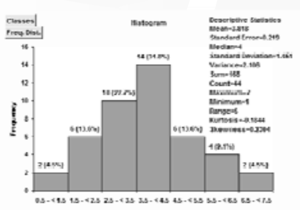
Chapter 4:
Displaying and Summarizing Quantitative Data

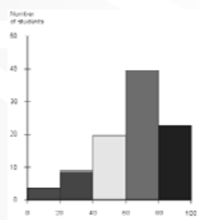
Histogram

- Histograms allow a visual interpretation of **quantitative (numerical) data** by indicating the number of data points that lie within a range of values, called a class, width or a bin. The frequency of the data that falls in each class is depicted by the use of a bar.

Graphs for *Quantitative Data: Histograms*

- ***Histogram***
  - Breaks the range of values of a variable into ***classes*** and displays only the **count** or **percent** (***relative frequency histogram***) of the observations that fall into each class.





Graphs for *Quantitative Data: Histograms*

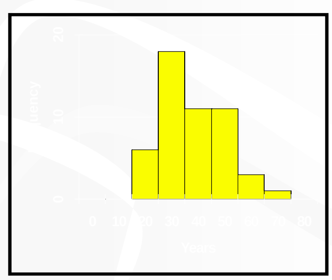
- ***Histogram***
  - ***a "bar graph" in which the horizontal scale represents classes and the vertical scale represents frequencies***
  - Data points cannot be seen on the plot
  - For large quantity of data points, group nearby values
  - The bins and the counts in each bin give the distribution of the quantitative variable. Your calculator will give you a bin width, but you may need to make adjustments to get a better display. The heights of the bins are plotted. Shape, Center and Spread are important.

Graphs for *Quantitative Data: Histograms*

- ***Construction Method:***
  - Draw a horizontal axis that covers the full range of values for the variable
  - Decide bar width (also called class width) so that 5 to 10 bars will cover the full range of data
  - Set borders for bars, count frequencies, draw bars

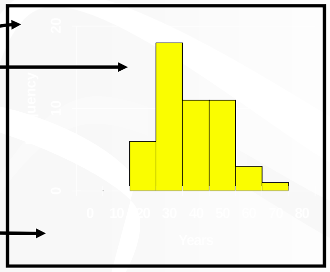
Histogram

Class Interval	***Frequency***
20-under 30	**6**
30-under 40	**18**
40-under 50	**11**
50-under 60	**11**
60-under 70	**3**
70-under 80	**1**



Histogram

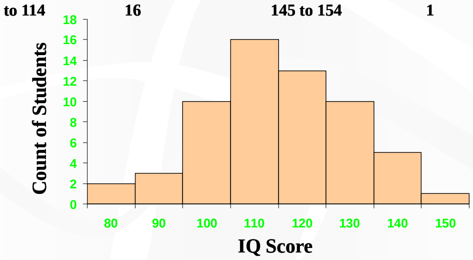
Class Interval	*Frequency*
20-under 30	6
30-under 40	18
40-under 50	11
50-under 60	11
60-under 70	3
70-under 80	1



Histogram

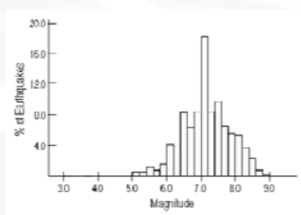
• *Example*

Class Interval	*Frequency*	*Class Interval*	*Frequency*
75 to 84	2	115 to 124	13
85 to 94	3	125 to 134	10
95 to 104	10	135 to 144	5
105 to 114	16	145 to 154	1



Histograms: Displaying the Distribution of Earthquake Magnitudes

- A **relative frequency histogram** displays the ***percentage*** of cases in each bin instead of the count.
  - In this way, relative frequency histograms are faithful to the area principle.
- Here is a relative frequency histogram of earthquake magnitudes:

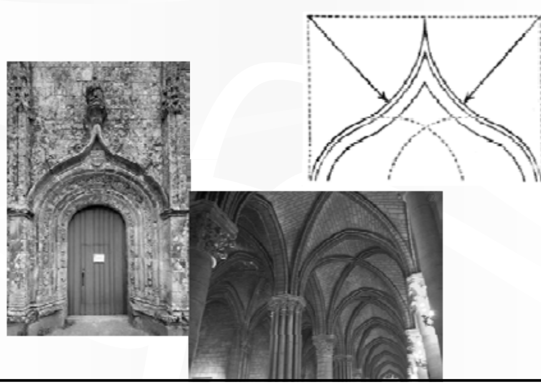


Relative Frequency and Cumulative Frequency

- ***Tells about the relative standing of an individual***
  - Construct a relative cumulative frequency histogram (ogive--*pronounced "oh jive"*)
  - Decide on class intervals and make a frequency table. Add three columns: relative frequency, cumulative frequency, and relative cumulative frequency.
  - Complete the table.

$$Relative\ frequency = \frac{frequency}{total\ frequency}$$

Ogives - *Ogival arches in Gothic Architecture*



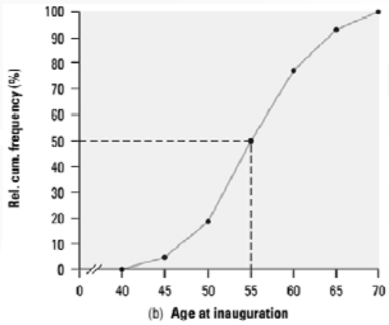
Relative Cumulative Frequency

• *Example*

Class	Freq.	Rel. Freq.	Cum. Freq.	Rel. Cum. Freq.
40-44	2	4.7%	2	4.7%
45-49	6	14.0%	8	18.6%
50-54	13	30.2%	21	48.8%
55-59	12	27.9%	33	76.7%
60-64	7	16.3%	40	93.0%
65-69	3	7.0%	43	100.0%
TOTAL	43	100.0%		

Relative Cumulative Frequency

- ***Example***



Graphs for *Quantitative Data*

- ***Stemplot (stem-and-leaf plot)*** ***Stem | leaf***
  - Organizes and groups data.
  - Make each observation into a ***stem***, consisting of all but the final (right-most) digit, and a ***leaf***, the final digit. Stems may have as many digits as needed, but ***each leaf contains only a single digit***.
  - Write the stems in a vertical column with the smallest at the top, and draw a vertical line at the right of this column.
  - Write each leaf in a row to the right of the stem, in increasing order out from the stem.
  - Label to include magnitude or decimal point.

Graphs for *Quantitative Data*

- ***Stemplot (stem-and-leaf plot)***
  - When the display is a little crowded, split each line (stem) into two bars.

Example: Pulse rates

8	**0 0 0 0 4 4 8**
7	**2 2 2 2 6 6 6 6**
6	**0 4 4 4 8 8 8 8**
5	**6**

Pulse Rate

5 | 6 means 56 beats/min

8	**8**
8	**0 0 0 0 4 4**
7	**6 6 6 6**
7	**2 2 2 2**
6	**8 8 8 8**
6	**0 4 4 4**
5	**6**

Pulse Rate

5 | 6 means 56 beats/min

Graphs for *Quantitative Data*

- ***Stemplot (stem-and-leaf plot)***
  - For numbers with three or more digits, you'll often decide to truncate (or round) the number to two places, using the first digit as the stem and the second as the leaf.

Example: 432, 540, 571, and 638

(indicate 6|3 as 630-639)

6	**3**
5	**4 7**
4	**3**

Stemplot

First data value = 35

stems		
	2	
	3	5 ← **leaf**
	4	
	5	
	6	
	7	

Raw Data:

35, 45, 42, 45, 41, 32, 25, 56, 67, 76, 65, 53, 53, 32, 34, 47, 43, 31

Stemplot

2	5
3	5 2 2 4 1
4	5 2 5 1 7 3
5	6 3 3
6	7 5
7	6

Raw Data:

35, 45, 42, 45, 41, 32, 25, 56, 67, 76, 65, 53, 53, 32, 34, 47, 43, 31

2 | 5 indicates 25 years

- Use **stemplots** for **small to fairly moderate** sizes of data (25 – 100)
- Try to use **graph paper**
  *(or **make sure** that your numbers **line up**)*

3	4 **(this is okay…)**		3	4 **(this is NOT)**
4	1 1 5 9		4	1159
5	2 3 6		5	236
6	0 1 7 8		6	0178
7	0 0 1 1 1 7		7	001117
8	4 4		8	44

Stemplot

- ***Example (DO THIS~!)***

2	**9**
3	**1 8**
4	**6**
5	
6	**0 3 3**
7	**1 1 0 8**
8	**6 2 5**
9	**9 9 9**

Raw Data:

60, 31, 46, 71, 86, 99, 82, 71, 85, 38, 70, 63, 99, 63, 78, 99, 29

2 | 9 indicates 29 percent

Stemplot

- ***Example (DO THIS~!)***

2	**9**
3	**1 8**
4	**6**
5	
6	**0 3 3**
7	**0 1 1 8**
8	**2 5 6**
9	**9 9 9**

Raw Data:

60, 31, 46, 71, 86, 99, 82, 71, 85, 38, 70, 63, 99, 63, 78, 99, 29

2 | 9 indicates 29 percent

Stemplot

- ***Example***
  - The overall pattern of stemplot is irregular, as is often the case when there are only few observations. There do appear to be two ***clusters*** of countries. For example, why do the three central Asian countries (Kazakhstan, Tajikistan, and Uzbekistan) have very high literacy rates?

Graphs for *Quantitative Data*

- ***Back-to-back stemplot***
  - comparing two related distributions
  - the leaves on each sides are ordered out from the ***common stem***.
- ***Literacy is generally higher among males than among females in these countries.***

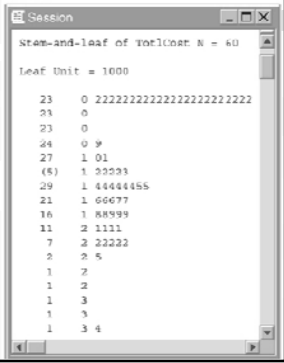
Female		Male
9	2	
8 1	3	
6	4	
	5	0
3 3 0	6	8 8
8 1 1 0	7	0 8
6 5 2	8	3 4 5 9
9 9 9	9	2 2 4 5 6
	10	0 0 0

Stemplot

- ***Stemplots do not work well for large data sets where each stem must hold large number of leaves.***
  - To plot the distribution of a moderate number of observations, double the number of stems in a plot by ***splitting stems*** into two: one with leaves 0 to 4 and the other with leaves 5 through 9.
  - When the observed values have many digits, ***trimming*** the numbers by removing the last digit or digits before making a stemplot is often best.
  - Use your judgment in deciding whether to split stems and whether to trim.
  - *Remember*: the purpose of a stemplot is to display the shape of the distribution.

Stemplot

- ***Example***
  - Stemplot of tuitions and fees for 60 colleges and universities in Virginia, made in Minitab.
- **Leaf unit: 1000**
- **$34,850 means $34,000.**
- **Minitab has truncated the last three digits, leaving 34 thousand.**
- **this is called "*trimming*"**


Minitab

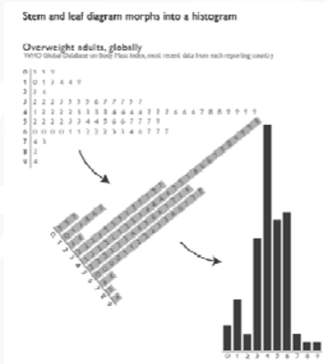
```
Session
Stem-and-leaf of TotlCost N = 60

Leaf Unit = 1000

   23    0 2222222222222222222222
   23    0
   23    0
   24    0 9
   27    1 01
  (5)    1 22233
   29    1 44444455
   21    1 66677
   16    1 88999
   11    2 1111
    7    2 22222
    2    2 5
    1    2
    1    2
    1    3
    1    3
    1    3 4
```

Stemplot

- ***Limitations:***
  - Stemplots display the actual values of the observations which makes stemplots awkward for large data sets.
  - The picture presented by a stemplot divides the observations into groups (stems) determined by the number system rather than by judgment.

Stemplot: Final Note

- A stem-and-leaf plot is really just a ***"sideways histogram"***



Stemplot: Final Note

- A stem-and-leaf plot is really just a ***"sideways histogram"***
  - The choice of stems is like choosing bar widths. The ***stem*** consists of all but the rightmost digit and the ***leaf***, the final digit. *Ex*. 20 mg (2 is the stem and 0 is the leaf)
  - Leaves should be arranged from least to greatest on each line
    - That may mean doing the plot twice: a "draft" and final copy
  - Start by drawing your vertical line and filling in stems
  - If less than 5, "***split the stems***"
  - Too few stems will result in a skycraper-shaped plot, while too many stems will yield a very flat "pancake" graph. Five stems is a good minimum
  - Provide a ***key*** to your coding

Horsepower of cars reviewed by *Consumer Reports:*

155	103	130	80	65
142	125	129	71	69
125	115	138	68	78
150	133	135	90	97
68	105	88	115	110
95	85	109	115	71
97	110	65	90	
75	120	80	70	

15	05
14	2
13	0358
12	0559
11	00555
10	359
9	00577
8	0058
7	01158
6	55889

6|5 = 65 horsepower

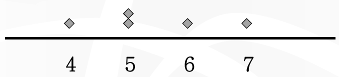
(not always necessary to use split stems)

Visual representation of Quantitative Data: *Dotplots*

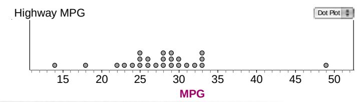
- The most basic method is a **dotplot**
  - Every data point can be seen on the plot
- *Construction method:*
  - Draw a horizontal axis (number line) that covers the full range of values for the variable and label it with the variable name. *(usually there is **no** vertical axis)*
  - Scale and number the axis—look for the min and max values
  - Put a dot on the axis for each data point
  - If data duplicate, stack them vertically

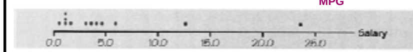
Visual representation of Quantitative Data: *Dotplots*

- ***Example:*** Construct the dotplot for the set 4, 5, 5, 7, 6

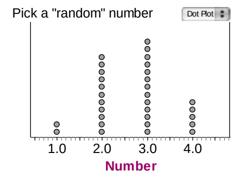


Dotplots



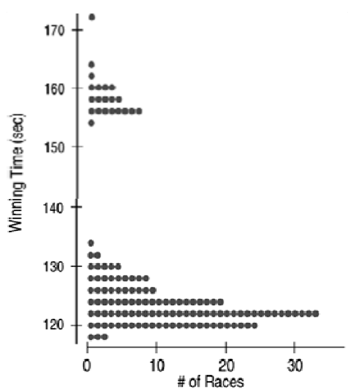


Salaries of Los Angeles Lakers players, 2002-03, in millions of dollars

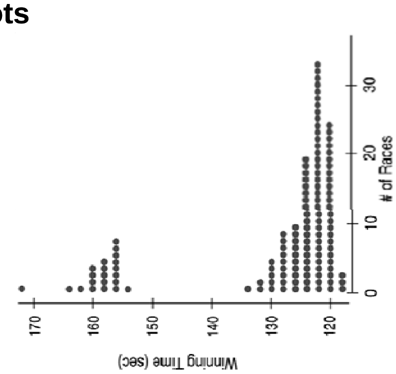


Dot plots work well for **relatively small** data sets **(50 or less)**

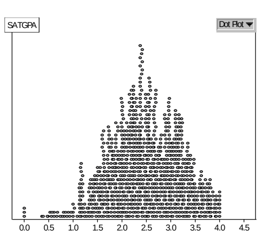
Dotplots



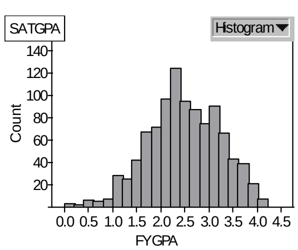
Dotplots



What's wrong with this picture?!!



Too much data for a dot plot!



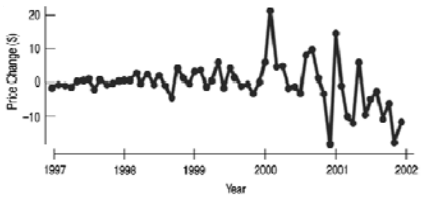
The histogram works much better!

Time Plots

- Data sets composed of similar measurements taken at ***regular intervals over time***
- Shows data values in chronological order
- Place ***time*** on ***horizontal*** scale
- Place the variable being measured on the vertical scale
- ***Connect*** data points with line segments

Timeplots: Order, Please!

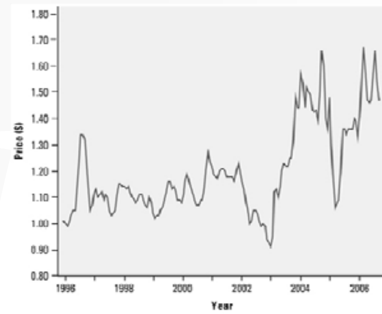
- For some data sets, we are interested in how the data behave over time. In these cases, we construct **timeplots** of the data.



Time Plot

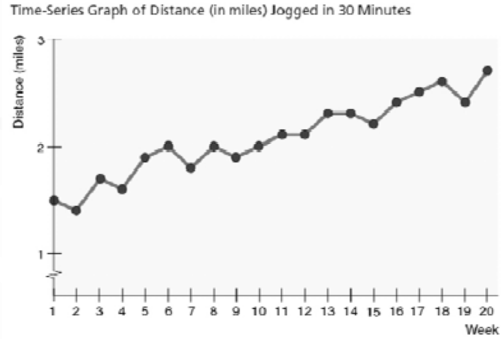
- ***Example***

Observation?



Time Plot

Time-Series Graph of Distance (in miles) Jogged in 30 Minutes



Rules For Any Graph

- **Provide a title.**
- **Label axes.**
- **Identify units of measure.**
- **Present information clearly.**

Shape, Outlier, Center, and Spread (SOCS)

- **When describing a distribution, make sure to always tell about three things: *shape, outlier/unusual feature, center,* and *spread…***

What is the Shape of the Distribution?

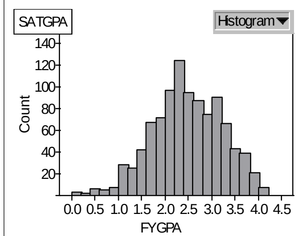
1. Does the graph of the data (histogram) have a single, central hump or several separated humps?
2. Is the histogram symmetric?
3. Do any unusual features stick out?

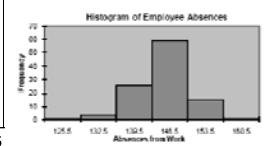
Shape of a Distribution: Humps

- Does the histogram have a single, central hump or several separated bumps?
  - Humps in a histogram are called **modes**.
  - A histogram with one main peak is dubbed **unimodal**; histograms with two peaks are **bimodal**; histograms with three or more peaks are called **multimodal**.

Shape of a Distribution

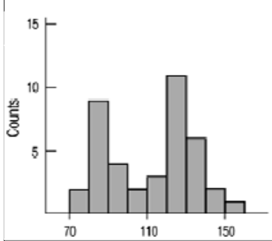
- **Unimodal**
  - **One peak value ("hump") that occurs more frequently than the rest**





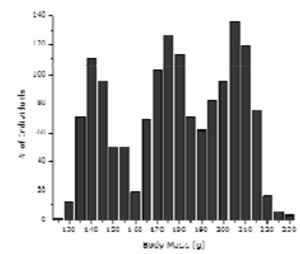
Shape of a Distribution

- **Bimodal**
  - **Two peak values that occur more frequently than the rest**



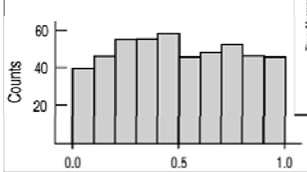
Shape of a Distribution

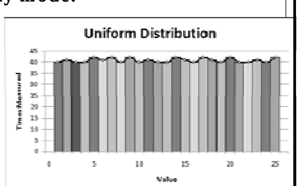
- **Multimodal**
  - **Three or more peak values**



Shape of a Distribution

- **Uniform**
  - **Bars in histogram are all about the same height. It doesn't appear to have any mode.**





Shape

Is the histogram **symmetric**?

ALWAYS say "**approximately** symmetric" or "**roughly** symmetric"

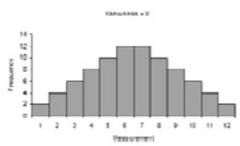
(unless it truly is ***perfectly*** *symmetric)*

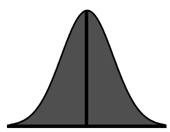
A symmetric histogram ...

...can fold in the middle so that the two sides almost match.

Symmetry

- Does the data look symmetric relative to the middle?
  - Can you fold it along a vertical line through the middle and have the edges match pretty closely, or are more of the values on one side? Does the distribution of the left half look like the right half?



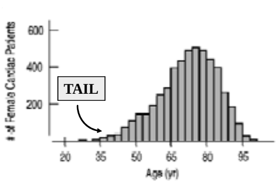


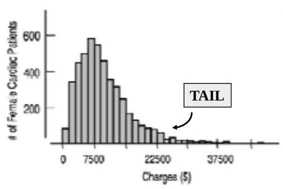
Symmetry

- Is the data skewed?
  - Are there tails on the data that stretch out away from the center?
  - Skewed to the Left: tail is on the left
    Skewed to the Right: tail is on the right

Skewed to the left/right

The thinner ends of a distribution are called **tails.**

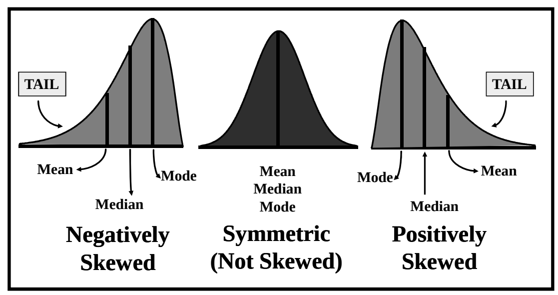




Skewed to the left (to the **lower "numbers"**)

Skewed to the right (to the **higher "numbers"**)

Symmetry



Where is the Center of the Distribution?

- If you had to pick a single number to describe all the data what would you pick?
- It's easy to find the center when a histogram is unimodal and symmetric—it's right in the middle.
- On the other hand, it's not so easy to find the center of a skewed histogram or a histogram with more than one mode.

The Measures of Central Tendency

- Mean
- Median
- Mode

Mean

- The **mean** of a data set is the average of all the data values.
- If the data are from a ***sample***, the mean is denoted by $\bar{x}$

$$\bar{x} = \frac{\sum x_i}{n}$$

- If the data are are from a **population**, the mean is denoted by μ (mu).

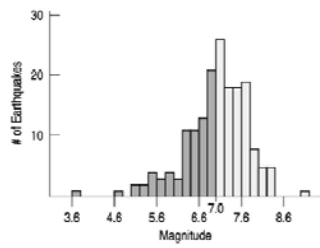
$$\mu = \frac{\sum x_i}{N}$$

Median

- It is the value in the middle when the data items are arranged in ascending order (**Q_2** or **M**).
- It is insensitive to extreme scores or skewed distribution.
- It is the 'middle point' in a distribution. Middle value in ordered sequence
  - If odd ***n***, middle value of sequence.
  - If even ***n***, average of 2 middle values.
- It is the measure of location most often reported for annual income and property value data.
- A few extremely large incomes or property values can inflate the mean but not the median.

Median

- The median is the value with exactly half the data values below it and half above it. It is the middle data value (once the data values have been ordered) that divides the histogram into two equal areas.



Mean vs. Median

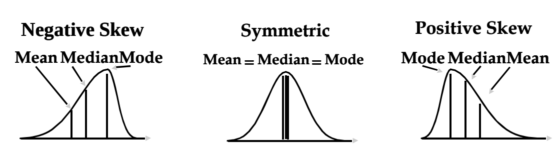
- The ***mean*** and the ***median*** are the most common measures of center.
  - If a distribution is perfectly symmetric, the *mean* and the *median* are the same.
  - The ***mean*** is ***not resistant to outliers***.
- *You* must decide which number is the most appropriate description of the center…

Mode

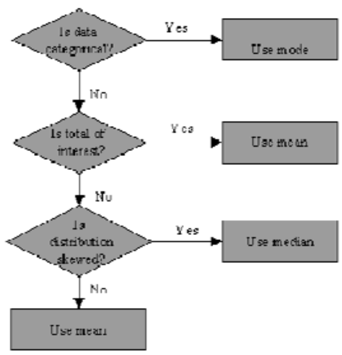
- It is the value that occurs most often (with greatest frequency).
- Not affected by extreme values.
- The greatest frequency can occur at two or more different values.
- May be no mode or several modes.
- If the data have exactly two modes, the data are <u>bimodal</u>.
- If the data have more than two modes, the data are <u>multimodal</u>.
- May be used for quantitative & qualitative data

Which average?

Mean	Median	Mode
• not appropriate for describing highly skewed distributions • not appropriate for describing nominal and ordinal data	• choose median when mean is inappropriate, except when describing nominal data	• choose mode when describing nominal data. However, for nominal data, an average may not be needed (use percentage instead)



Selecting Among the Mean, Median, and Model



How Spread out is the Distribution?

- Variation matters, and Statistics is about variation. Without variability, there would be no need for the subject ☺.
  - When describing data, **never** rely on center alone.
- Are the values of the distribution tightly clustered around the center or more spread out?
- Always report a measure of **spread** (or **variation**) along with a measure of center when describing a distribution numerically.

Measures of Spread (Variability)

- Measures of variability "***describe the spread or the dispersion of a set of data***."
- Common Measures of Variability
  - Range
  - Interquartile Range (IQR)
  - Variance
  - Standard Deviation
- Like measures of Center, ***you*** must choose the most appropriate measure of spread.

The Range

- The **range** of a data set is the difference between the largest and smallest data values.
- It is the **simplest measure** of variability.
- It is **very sensitive** to the smallest and largest data values.
- A disadvantage of the range is that a single extreme value can make it very large and, thus, not representative of the data overall.

Example:

Range = Largest – Smallest
= 48 - 35
= 13

35	41	44	45
37	41	44	46
37	43	44	46
39	43	44	46
40	43	44	46
40	43	45	48

Quartiles

- **Quartiles** divide the data into four equal sections.
  - Q_1 : 25% of the data is set below the first quartile (also the 25th percentile).
  - Q_2 : 50% of the data is set below the second quartile (this is also 50th percentile and the median).
  - Q_3 : 75% of the data is set below the third quartile (also the 75th percentile).
- The quartiles border the middle half of the data.
- Quartile values are not necessarily members of the data set.

Quartiles

- **To findQ1 and Q3, order data from min to max.**
- **Determine the median, if necessary.**
- **The first quartile is the middle of the 'bottom half'.**
- **The third quartile is the middle of the 'top half'.**

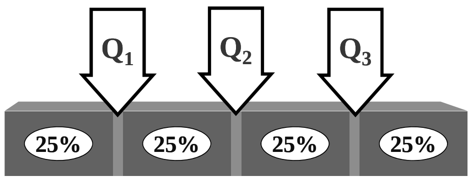
19	22	23	23	23	26	26	27	28	29	30	31	32

Q1=23 med Q3=29.5

45	68	74	75	76	82	82	91	93	98

Q1 med=79 Q3

Quartiles



Example: Quartiles

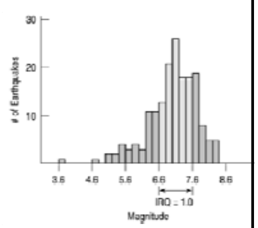
Ordered array: 106, 109, 114, 116, 121, 122, 125, 129

- Q_1: $i = \frac{25}{100}(8) = 2 \quad Q_1 = \frac{109+114}{2} = 1115$
- Q_2: $i = \frac{50}{100}(8) = 4 \quad Q_2 = \frac{116+121}{2} = 1185$
- Q_3: $i = \frac{75}{100}(8) = 6 \quad Q_3 = \frac{122+125}{2} = 1235$

InterQuartile Range (IQR)

- It is the range for the **<u>middle 50%</u>** of the data.
- It ***overcomes the sensitivity*** to extreme data values.
- Also known as **Midspread**:
  Spread in the Middle 50%
- The **<u>IQR</u>** of a data set is the difference between the third quartile and the first quartile.

$IQR = Q_3 - Q_1$



Example] 11 12 13 16 16 17 17 18 21

$IQR = Q_3 - Q_1 = 17.5 - 12.5 = 5$

Standard Deviation

- **Standard Deviation** is a measure of the ***"average"*** deviation of all observations from the mean. It is the most frequently used measure of variability/spread.
- It is the positive square root of the variance of a data set.
- It is measured in the **<u>same units as the data</u>**, making it more easily comparable, than the variance, to the mean.
- It provides an overall measurement of how much participants' scores differ from the ***<u>mean</u>*** score of their group. It is a special type of average of the deviations of the scores from their mean.
- The more spread out participants are around their mean, the larger the standard deviation.

Standard Deviation

- To calculate **Standard Deviation**:
  - Calculate the **mean**.
  - Determine each observation's **deviation (x - xbar).**
  - "Average" the ***squared*-deviations** by dividing the total *squared* deviation by **(*n* - 1)**.
  - This quantity is the **Variance**.
  - Square root the result to determine the **Standard Deviation.**

Standard Deviation

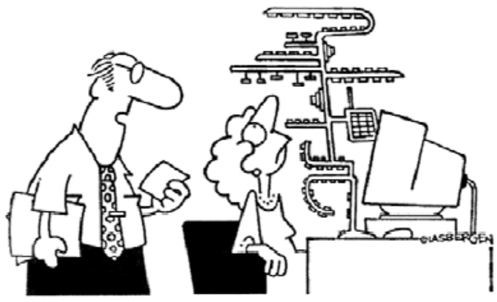
- If the data set is a ***sample***, the standard deviation is denoted ***s***.

$$s = \sqrt{s^2} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$$

- If the data set is a ***population***, the standard deviation is denoted **σ** (sigma).

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum (x_i - \mu)^2}{N}}$$

Standard Deviation



"It's the new keyboard for the statistics lab. Once you learn how to use it, it will make computation of the standard deviation easier."

Pattern of a Distribution "SOCS"

- ***Shape***
  - Modes: Major peaks in the distribution
  - Symmetric: The values smaller and larger than the midpoint are mirror images of each other
  - Skewed to the right: Right side of the graph extends much farther out than the left side.
  - Skewed to the left: Left side of the graph extends much farther out than the right side.
- ***Center (Location)***
  - Mean: The arithmetic average. Add up the numbers and divide by the number of observations.
  - Median: List the data from smallest to largest. If there is an odd number of data values, the median is the middle one in the list. If there is an even number of data values, average the middle two in the list

Pattern of a Distribution "SOCS"

- ***Spread***
  - Range: The difference in the largest and smallest value. (Max – Min)
  - Standard Deviation: Measures spread by looking at how far observations are from their mean.
    The computational formula for the standard deviation is

$$s = \sqrt{\frac{1}{n-1}\sum(x_i - \bar{x})^2}$$

  - Interquartile Range (IQR): Distance between the first quartile (**Q_1**) and the third quartile (**Q_3**). **IQR = $Q_3 - Q_1$**
    Q_1 – 25% of the observations are less than Q_1 and 75% are greater than Q_1.
    Q_3 – 75% of the observations are less than Q_3 and 25% are greater than Q_3.

Pattern of a Distribution "SOCS"

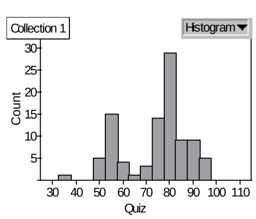
- ***Outlier/Unusual Feature***
  - An individual value that falls outside the overall pattern.
  - Identifying an outlier is a matter of judgment. Look for points that are clearly apart from the body of the data, not just the most extreme observations in a distribution.
  - You should search for an explanation for any outlier.
  - Sometimes outliers points to errors made in recording data.
  - In other cases, the outlying observation may be caused by equipment failure or other unusual circumstances.

Rule of Thumb

$1.5 \times$ IQR

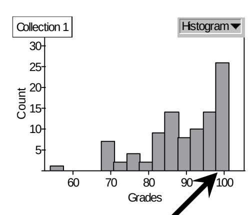
SOCS

- **Shape:** The shape is bimodal, and around each mode the shape is roughly symmetric.
- **Outlier/Unusual features:** There is a gap in the lower 40's, with a possible outlier in the mid 30's.
- **Center:** This distribution of quiz scores appears to have two modes, one at around 55, and another at around 80.
- **Spread:** The spread is from the mid-30's to the mid-90's.



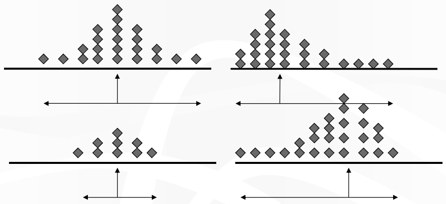
More SOCS…

- **Shape:** The shape is unimodal and skewed to the left (to the lower grades)
- **Outlier/Unusual features:** There is a gap from the upper 50's to the upper 60's, with a possible outlier in the mid 50's.
- **Center:** This distribution of grades has a single mode at around 100.
- **Spread:** The spread is from the mid-50's to about 100.



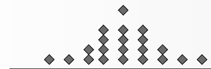
this does NOT mean that someone had a grade of above 100. ***(more likely, a lot of 98's and/or 99's)***

Interpreting Graphs: *Location and Spread*

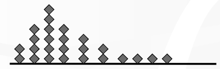


- Where is the data centered on the horizontal axis, and how does it spread out from the center?

Interpreting Graphs: *Shapes*



Mound shaped **and** ***symmetric*** **(mirror images)**

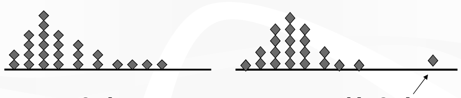


Skewed right**: a few unusually large measurements**

Skewed left**: a few unusually small measurements**

Bimodal**: two local peaks**

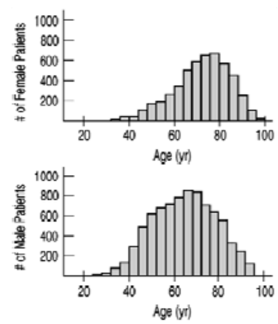
Interpreting Graphs: *Outliers*



- *Are there any strange or unusual measurements that stand out in the data set?*

Comparing Distributions

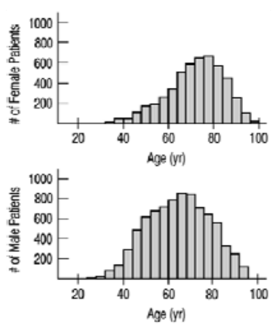
Compare the following distributions of ages for female and male heart attack patients.



Comparing Distributions

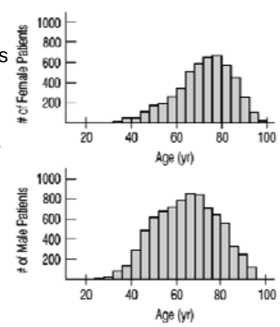
Be sure to use language of comparison.

- **Center:** This distribution of ages for females has a higher center (at around 78) than the distribution for male patients (around 62).
- **Shape:** Both distributions are unimodal. The distribution for males is nearly symmetric, while the distribution for females is slightly skewed to the lower ages.



Comparing Distributions

- **Spread:** Both distributions have similar spreads: females from around 30 – 100, and males from about 24 – 96. Overall, the distribution for female ages is slightly higher than that for male ages.
- (There are no **outliers or unusual features**)
- **YOU MUST USE COMPLETE SENTENCES!!!**



Data Change

- A survey conducted in a college intro stats class asked students about the number of credit hours they were taking that quarter. The number of credit hours for a random sample of 16 students is given in the table below.

10	10	12	14	15	15	15	15
17	17	19	20	20	20	20	22

- Compute the following:

a) Mean
b) Median
c) Range
d) IQR
e) Standard Deviation

Data Change

- Suppose that the student taking 22 credit hours in the data set in the previous question was actually taking 28 credit hours instead of 22 (so we would replace the 22 in the data set with 28). Indicate whether changing the number of credit hours for that student would make each of the following summary statistics increase, decrease, or stay about the same:

10	10	12	14	15	15	15	15
17	17	19	20	20	20	20	28

a) Mean
b) Median
c) Range
d) IQR
e) Standard Deviation