

**5.1 Exploring Data – Mean, Median , Mode and Range.**

**Three different Types of “Averages” in Statistics**

**Mean – THE AVERAGE ;** Determine the sum of all the data entries and then divide by the number of items. (Symbol =  $\bar{x}$  )

Ex. Find the average of: 2, 2, 2, 4, 20 =  $\bar{x} = \frac{2+2+2+4+20}{5} = \frac{30}{5} = 6$

: 5.0, 5.5, 6.0, 6.5, 0.0, 7.0 =  $\bar{x} = \frac{30}{6} = 5$

How are these two sets of data different? *1<sup>st</sup> set has an outlier (a data*

**Mode – The number that is represented most often.** If no numbers repeat there is no mode. There can be more than one mode. *Find the mode in ex. above* ① 2 = mode

② No mode

**Median – The middle # when the data is arranged in order from smallest to largest.**

Ex. Find the median of: 40, 2, 17, 2, 2 ( note: Odd number of data) = 2, 2, 2, 17, 40 median = 2

: 6, 11, 2, 2, 10, 3 ( Note: Even number of data) = 2, 2, 3, 6, 10, 11 median =  $\frac{3+6}{2} = \frac{9}{2} = 4.5$

**Other Measure in Statistics:**

**Dispersion** - a measure that varies by the spread among the data in a set ; dispersion has a value of zero if all data in the set is identical, and it increases in value as the data becomes more spread out.

One way to measure dispersion is by the **Range** = Highest - lowest value

**Example 1:** Read on Pg 210

Measured Lifespans of 30 car batteries (years)									
Brand X					Brand Y				
5.1	7.3	6.9	4.7	5.0	5.4	6.3	4.8	5.9	5.5
6.2	6.4	5.5	5.7	6.8	4.7	6.0	4.5	6.6	6.0
6.0	4.8	4.1	5.2	8.1	5.0	6.5	5.8	5.4	5.1
6.3	7.5	5.0	5.7	8.2	5.7	6.8	5.6	4.9	6.1
3.3	3.1	4.3	5.9	6.6	4.9	5.7	6.2	7.0	5.8
5.8	6.4	6.1	4.6	5.7	6.8	5.9	5.3	5.6	5.9

**How can you compare the data to help Paulo decide which brand of battery to buy?**

- Determine the mean, median, mode and range of each Brand.
- Explain why the mean and median do not fully describe the difference between these two brands of batteries. Consider the range. What additional information can be learned from the range?
- Is the mode useful in this situation?
- Suppose one battery from brand Y is defective and has a lifespan of 0.5 years instead of 5.9 years. Would this affect Paulo's decision?

a)  $\bar{x} = \frac{172.3}{30} = 5.74$  ✓

$\bar{x} = \frac{171.6}{30} = 5.72$   $\rightarrow \frac{171.7}{30} = 5.72$

Median = put all #'s in order low-high  
 $= \frac{5.7 + 5.8}{2} = 5.75$

Median =  $\frac{5.75 + 5.8}{2} = 5.75$

Mode = 5.7

Mode = 5.9

Range =  $8.2 - 3.1 = 5.1$

Range =  $7 - 4.5 = 2.5$

b) If you only know the mean + median you do not know if the data is clustered or if any of the data is extreme. The range will tell you how far the data is spread out. Brand X has a greater range than Brand Y, this means that a Brand X battery may last a lot longer than a Brand Y battery, or far less.

Assignment Pg 211 #1,2

#### Extra questions

- Determine if each of the following estimates or predictors are based on mean, median or mode.
  - The average salary for students with part-time jobs is \$86.78/week.
  - The middle mass of a collage football player is 70.62 kg
  - Most snowmobile accidents occur in the winter
  - The average speed on a Canadian highway is 95km/h
  - More people play golf then play billiards.
- The following data is from two math tests given by the same teacher to the same class in the same semester

#### Unit Test #1

81	80	79	79	78	76	75	75	74	73	73	73
73	73	72	71	71	68	67	66	64	63	61	58

#### Unit Test #2

98	95	93	89	87	84	81	79	79	76	73	73
73	73	73	71	69	64	59	59	57	53	44	41

- Find the range of the data for both tests
- Find the mean of the data for both tests
- Find the mode of the data for both tests
- Find the median of the data for both tests
- Which test did the class perform better on? Justify your answer

#### Answers:

- a) mean b) median c) mode d) mean e) mode
2. a) 23, 57    b) 71.8, 72.6    c) 73, 73    d) 73, 73    e) Discussion

c) No  $\rightarrow$  the values are so similar they don't inform any differences that well.

$\hookrightarrow$  Mode can tell us consistency of actual expectancy.  
The mode doesn't help me decide which brand of battery is better to buy.

d) It depends whether the new value of O.S is considered an outlier or an accepted part of the range.

Note: an outlier should not be used in a calculation it will skew the data.

If it is included the data will look less favorable and may not be chosen. As well if 1 out of 30 batteries is defective, it may reduce the chance of it being chosen unless there is warranty.