

## GSE Algebra 1 Statistics Notes

### **Measures of Center**

- Is a single measure used to represent the middle value.

Median- is the middle most number

Mean- average number, it is the sum divided by the number of values.

$$\bar{x} =$$

Mode – the number(s) that appear most often.

Ex.1 Find the mean, median, and mode of the test data.

100, 95, 65, 70, 80, 80, 90, 90

### **Measures of Spread**

- Numbers to describe how far apart certain key values are from each other.

Minimum – the smallest value in a data set.

Maximum – the largest value in a data set.

First quartile – median of lower half of data.

Second quartile – median

Third quartile – median of upper half of data.

Range – the difference between the maximum and minimum.

$$Range =$$

Interquartile range – the difference between quartile 1 and quartile 3.

$$IQR =$$

Mean absolute deviation – the average absolute value of the difference between each data set and mean.

$$MAD =$$

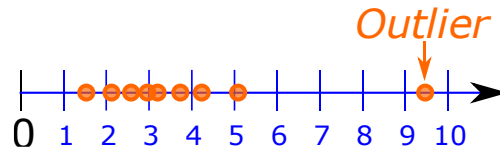
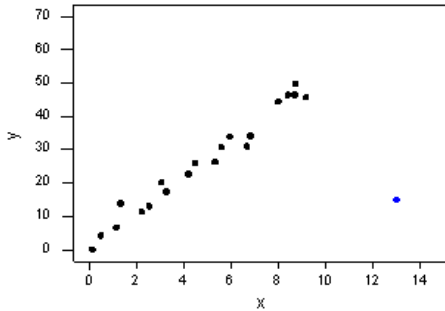
Ex.2 Below represents the April high temperatures for seven years. Find all the measures of spread for the data.

77, 86, 84, 93, 90, 81, and 80

## Outliers

- Data values that are much less/greater than most of the data set.

Extreme values – are values that appear to be outliers.



Steps to determine if a value is an outlier.

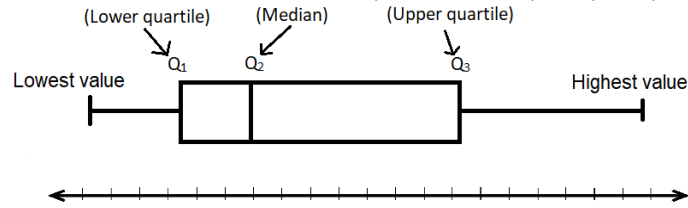
- A data value is an outlier if it is less than...  
 $Q1 - 1.5(IQR)$
- Or if a data value is greater than  
 $Q3 + 1.5(IQR)$

If there is an outlier, use the median as the measure of center and IQR as the measure of spread. Outliers greatly affect mean and range.

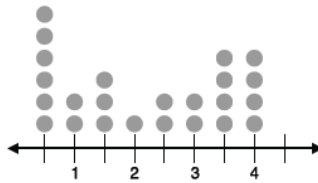
Ex.3 The following is a list of salaries in thousands. Determine if there is an outlier.  
25, 30, 35, 35, 35, 40, 40, 40, 45, 45, 50, 60, 150

## Graphs

Box Plot – is a graph that shows the minimum, maximum, Q1, Q2, and Q3.

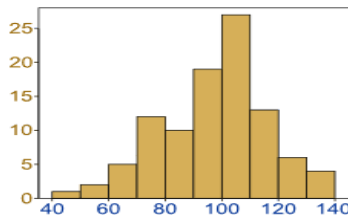


Dot plot – is a graph that uses dots to show the number of times each value in a data set appears in the data set.



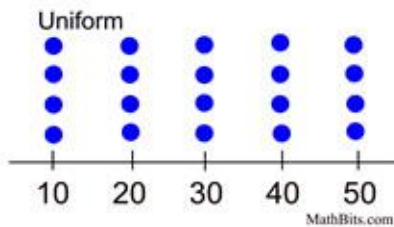
Histogram – a bar graph which shows frequency distribution.

1. Divide the range into even sections
2. Tally each frequency

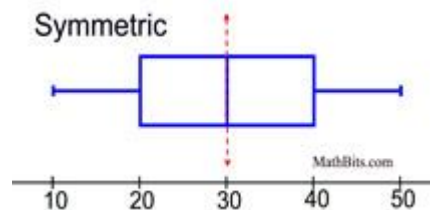
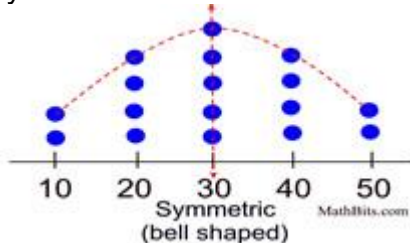


## Shape

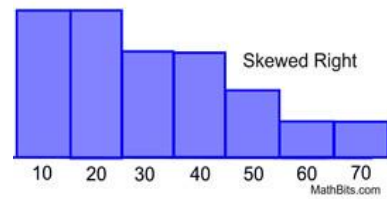
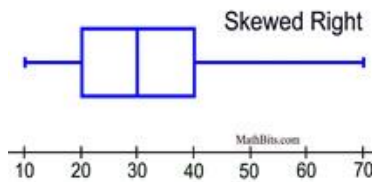
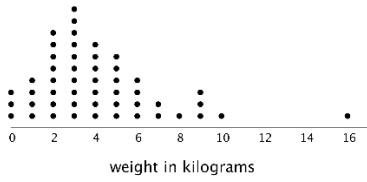
Uniform – data is evenly distributed



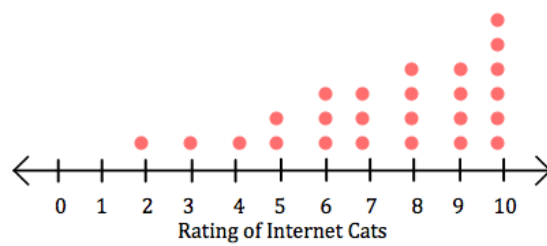
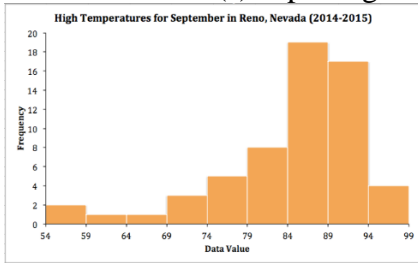
Symmetric – data is centered toward middle. Also know as normal distribution.



Skewed right – described where the outlier is pulling the data.



Skewed left – the outlier(s) is pulling the data left.



Ex.4 Find any outlier then create a dot plot, box plot, and a histogram. Then describe the shape.

2, 2, 9, 10, 10, 11, 11, 11, 12

## Two-Way Frequency Tables

- Two-way frequency table – a table that divides responses into categories.
- Joint relative frequency – the number of times a specific response is given divided by the sample.
- Marginal relative frequency – the total number of times a specific response is given divided by the sample.
- Conditional relative frequency – the percent of a joint frequency compared to the subtotal (often indicate by the word given).

Ex.1 Constructing tables.

	High School Diploma	Bachelor's Degree	Master's/ Doctoral Degree	Total
Male	16	46		65
Female		51	3	
Total	28		6	

- a) How many males have a bachelor's degree?
- b) Find the joint relative frequency of males who have a bachelor's degree.
- c) Find the marginal frequency of people with a masters/doctor's degree.
- d) Given someone is male, what is the probability of having a high school diploma?

Ex.2 Two-way relative frequency tables

<b>What is your favorite sport to watch on television?</b>			
	Football	Basketball	Baseball
Males	40	22	15
Females	12	16	45
Total	52	38	60

- a) How many females like baseball?
- b) What percent of girls prefer basketball?
- c) What is the probability of liking football?
- d) What is the probability of liking basketball or baseball?