these data 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 (Stemplot)



# Making a Stem-and-Leaf Plot (Stemplot)

5

6

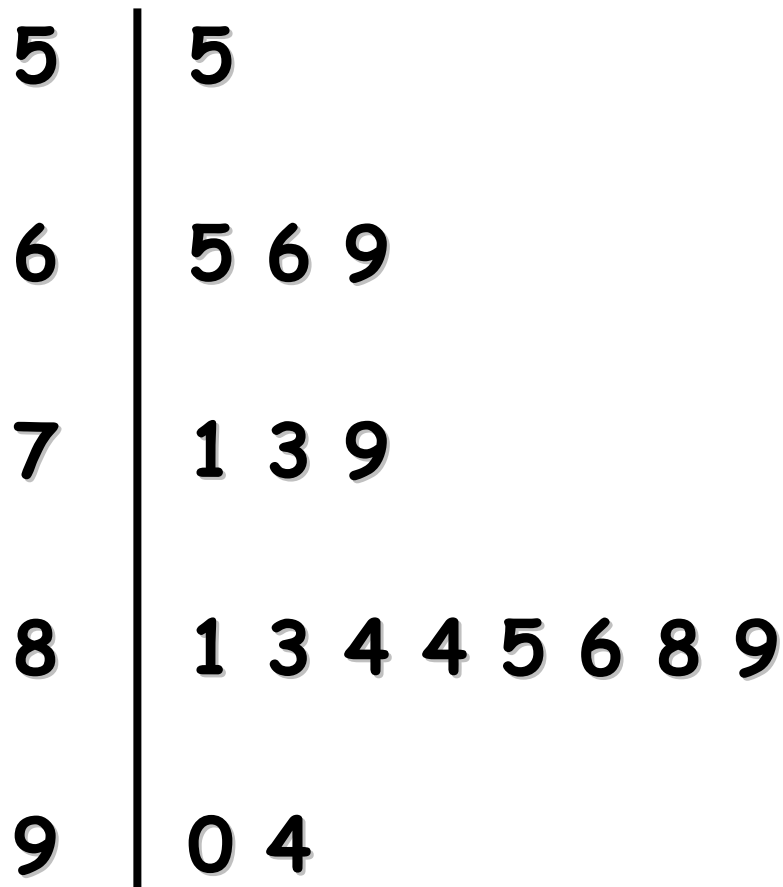
7

8

9

Now, we record the leaves, the ones digit for each value

# Making a Stem-and-Leaf Plot (Stemplot)



# 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.*

# Stemplots to compare distributions

- **Back-to-back stemplots**

- **Two distributions**

**or**

- **One distribution that can be separated into two groups using a qualitative variable**

# Stemplots to compare distributions

- **Back-to-back stemplots**

- **Two distributions**

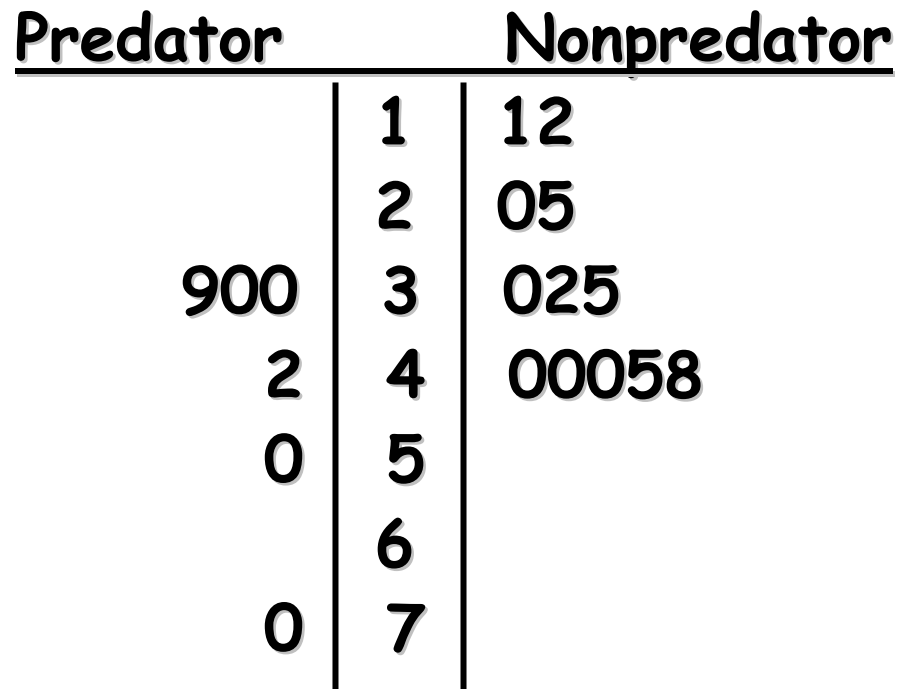
or

- **One distribution that can be separated into two groups using a qualitative variable**

- *Example:* predators and nonpredators for the sample of mammals

# Stemplot for the Speeds of the Predators and the Nonpredators

- What does this stemplot tell you about the distributions?



# 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?