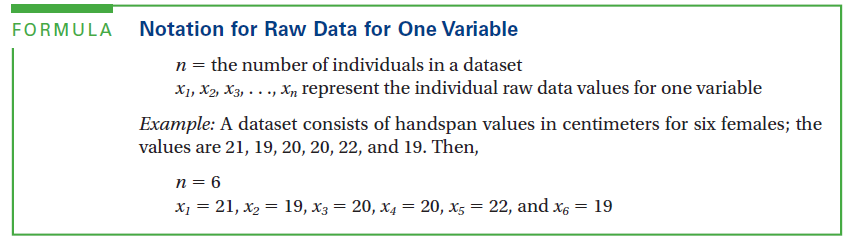
Summary statistics

Raw data:

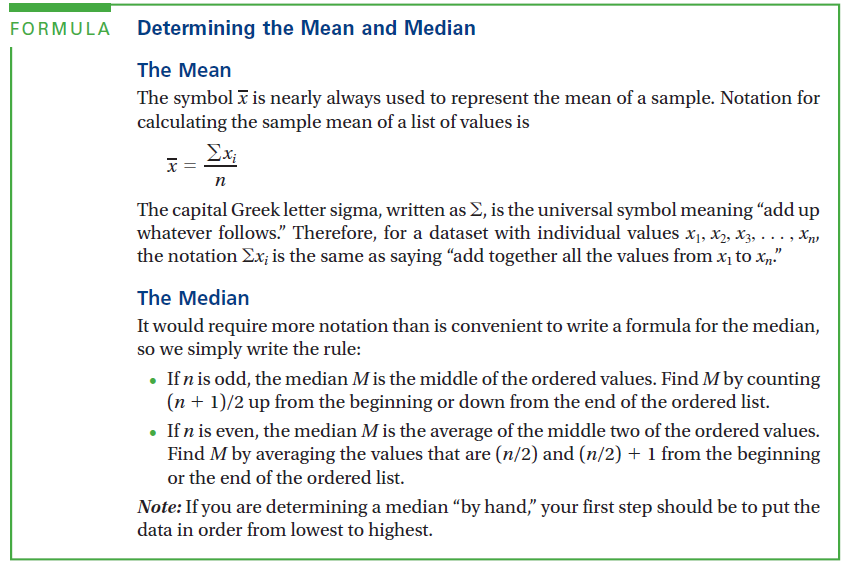
X1,x2,x3,…..xn………..

Sample size: n



Describing location (measures of center)

Mean: arithmetic average



Mode: most common value of the data set

Might be 2,3,4… or 0,1 …… mode(s)

Midrange: average of max and min data values -- (x1+xn)/2

Effect of outliers:

**Resistant**: Value is not affected by outliers

Median is resistant

Mean is not

In a strong bimodal data set, mean and median are not typical values\*\*\*\*\*

Effect of shape

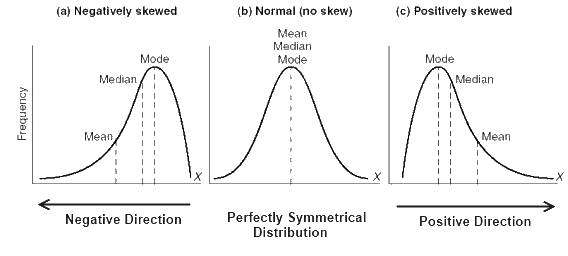
Perfectly symmetric data have mean=median

In general, symmetric data mean is approx. median

**Skewed data:**

**Skewed left: mean<median**

**Skewed right: mean>median**



Describing spread

(measures of variability)

Range: difference between max and min valve: Max minus min(Q4-Q0)

Interquartile range(IQR): difference between upper quartile and lower quartile

**IQR=Q3(upper)-Q1(lower)**

**Lower quartile: Q1 🡪 median of the data below the median**

**Upper quartile: Q3🡪 median of the data above the median**

**Q2=Median**

**Example finding the L/U quartiles**

**When n=odd**

**Eg. 1,2,3,4,4,5,6,7,8**

**M=4 Q1=(2+3)/2=2.5 Q3=(6+7)/2=6.5**

**When n=even**

**Eg. 1,2,3,4,4,5,6,7 the median of this number set in-between the 2 4s in the middle(4 M 4), even the 2 4s are both 4**

**M=(4+4)/2=4 Q1=the median of set(1,2,3,4)=(2+3)/2=2.5 Q3=the median of set(4,5,6,7)=(5+6)/2=5.5**

**5 number summary:**

*Min=Q0 Lower Q=Q1 Median=Q3 Upper Q=Q3 Max=Q4*

**Percentiles:**

*Kth percentile divides data into K% at or below, (100-K)% at or above \*\**

***Is it an outlier?***

***Values below Q1-1.5 x IQR***

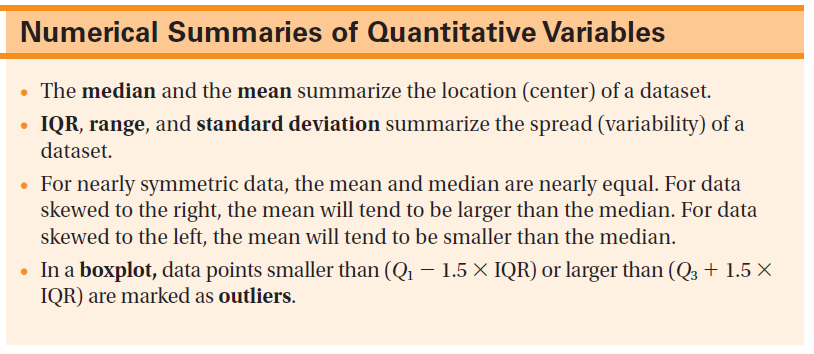
***Values above Q3+1.5 x IQR are outliers (x=times)***

**Boxplot**

-box from Q1toQ3 with line for Median

-whiskers go up/down to max/min value excluding outliers

-show outlier(s) with asterisk



**Outliers**

1. It’s a legit data value 🡪 keep it
2. It’s a mistake 🡪 if possible, fix it/if not, discard
3. It’s real, but belongs to a different group 🡪 which group are you studying?