Statistics Warm Up

1. List the first nine even integers greater than zero

   Ans:

2. Find the mean of the first nine even integers greater than zero.

   $$\bar{x} = \frac{2+4+6+8+10+12+14+16+18}{9}$$

   $$= \frac{90}{9} = 10$$

   Ans: mean =

3. Find the mode and median for the data below.

   5, 3, 1, 2, 1, 4, 6, 5, 3, 1, 5, 6, 2, 1

   1, 1, 1, 1, 2, 2, 3, 3, 4, 5, 5, 5, 6, 6

   Ans: mode =

   median =

   $$\frac{14+1}{2}$$

   $$= \frac{15}{2}$$

   $$= 7.5$$

Before we begin, are there any questions from last day's work?

## Today's Learning Goal(s):   Word Wall

By the end of the class, I will be able to:

a) understand the effect of an outlier
       on a small *vs*. large set of data
b) calculate the range for a set of data

Use Excel File:

Central Tendency - With Outliers, Large vs Small Data Sets

**This is a *small-sized* data set.**

| 4 | 5 | 7 | 2 | 9 | 3 | 6 | 9 | 8 | 5 mean....... 5.5 |
| 8 | 5 | 6 | 2 | 3 | 9 | 7 | 3 | 6 | 2 median.... 5.5 |

**This is a *small-sized* data set.**

| 4 | 5 | 7 | 2 | 9 | 3 | 6 | 99 | 8 | 5 mean....... 10.0 |
| 8 | 5 | 6 | 2 | 3 | 9 | 7 | 3 | 6 | 2 median.... 5.5 |

**This is a LARGE-sized data set.**

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 9 | 8 | 5 | 3 | 1 | 4 | 6 | 7 | 8 | 9 | 8 | 9 | | mean...... | 6.0 |
| 9 | 9 | 4 | 9 | 8 | 4 | 10 | 7 | 2 | 1 | 10 | 8 | | median.... | 6.0 |
| 3 | 1 | 4 | 2 | 10 | 3 | 8 | 9 | 9 | 4 | 5 | 2 | | | |
| 2 | 5 | 9 | 8 | 6 | 5 | 2 | 9 | 7 | 10 | 10 | 3 | | | |
| 9 | 4 | 3 | 3 | 7 | 7 | 8 | 3 | 8 | 4 | 10 | 8 | | | |
| 8 | 6 | 9 | 5 | 3 | 7 | 10 | 10 | 8 | 4 | 10 | 2 | | | |
| 10 | 5 | 1 | 7 | 7 | 9 | 5 | 9 | 9 | 4 | 10 | 10 | | | |
| 6 | 2 | 8 | 7 | 8 | 9 | 2 | 1 | 7 | 10 | 6 | 8 | | | |
| 7 | 2 | 10 | 9 | 7 | 3 | 5 | 8 | 5 | 7 | 8 | 10 | | | |
| 3 | 3 | 6 | 9 | 4 | 10 | 2 | 1 | 2 | 1 | 3 | 6 | | | |
| 7 | 8 | 7 | 3 | 9 | 9 | 8 | 6 | 3 | 5 | 4 | 5 | | | |
| 8 | 3 | 3 | 10 | 10 | 1 | 6 | 7 | 8 | 4 | 5 | 10 | | | |
| 7 | 8 | 7 | 9 | 10 | 8 | 8 | 9 | 5 | 6 | 10 | 5 | | | |
| 4 | 6 | 8 | 1 | 8 | 1 | 3 | 6 | 4 | 5 | 8 | 9 | | | |
| 9 | 2 | 7 | 8 | 8 | 6 | 4 | 9 | 9 | 7 | 4 | 7 | | | |
| 1 | 2 | 4 | 1 | 5 | 7 | 5 | 5 | 7 | 6 | 9 | 8 | | | |
| 1 | 5 | 9 | 6 | 2 | 2 | 10 | 8 | 4 | 3 | 2 | 6 | | | |
| 8 | 5 | 9 | 7 | 4 | 6 | 8 | 9 | 1 | 8 | 1 | 5 | | | |
| 8 | 4 | 1 | 2 | 4 | 9 | 4 | 10 | 9 | 3 | 6 | 7 | | | |
| 5 | 4 | 8 | 6 | 3 | 5 | 9 | 6 | 5 | 8 | 9 | 10 | | | |
| 2 | 9 | 8 | 9 | 9 | 8 | 2 | 10 | 10 | 6 | 1 | 8 | | | |
| 4 | 3 | 5 | 5 | 2 | 9 | 1 | 8 | 6 | 6 | 4 | 7 | | | |
| 4 | 9 | 2 | 8 | 8 | 2 | 6 | 2 | 4 | 6 | 10 | 2 | | | |
| 7 | 2 | 10 | 9 | 3 | 6 | 10 | 2 | 3 | 5 | 4 | 6 | | | |
| 1 | 3 | 5 | 2 | 5 | 10 | 4 | 4 | 5 | 5 | 4 | 6 | | | |

**This is a LARGE-sized data set.**

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 9 | 8 | 5 | 3 | 1 | 4 | 6 | 7 | 8 | 9 | 8 | 99 | | mean...... | 6.3 |
| 9 | 9 | 4 | 9 | 8 | 4 | 10 | 7 | 2 | 1 | 10 | 8 | | median.... | 6.0 |
| 3 | 1 | 4 | 2 | 10 | 3 | 8 | 9 | 9 | 4 | 5 | 2 | | | |
| 2 | 5 | 9 | 8 | 6 | 5 | 2 | 9 | 7 | 10 | 10 | 3 | | | |
| 9 | 4 | 3 | 3 | 7 | 7 | 8 | 3 | 8 | 4 | 10 | 8 | | | |
| 8 | 6 | 9 | 5 | 3 | 7 | 10 | 10 | 8 | 4 | 10 | 2 | | | |
| 10 | 5 | 1 | 7 | 7 | 9 | 5 | 9 | 9 | 4 | 10 | 10 | | | |
| 6 | 2 | 8 | 7 | 8 | 9 | 2 | 1 | 7 | 10 | 6 | 8 | | | |
| 7 | 2 | 10 | 9 | 7 | 3 | 5 | 8 | 5 | 7 | 8 | 10 | | | |
| 3 | 3 | 6 | 9 | 4 | 10 | 2 | 1 | 2 | 1 | 3 | 6 | | | |
| 7 | 8 | 7 | 3 | 9 | 9 | 8 | 6 | 3 | 5 | 4 | 5 | | | |
| 8 | 3 | 3 | 10 | 10 | 1 | 6 | 7 | 8 | 4 | 5 | 10 | | | |
| 7 | 8 | 7 | 9 | 10 | 8 | 8 | 9 | 5 | 6 | 10 | 5 | | | |
| 4 | 6 | 8 | 1 | 8 | 1 | 3 | 6 | 4 | 5 | 8 | 9 | | | |
| 9 | 2 | 7 | 8 | 8 | 6 | 4 | 9 | 9 | 7 | 4 | 7 | | | |
| 1 | 2 | 4 | 1 | 5 | 7 | 5 | 5 | 7 | 6 | 9 | 8 | | | |
| 1 | 5 | 9 | 6 | 2 | 2 | 10 | 8 | 4 | 3 | 2 | 6 | | | |
| 8 | 5 | 9 | 7 | 4 | 6 | 8 | 9 | 1 | 8 | 1 | 5 | | | |
| 8 | 4 | 1 | 2 | 4 | 9 | 4 | 10 | 9 | 3 | 6 | 7 | | | |
| 5 | 4 | 8 | 6 | 3 | 5 | 9 | 6 | 5 | 8 | 9 | 10 | | | |
| 2 | 9 | 8 | 9 | 9 | 8 | 2 | 10 | 10 | 6 | 1 | 8 | | | |
| 4 | 3 | 5 | 5 | 2 | 9 | 1 | 8 | 6 | 6 | 4 | 7 | | | |
| 4 | 9 | 2 | 8 | 8 | 2 | 6 | 2 | 4 | 6 | 10 | 2 | | | |
| 7 | 2 | 10 | 9 | 3 | 6 | 10 | 2 | 3 | 5 | 4 | 6 | | | |
| 1 | 3 | 5 | 2 | 5 | 10 | 4 | 4 | 5 | 5 | 4 | 6 | | | |

MBF 3CI          **<u>3.4 Measures of Central Tendency</u>** (Day 2)

*"Which measure of central tendency should I use?"*          Date: <u>Mar. 24/17</u>

Some general guidelines:
- Outliers will affect the mean the most, especially if the sample size is small.
  Hence, use the median if the data contains outlier(s) in a   **smaller** data set.

*Ex. 5:*  Calculate the mean and median for the data:   6, 99, 4, 8, 2

$$Mean = \frac{6+99+4+8+2}{5}$$
$$= \frac{119}{5}$$
$$= 23.8$$

Median: 2, 4, 6, 8, 99
$$\to 6$$

- If the data are mainly symmetric, the mean and median will be close, or exactly the same,
  so either is fine to use.  This applies in a small   **or** large data set!

*Ex. 6:*  Calculate the mean and median for the data:   6, 7, 4, 8, 2

$$Mean = \frac{6+7+4+8+2}{5}$$
$$= \frac{27}{5}$$
$$= 5.4$$

Median: 2, 4, 6, 7, 8
$$\to 6$$

- Use the mode when the frequency of the data is more important than the calculated value (i.e. shoe size),
  and when the data is discrete (i.e. hair colour).

The **range** of a data set is the difference between the largest piece of data and the smallest.

*Ex. 7:*   A simple random sample of car owners was performed.
Each owner was asked how old they were when they got their first car and results are summarized:

Recall interval notation:   "Square brackets" = [ …include the value,  "Round brackets"( = .Do Not include the value.

| Age | [15 to 20) | [20 to 25) | [25 to 30) | [30 to 35) | [35 to 40) |
|---|---|---|---|---|---|
| Frequency | 10 | 18 | 12 | 8 | 2 |

*(handwritten annotations above table:)* 17 — 15,16,17,18,19   22 — 20,21,22,23,24   27   32   37 — 35,36,37,38,39

a)  Calculate the range of the data.

$$\text{Range} = 39 - 15$$
$$= 24$$

b)  *Approximate* the mode, median and mean age for this sample.

Mode: 20-25
∴ 22

Median: Add Frequency = 50
∴ ½ between 25 + 26th person
∴ in the 20-25 interval
∴ approx. 22

$$\text{Mean} = \frac{10(17) + 18(22) + 12(27) + 8(32) + 2(37)}{50}$$
$$= \frac{1220}{50}$$
$$= 24.4$$

*Entertainment:* pp. 136-138 #4, 5, 6, 7c (do 7a and 7b if you didn't do it last class), 9
                    ARE YOU CHECKING YOUR ANSWERS?