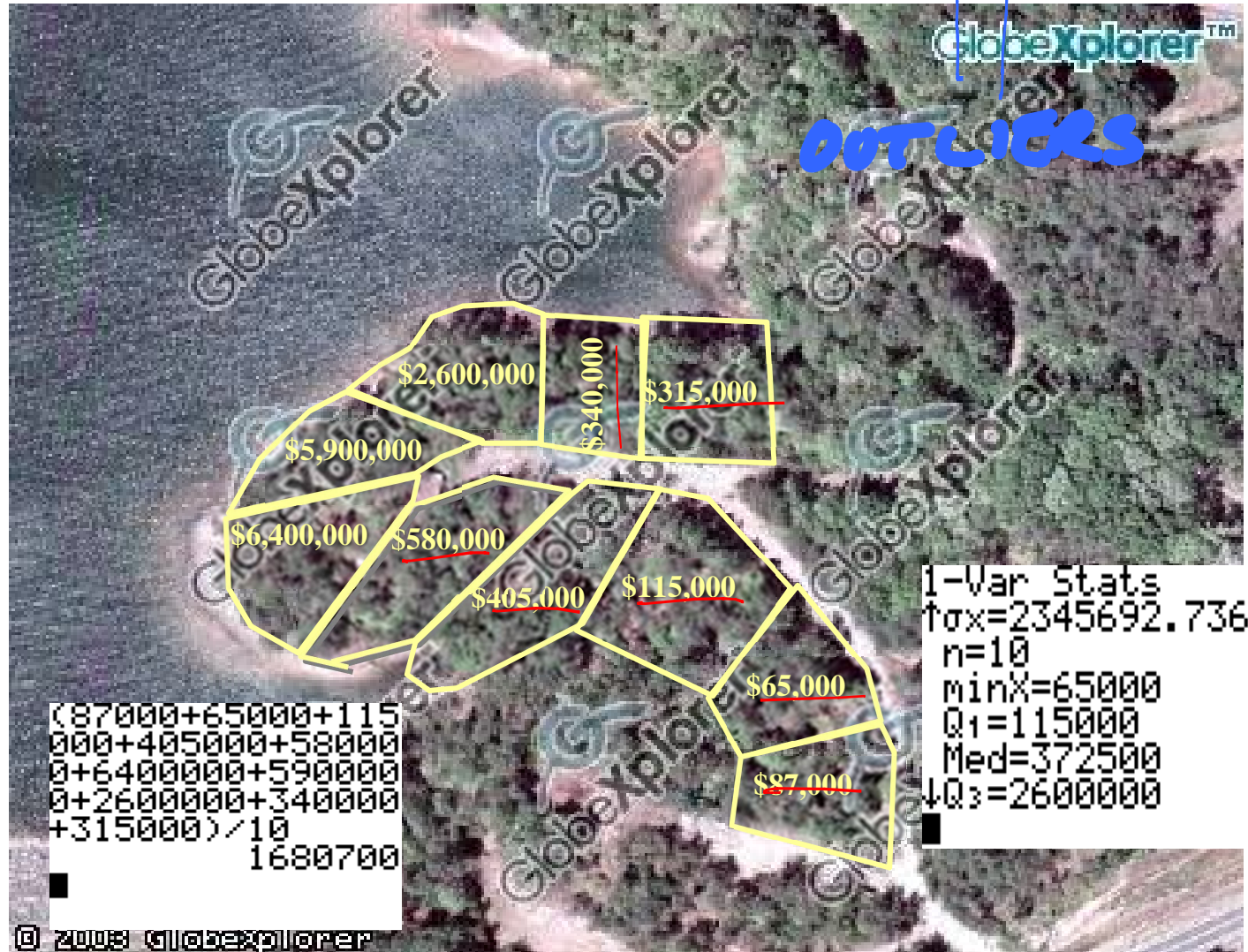




Describing Data
Central Tendencies

Name: _____

A realtor is showing a client some potential properties around a lake are shown in the map below.



Using your calculator's list menu find the Mean (\bar{x}), Median, Range, Population Standard deviation (σ).

First Reset the Stat Menu.

Press **STAT**, **5**, **ENTER**

1. Go to the LIST Editor

Press **STAT**, **1**

2. Clear out any old data. Highlight the "L1" and press CLEAR, ENTER.

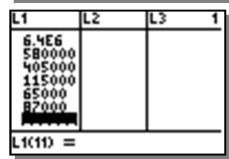
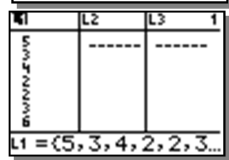
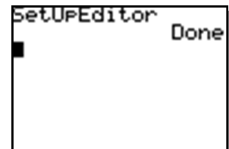
Press **▲**, **CLEAR**, **ENTER**

3. Enter each value in the "L1" column press ENTER after each value.

4. Request the calculator's One Variable Statistics.

Press **STAT**, **▶**, **1**, **ENTER**

MED = \$372,500



Which Central Tendency, Mean or Median, is a better representation of the center in this

example? (Mean (\bar{x}), Median (Med), Number of Data Points (n), Population Standard deviation (σ))

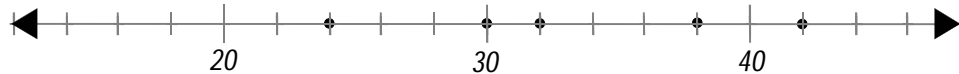
The potential client is looking to purchase property on or near a lake for somewhere around \$300,000 - \$400,000. As a realtor, would you show the client the MEAN property value for this particular cul-de-sac or would you share the MEDIAN which shows 50% of prices are above and 50% of the prices are below that price.

~~65K, 87K, 115K, 315K~~, 340K, 405K, ~~580K, 2.6M, 5.9M, 6.4M~~

MEAN (\bar{x}): The **mean** is the usual referred to as the average or **arithmetic** average. For example, consider the data set 30, 38, 24, 32, 42. The mean would be the sum of these numbers divided by the number of data points.

$$\bar{x} = \frac{30 + 38 + 24 + 32 + 42}{5} = 33.2$$

This measure of center should be used when all the data points are relatively close to one another. The above points are plotted on a number line and all of the points are relatively close.

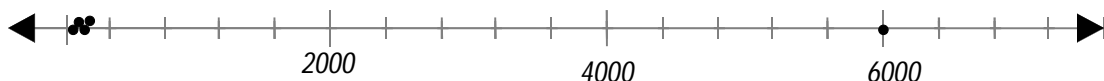


MEDIAN (MED): The median is the middle number after lining up the numbers from least to greatest. If there are 2 middle numbers the median is the average of the two middle numbers. For example given the data set 9, 31, 12, 6000, 40. The mean would be the sum of these numbers divided by the number of data points.

9, 12, **31**, 40, 6000

Median = 31

This measure of center should be used when a data set has a few outliers. Outliers are numbers that are significantly distant from the rest of the group of numbers. An outlier can drastically change the **MEAN** but has little effect on the **MEDIAN**. In the above example $\bar{x} = 1218.4$ which doesn't really represent the data set well. The median is 31 which states that half of the numbers are above 31 and half are below 31.



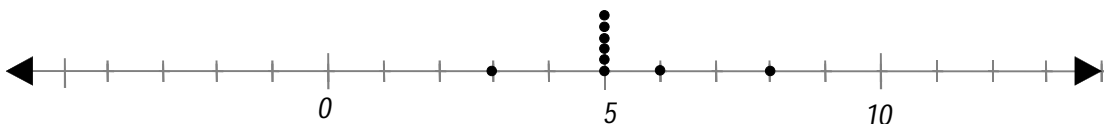
MODE : The mode is simply the number that occurs the most often in a data set. If there aren't any duplicate numbers in the data set then there isn't a mode. However if there are two numbers that occur the most, it is possible to have more than one mode.

For example given the data set 20, 30, 20, 15, 6. The mode for this set is 20.

For example given the data set 20, 30, 20, 30, 6. The mode for this set is 20 and 30.

For example given the data set 20, 10, 30, 15, 8. This set doesn't have a mode.

This measure of center should be used when there are an overwhelming number of identical data points. Example Data set: 5, 5, 6, 5, 8, 5, 3, 5, 5



STAT 101
n=10
Σx=71000
Σx²=32523

STAT 101
Σx=73000
ΣMed=84500
HQS=91000

1. A person is trying to find new employees to work for his company. She wants to show applicants a statistic about how much her employees make. Which measure of a center would be the best to describe salaries at her company if she knows that 10 of her employees make the following annual salaries?

11,000 ← **OUTLIERS** 10,000 91,000 92,000 80,000
81,000 88,000 93,000 91,000 73,000

MEAN = 71,000 **MEDIAN = 84,500** MODE = 91,000

Which center measure is most appropriate and why? MEDIAN B/C OF 11,000 10,000 ARE OUTLIERS

2. A company is selling boxes of crackers and wishes to list the number of crackers that they are contained in a box. A market research found 20 different boxes. Each box contained the following number of crackers. What number should they list on the box as the number of crackers in the box mean median or mode?

45 50 48 50 50 51 50 50 50
49 50 50 51 50 50 50 54 50 51

MEAN = 49.95 MEDIAN = 50 **MODE = 50**

Which center measure is most appropriate and why? CONSIDER MODE B/C MORE THAN HALF THE DATA SO

3. A student has the following grades in a Math class:

94 92 85 90 96 94 100 92 0 80

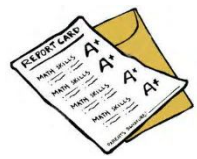
MEAN = 82.3 **MEDIAN = 92** MODE = 92 & 94

Which measure of center do you think should be used to determine the students grade and why? MEDIAN B/C "0" IS AN OUTLIER

4. A student has the following grades in a Math class:

70, 72, 68, 67, 71, 98, 100, 92, 71, 88

MEAN = 79.7 MEDIAN = 71.5 MODE = 71



Which measure of center do you think should be used to determine the students grade and why? MEAN B/C ALL DATA "RELATIVELY" CLOSE TO ONE ANOTHER

5. An employee is working at a clock shop. The clocks on the wall show the following times:



4:40pm, 4:43pm, 4:43pm, 4:46pm, 4:59pm, 4:43pm, 4:43pm, 4:43pm, 4:55pm, 4:43pm

MEAN = 4:45:48 MEDIAN = 4:43 MODE = 4:43pm

Which measure of center do you think should be used to determine the correct time and why? MODE B/C MORE THAN HALF OF THE

DATA POINTS ARE THE SAME:

6. A college wanted to list a meaningful statistic in a brochure showing the ages of the students attending their college. Below is a random sample of their students:

20, 24, 23, 26, 24, 29, 19, 21, 66, 62, 24, 28, 20, 67, 22, 20, 20, 79, 20, 20

MEAN = 31.7 yrs MEDIAN = 23.5 yrs MODE = 20



Which measure of center do you think should be used on the brochure and why? MEDIAN B/C 66, 62, 67, 79 ARE OUTLIERS

7. A student scored 92, 85, and 87 on their first three tests. What would they need to make on their fourth test to have an 'A' test average (90%)?

$$\begin{array}{r} 264 + x = 360 \\ - 264 \quad - 264 \end{array}$$

$$x = 96$$

$$4 \cdot \frac{92 + 85 + 87 + x}{4} = 90.4$$

8. A person at the mall purchased 4 shirts at an average price of \$14. How much did the 5th shirt cost if the average after the 5th shirt was purchased changed to \$16 per shirt?

$$\frac{14 + 14 + 14 + 14 + x}{5} = 16$$

$$5 \cdot \frac{56 + x}{5} = 16 \cdot 5$$

$$56 + x = 80$$

$$\begin{array}{r} 56 + x = 80 \\ - 56 \quad - 56 \end{array}$$

$$x = 24$$

9. Consider the data set 5, 2, 8, 14, 11

a. What is the mean of the data set?

$$\bar{x} = \frac{5+2+8+14+11}{5}$$

$$\bar{x} = \frac{40}{5} = 8$$

b. What is the Median of the data set?

~~2, 5, 8, 11, 14~~

MEDIAN = 8

c. If 3 were added to each number in the data set, what effect would this have on the mean and the median?

$\begin{matrix} 5+3 & 2+3 & 8+3 & 14+3 & 11+3 \\ 8 & 5 & 11 & 17 & 14 \end{matrix}$

$$\bar{x} = \frac{8+5+11+17+14}{5}$$

$$\bar{x} = \frac{55}{5} = 11$$

THE MEAN ALSO INCREASED BY 3

MED: ~~5, 8, 11, 14, 17~~

MED: 11 THE MEDIAN ALSO INCREASED BY 3.

d. If each number in the data set was doubled, what effect would this have on the mean and the median?

$\begin{matrix} 2 \cdot 5 & 2 \cdot 2 & 2 \cdot 8 & 2 \cdot 14 & 2 \cdot 11 \\ 10 & 4 & 16 & 28 & 22 \end{matrix}$

$$\bar{x} = \frac{10+4+16+28+22}{5}$$

$$\bar{x} = \frac{80}{5} = 16$$

THE MEAN WAS ALSO DOUBLED.

MED: ~~4, 10, 16, 22, 28~~

MED = 16

THE MEDIAN WAS ALSO DOUBLED.