Comparing Data Sets

Center – the middle of your set of data; represented by the mean, median, and/or mode

- ***Mean***—Susceptible to be pulled in one direction by outliers
- ***Median***—Stable center because it is NOT pulled in one direction by OUTLIERS
- ***Mode***—Can have multiple values and may not always be in the center

75, 80, 80, 70, 40

Comparing Data Sets using Measure of Center:

Example:

Data Set A: 20, 64, 70, 70, 75, 80, 82, 90, 100

Data Set B: 60, 64, 70, 70, 75, 80, 82, 90, 100

a. Find the measures of centers.

Data Set A:

Mean	72.3
Median	75
Mode	70

Data Set B:

Mean	76.8
Median	75
Mode	70

b. What measure of center would be the best to compare the two data sets? Why?

mean, b/c lower-outlier

c. Looking at the measures of center, how do the data sets compare?

performed almost the same
Data A - lower mean

Spread – the variability of your set of data; represented by range, IQR, MAD, and _____ _____.

- ***Range***—describes the entire spread of data; susceptible to outliers
- ***IQR***—describes the middle 50% of data; NOT susceptible to outliers
- ***MAD***—tells you the average spread; susceptible to outliers

Comparing Data Sets using Measure of Spread:

Example:

$Q_1 = 67$ $Q_3 = 86$

Data Set A: 20, 64, 70, 70, 75, 80, 82, 90, 100

Data Set B: 60, 64, 70, 70, 75, 80, 82, 90, 100

$Q_1 = 67$ $Q_3 = 86$

a. Find the measures of centers.

Data Set A:

Range	80
IQR	19
MAD	14.53

$Q_3 - Q_1 = 19$

Data Set B:

Range	40
IQR	19
MAD	9.98

b. What measure of spread would be the best to compare the two data sets? Why?

Range / MAD

c. Looking at the measures of spread, how do the data sets compare?

Class B is more consistent

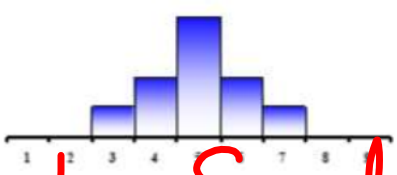
Warm-up: 1.10.22

1) Pick up a handout on the back table.

2) Get calculator.

3) Get out HW and calendar.

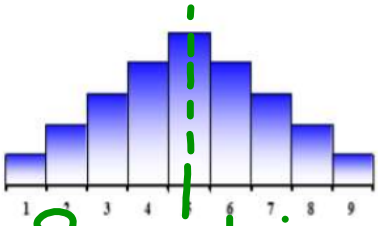
4) Complete questions for Day 1 on handout.



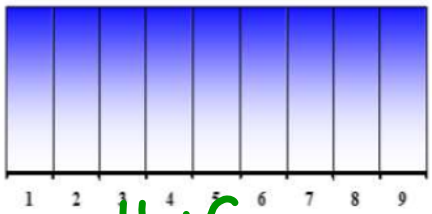
Less Spread

More Spread

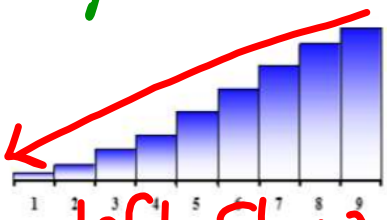
Shape – the shape of a distribution is described by Symmetry, Skew, number of peaks



Symmetric

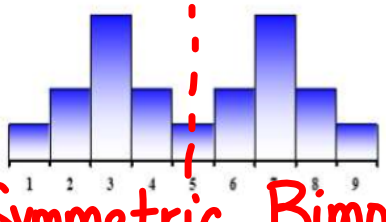


Uniform

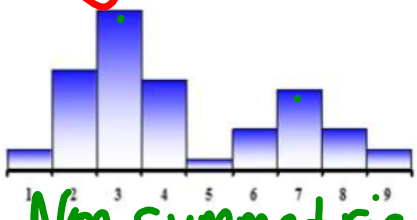


Left-Skew

Right Skew

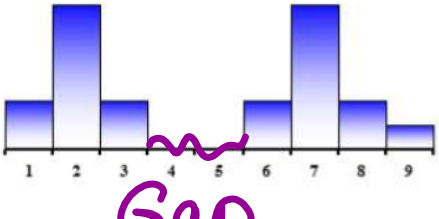


Symmetric Bimodal

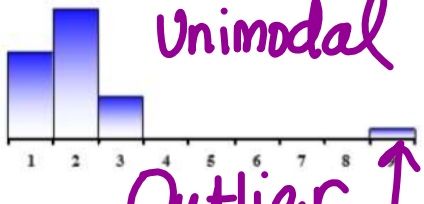


Non-symmetric Bimodal

Unusual Features



Gap



Unimodal

Outlier