## *MATH SPEAK - TO BE UNDERSTOOD AND MEMORIZED*

1) **STATISTICS** = the branch of mathematics used to analyze and interpret ***data***.
2) **DATA** = information, often in the form of numbers, that are gathered, organized, analyzed and interpreted.
3) **DATA SET** = a group of numbers to be statistically analyzed.
4) **POPULATION** = all the members of a group to be statistically analyzed.
5) **POPULATION SIZE** = the number of pieces of data in a ***data set*** and is represented by ***n***.
6) **SAMPLE** = some of the members of a group to be statistically analyzed.
7) **SAMPLE SIZE** = the number of pieces of data in a ***data set*** and is represented by ***n***.
8) **RANGE** = ***range*** $=$ largest number $-$ smallest number . ***Range*** gives a rough indication of the ***dispersion*** of the numbers within a ***data set***.
9) **LINE PLOT** = a number line having dots the representing each of the numbers in a ***data set*** above the corresponding number on the number line.
10) **MODE** = the number in a ***data set*** that appears most often.
11) **MEDIAN** = the number that divides the ***data set*** into two equal upper and lower halves.
12) **MEAN** = the average of all the numbers in a ***data set***. ***Mean*** is represented by these symbols: $\bar{x}$ = ***mean*** for a sample, $\mu$ = ***mean*** for an entire population.
13) **OUTLIER** = a ***number*** that is judged to be very far from the other ***data*** in the ***data set***. **i.e.** in the ***data set*** below, 35 is the outlier because it is 20 units away from its closest number.

    **e.g.** $x = {}^{-}7,\ {}^{-}3,\ 0,\ 1,\ 2,\ 5,\ 9,\ 11,\ 15,\ 35$
14) **STANDARD DEVIATION** = a number that describing the variation/dispersion (spread) of the ***data*** relative to the ***mean*** of the ***data***.

---

## *STANDARD DEVIATION*

I) ***STANDARD DEVIATION is a number describing that describes how spread out the data is relative to the data's mean***

A) If the ***standard deviation*** is small, the ***data*** points are close to each other and are close to the ***mean***. If the ***standard deviation*** is large the ***data*** points are spread throughout the ***range*** of the ***data*** with some ***data*** points being close to the ***mean*** with others being far from the ***mean***.

1) ***Standard deviation*** represented with the symbol $\sigma$ and is calculated using formulae given below.

$$\sigma = \sqrt{\frac{\sum\left(x-\bar{X}\right)^2}{n}} \quad \text{or} \quad \sigma = \sqrt{\frac{\sum\left(x-\mu\right)^2}{n}}$$

where $\sigma =$ ***standard deviation***, $\Sigma =$ the sum of, $x =$ piece of data, $\bar{X}$ or $\mu =$ ***mean*** and $n =$ ***sample*** or ***population size***

a) You are not required to know or use these ***standard deviation*** formulae to answer questions. They are included in this notes package for your information only.

b) Compare the formulae for ***standard deviation*** to the formula used to calculate ***mean***. **NOTICE** that they are similar in that they create a number that is divided by ***n*** (***sample*** or ***population size***). This is an indication that the ***standard deviation*** is another way to calculate an ***average***, specifically the ***average distance*** the data points are from the ***mean***.

2) ***SAMPLE PROBLEM 1:*** Calculate the ***standard deviation*** for this ***data set:*** $x = 2,\ 3,\ 4,\ 7$

$$\bar{X} = \frac{\sum x}{n} = \frac{2+3+4+7}{4} = \frac{16}{4} = 4$$

$$\sigma = \sqrt{\frac{\sum (x-\bar{X})^2}{n}} = \sqrt{\frac{(2-4)^2+(3-4)^2+(4-4)^2+(7-4)^2}{4}} = \sqrt{\frac{(^-2)^2+(^-1)^2+(0)^2+(3)^2}{4}}$$

$$\sigma = \sqrt{\frac{4+1+0+9}{4}} = \sqrt{\frac{4+1+0+9}{4}} = \sqrt{\frac{14}{4}} = \sqrt{3.5} \doteq 1.87828 \quad \rightarrow \quad \sigma \doteq 1.9$$

The ***standard deviation*** of the above ***data set*** is: $\sigma \doteq 1.9$. This means that the average distance between the each number in the ***data set*** and the ***mean*** is 1.9.

---

***ASSIGNMENT: PRINT THIS INFORMATION ON YOUR OWN GRID PAPER***

LAST then FIRST Name S2– STANDARD DEVIATION Block:

***Show the process required to complete each problem to avoid receiving a zero grade. Neatness Counts!!!*** *(Marks indicated in italicized brackets.)*

***REMEMBER TO USE GRID PAPER FOR ALL ASSIGNMENTS!!!***

Two friends are in a bowling league. They recorded their scores over twelve games in the table below.

| **Bowling Scores For 12 Games** | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Angela** | 135 | 141 | 109 | 146 | 127 | 131 | 118 | 124 | 129 | 133 | 139 | 123 |
| **Kim** | 89 | 93 | 100 | 101 | 115 | 125 | 123 | 135 | 155 | 140 | 148 | 150 |

**1)** ***For each player*** answer parts **a)** through **e)**. Be sure you show all the calculations you used to determine the ***range***, ***median*** and ***mean***. Summarize your answers in a table. ***Don't forget to include units.***

**a)** Create a ***line plot***. *(3)* **b)** Determine the ***sample size***: ***n*** = # *(0.5)* **c)** Determine the ***range***. *(1)*

**d)** Determine the ***median***. *(2)* **e)** Determine the ***mean***. *(2)*

**2)** A statistical analysis revealed that Kim's standard deviation is $\sigma = 22.3$ points while Angela's standard deviation is $\sigma = 8.7$ points. List the players from most to least consistent? Justify your answer with at least 2 pieces of evidence. *(3)*

/20