Review Sheet for Test 1

1) a) The following data represents the fuel consumption, in miles per gallon, of a sample of 30 cars. Construct a frequency distribution for the data using 5 classes. Include the relative frequency.

   32 35 28 40 30 42 55 40 45 24 28 34 40 36 34 40 30 25 28 32

   40 35 25 44 26 39 38 42 45 32

2) The following frequency distribution shows the daily expenditure of food for a sample of 33 people

Daily Intake of Saturated Fat		
Class	Frequency	Midpoint
5 – 10	8	
11 – 16	5	
17 – 22	10	
23 – 28	6	
29 – 34	3	
35 – 40	1	

Construct a histogram and a frequency polygon from the data in the frequency distribution.

3) The following data shows the ages of the 30 highest paid CEOs.

   64 74 55 55 62 63 50 67 51 59 50 52 50 59 62 64 57
   61 49 63 62 60 55 56 48 58 64 60 60 57

   a) Create a stem and leaf plot for the data
   b) Find the range of the data.

4) The following data give the 2009-2010 tuition and fees (in thousands of dollars) for the top 10 liberal arts colleges.

   39 39 38 51 38 40 37 40 37 40 35 39

   a) Calculate the mean
   b) Calculate the median
   c) Calculate the mode
   d) Calculate the mode for the following data: 250, 390, 250, 390, 120, 120

5) a) Find the variance and standard deviation of the sample data. Round to the nearest **hundredth** where applicable.
   9 5 9 10 11 12 7 7 8 12
   b) Identify which values cover 68%, 95%, and 99.7% of the data.

6) The following data shows the age, x (in years) and systolic blood pressures, y, for 10 men.

Age, x	16	25	39	45	49	64	70	29	57	22
Blood Pressure, y	109	122	143	132	199	185	199	130	175	118

Calculate the r, the correlation coefficient. What kind of relationship does r indicate?

7) The following data shows the total payroll, in millions of dollars, and winning percentage for all the MLB teams in the American League for the 2015 baseball season.

Payroll	110	187	115	86	174	71	114	151	109	219	86	120	76	142	123
Winning %	50	48	47	50	46	53	59	53	51	54	41	47	49	54	57

   a) Calculate the equation of the regression line for the data.
   b) Use your equation from part a) to predict the winning percentage for a team with a $135 million payroll.

8) The following data give the number of hours of television watched per day by a sample of 28 people.
0 1 1 2 2 2 2 2 2 3 3 3 3 4 4 4 4 4 5 5 5 5 5 6 6 7 7 9
   a) Calculate the values of the three quartiles and the interquartile range.
   b) Find the approximate value of the 60th percentile.
   c) Find the percentile rank of 3. What does this rank mean?

9) For the data set, 3047, 4387, 3879, 2847, 4857, 9213
   a) Identify the outlier
   b) Calculate the mean and median, leaving the outlier in the data set.
   c) Calculate the mean and median, removing the outlier from the data set.

10)

Definition		**Vocabulary**
_____	A variable that can be measured numerically.	a. Continuous Variable
_____	A variable that cannot assume a numerical value. eg. Data elements that are TV shows, clothes, or apps.	b. Qualitative Variable
_____	A portion of the population, usually selected at random.	c. Quantitative Variable
_____	A variable whose values are countable. eg. 9,4,15,75	d. Discrete Variable
_____	A variable that can assume any numerical value over a certain interval. eg. $\frac{1}{2}$, $0.7 - 3.68$, $\frac{14}{19}$, 9.4	e. Sample