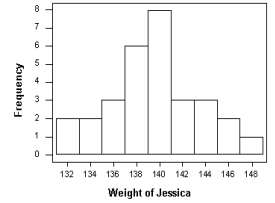


Chapter 4: Displaying Quantitative Data *The MOST relevant information is bolded, colored, or starred.*

Graphing Quantitative Data!

*Histogram

- Similar to a bar graph but the bars are not spaced apart.
- Histograms are best used with large data sets.
- Relative frequency histograms use the percentages of value that fall into ranges.



Stem-and-Leaf Display

- Shows the individual values and is better for small data
- There is a stem (leading digits) and the leaves (the trailing digits)
- Example: 63, 6 is the stem and 3 is the leaf

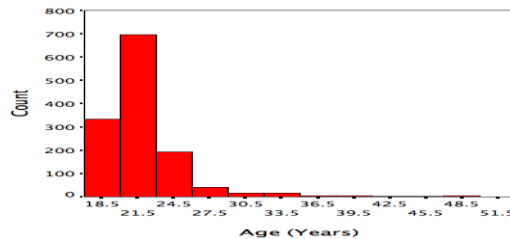
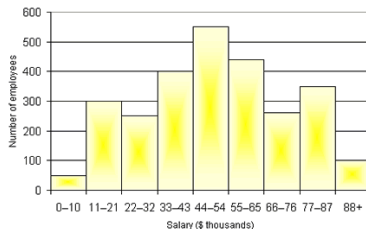
| Step 1 | Step 2 | Step 3 | | | | | | | | | | | | | | | |
|--|--|--|-------|--|--|------|------|--|---|---|---|---|---|-----|---|---|---|
| Make the stem by writing the tens digits from least to greatest. | Make the leaves by writing each ones digit in order to the right of its ten digit. | Draw a line to separate the stems and leaves. Add a title and key. | | | | | | | | | | | | | | | |
| 4 5 6 | 4 8 9 5 4 5 8 6 0 3 | <table border="1"> <thead> <tr> <th colspan="3">Title</th> </tr> <tr> <th>Stem</th> <th colspan="2">Leaf</th> </tr> </thead> <tbody> <tr> <td>4</td> <td>8</td> <td>9</td> </tr> <tr> <td>5</td> <td>4</td> <td>5 8</td> </tr> <tr> <td>6</td> <td>0</td> <td>3</td> </tr> </tbody> </table> | Title | | | Stem | Leaf | | 4 | 8 | 9 | 5 | 4 | 5 8 | 6 | 0 | 3 |
| Title | | | | | | | | | | | | | | | | | |
| Stem | Leaf | | | | | | | | | | | | | | | | |
| 4 | 8 | 9 | | | | | | | | | | | | | | | |
| 5 | 4 | 5 8 | | | | | | | | | | | | | | | |
| 6 | 0 | 3 | | | | | | | | | | | | | | | |
| | | Key: 4 8 means 48 inches | | | | | | | | | | | | | | | |

- Dotplot uses dots for each case.
- MAKE SURE scales are consistent and to use appropriate labels. Use **QUANTITATIVE data**.

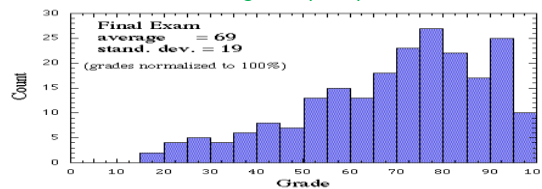
*How does one analyze and compare these histograms and stem-and-leaf displays?

Shape

- Uniform= no peaks, unimodal=one peak, bimodal=two peaks, and multimodal= three or more peaks
- Symmetric vs. Skewed
- Symmetric
- Skewed Right the long tail stretches farther right



Skewed Left the long tail (end) stretches farther left



- Look for **outliers** that stray far from the rest of the data and **gaps** in the display where there are no values.

Center

- The center is a rough estimate of a typical data value.
- If it's unimodal and symmetric then the center is the center of the display.
- If it's skewed data then the center is where the data set can be split in half, 50-50, to contain the same amount of values on both sides. This center would roughly represent the median.

Spread

- Describes the variation in the data, how tightly the values cluster around the "center."
- More variation means that values are less "predictable."
- Standard deviation, range, outliers, etc.

Timeplots represent a change over time, a trend in the data.

It's easier to summarize symmetric data. When given skewed data try to re-express or transform it by using log or square root re-expression to make data symmetric.