NAME: _____

This assignment has three parts. Part 1 requires you to define several statistical vocabulary words. In Part 2, you will use what you learn about the vocabulary words to apply it to actual data. For Part 3, you will revisit some mathematical concepts from the past that will need to be fresh on your mind throughout this course.

Part 1 – Vocabulary

Use any of the following websites to define the terms below. Feel free to use another website (or even the dictionary!) Make sure you find a good definition that *makes sense to you*!

http://stattrek.com/Help/Glossary.aspx (select term from drop down box)
http://nolanmath.com/APStat/u1voc.pdf
http://www.stats.gla.ac.uk/steps/glossary/presenting_data.html#bar
http://www.mathwords.com/index_prob_stats.htm
http://www.icoachmath.com/SiteMap/MathDictionaryMain.html

1. Population
2. Sample
3. Parameter
4. Statistic
5. Categorical Variables
14. Interquartile Range (IQR)
15. Center
16. Spread
17. Outliers
18. Symmetry

6. Quantitative Variables
7. Univariate Data
8. Bivariate Data
9. Mean
10. Median
11. Range
12. First Quartile (Lower Quartile; Q1)
13. Third Quartile (Upper Quartile; Q3)
19. Skewed Right
20. Skewed Left
21. Bar Chart/Graph
22. Histogram
23. Dot Plot
24. Stem and Leaf Plot
25. Box (and Whisker) Plot
26. Parallel Box Plots

Part 2 – Practicing with Data

Use your vocabulary words from Part 1 to help complete these problems.

I. Determine if the variables listed below are *quantitative* or *categorical*.

a. Time it takes to get to school _____

b. Hair color _____

c. Teacher salaries _____

d. Gender _____

e. Height _____

f. Age of Grammy winners _____

g. Type of headache medicine _____

h. Jellybean flavors _____

i. Country of origin _____

j. Number of shoes owned _____

II. A *statistic* is a number calculated from data. Determine the following statistics of the data below on the number of homeruns Chipper Jones has hit in each baseball season from 1995 to 2010.

Number of Homeruns: 23 30 21 34 45 36 38 26 27 30 21 26 29 22 18 10

a. Mean: _____

b. Minimum: _____

c. Lower Quartile (Q1): _____

d. Median: _____

e. Upper Quartile (Q3): _____

f. Maximum: _____

g. Range _____

h. IQR: _____

i. In the space provided to the right, make a Box and Whisker Plot using your answers above.

III. **Shopping Spree!** A marketing consultant observed 50 consecutive shoppers at a supermarket. The variable of interest was how much each shopper spent in the store. Here are the data arranged in increasing order (rounded to the nearest dollar).

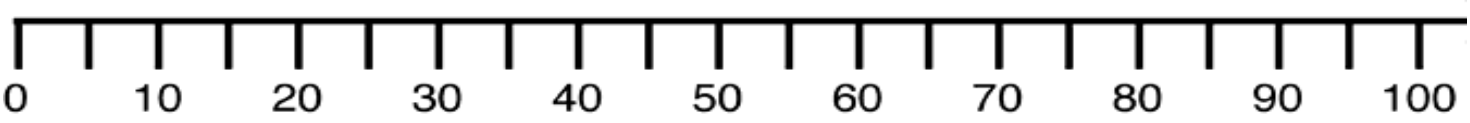
3	9	9	11	13	14	15	16	17	17
18	18	19	20	20	20	21	22	23	24
25	25	26	26	28	28	28	28	32	35
36	39	39	41	43	44	45	45	47	49
50	53	55	59	61	70	83	86	86	93

a. Make a Stem and Leaf Plot using tens of dollars as the *stem* and dollars as the *leaves*. (See the key below to get started)

Key: Stem | Leaf

1 | 2 = $12

b. Make a Dot Plot using the data above. Use the given number line below.



IV. **Accidental Deaths:** In 1997, there were 92,353 accidental deaths in the United States. Among these deaths, 42,340 were from motor vehicle accidents, 11,858 from falls, 10,163 from poisoning, 4051 from drowning, and 3601 from fires. The rest were listed as "other" causes.

a. Find the **percent** of accidental deaths from each of these causes, rounded to the nearest percent.

Motor Vehicle Accidents: _____

Falls: _____

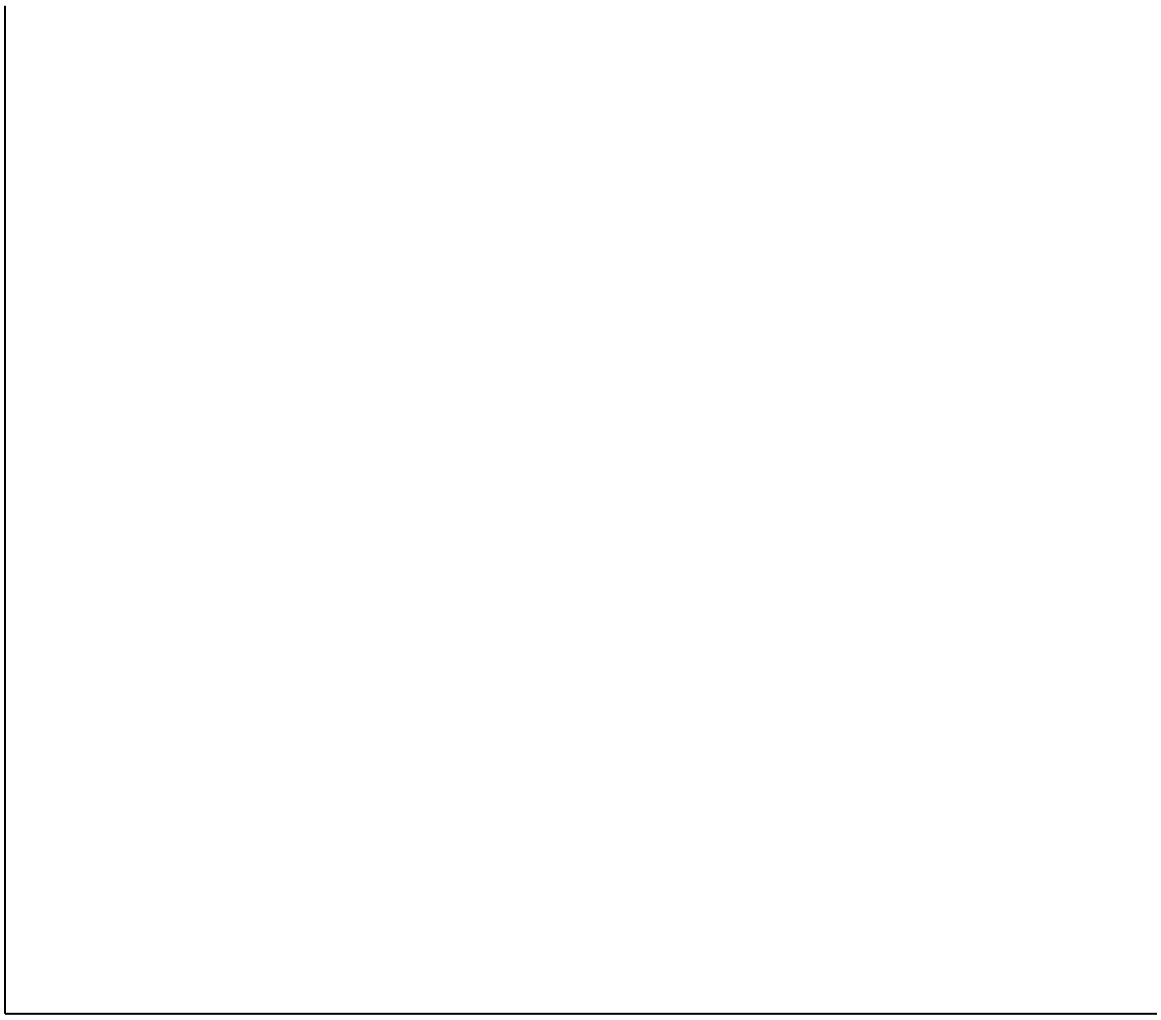
Poisoning: _____

Drowning: _____

Fires: _____

Other: _____

b. Create a well-labeled **bar graph** of the distribution of causes of accidental deaths. Be sure to include an "other causes" bar.



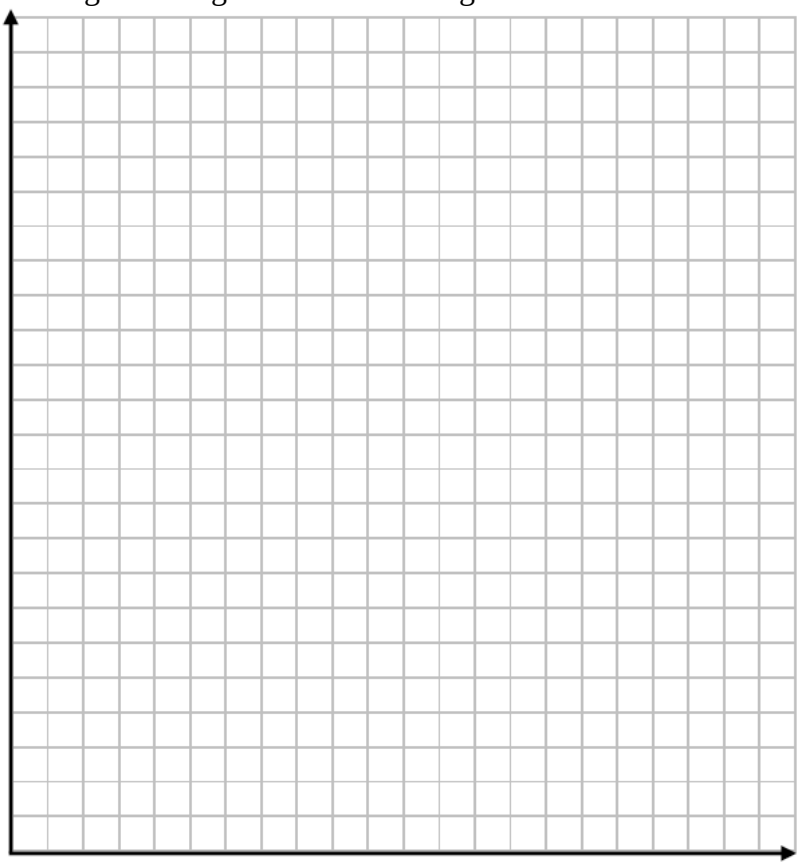
V. **Where Do Older Folks Live?** This table gives the percentage of residents aged 65 or older in each of the 50 states.

State	Percent	State	Percent	State	Percent
Alabama	13.1	Louisiana	11.5	Ohio	13.4
Alaska	5.5	Maine	14.1	Oklahoma	13.4
Arizona	13.2	Maryland	11.5	Oregon	13.2
Arkansas	14.3	Massachusetts	14.0	Pennsylvania	15.9
California	11.1	Michigan	12.5	Rhode Island	15.6
Colorado	10.1	Minnesota	12.3	South Carolina	12.2
Connecticut	14.3	Mississippi	12.2	South Dakota	14.3
Delaware	13.0	Missouri	13.7	Tennessee	12.5
Florida	18.3	Montana	13.3	Texas	10.1
Georgia	9.9	Nebraska	13.8	Utah	8.8
Hawaii	13.3	Nevada	11.5	Vermont	12.3
Idaho	11.3	New Hampshire	12.0	Virginia	11.3
Illinois	12.4	New Jersey	13.6	Washington	11.5
Indiana	12.5	New Mexico	11.4	West Virginia	15.2
Iowa	15.1	New York	13.3	Wisconsin	13.2
Kansas	13.5	North Carolina	12.5	Wyoming	11.5
Kentucky	12.5	North Dakota	14.4		

Histograms are a way to display groups of quantitative data into bins (the bars). These bins have the same width and scale and are touching because the number line is continuous. To make a histogram you must first decide on an appropriate bin width and count how many observations are in each bin. The bins for percentage of residents aged 65 or older have been started below for you.

a. Finish the chart of Bin widths and then create a histogram using those bins on the grid below. Make sure you include appropriate labels, title and scale.

Bin Widths	Frequency
4 to < 6	1
6 to < 8	
8 to < 10	



VI. **SSHA Scores:** The following are scores from the Survey of Study Habits and Attitudes (SSHA) for 18 first-year college women and for 20 first-year college men:

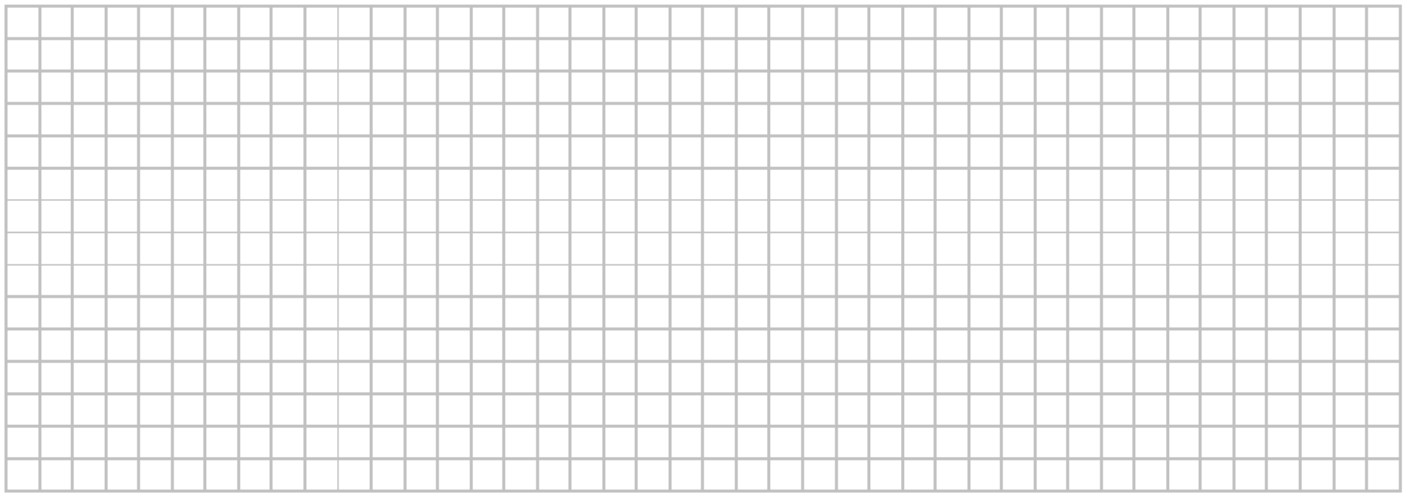
Women: 154 109 137 115 152 140 154 178 101 103 126 126 137 165 165 129 200 148

Men: 108 140 114 91 180 115 126 92 169 146 109 132 75 88 113 151 70 115 187 104

a. Put the data values in order for each gender. Compute numeral summaries for each gender:

	Women	**Men**
Mean		
Minimum		
Q1		
Median		
Q3		
Maximum		
Range		
IQR		

b. Using the minimum, Q1, Median, Q3, and Maximum from each gender, make parallel boxplots to compare the distributions.



Part 3 – More Prerequisite Practice

The prerequisite for AP Statistics is Math III or Accelerated Math II. You will not find very much equation solving in this course, but some quick review of Math I, Math II, and Math III content will be helpful.

I. **Here is a formula that is used often in AP Statistics:** $z = \frac{x - \bar{x}}{s}$

1. If z = 2.5, x = 102 and $\bar{x}$ = 100, what is s? Show your work.

2. If z = -3.35, x = 60, and s = 4, what is $\bar{x}$? Show your work.

II. **It is expected that you have thorough understanding of <u>linear functions</u>:**

<u>Example:</u> The USDA reported that in 1990 each person in the United States consumed an average of 133 pounds of natural sweeteners. They also claim this amount has decreased by about 0.6 pounds each year.

a. Write a linear equation that relates years since 1990 to the average consumption of natural sweeteners. Define your variables.

y = -0.6x + 133 ***y = pounds of sweeteners consumed***
x = # years since 1990

b. What is the slope and what is the y-intercept?

Slope = -0.6. As the year increases by 1, sweetener consumption decreases by 0.6 pounds
Intercept = 133. At year x=0 (or 1990), the amount of sweetener consumption was 133 pounds

c. Predict the average consumption of sweeteners per person for the year 2005.
y = -0.6 (15) + 133 = 124 pounds per person

<u>Now You Try!</u>

1. Sally makes $30 a day at her summer job. For every hour she works past the required work period, she makes an additional $5.

   a. Write a linear equation that relates the number of extra hours she works a day to her total paycheck for the day. Define your variables.

   b. What are the slope and the y-intercept?

   c. Predict Sally's paycheck for the day if she works 8.5 extra hours in a day.

2. The following equation can be used to predict the average height of boys anywhere between birth and 15 years old: y = 2.79x +25.64 where *x* is the age (in years) and *y* is the height (in inches).

   a. What does the slope represent in this problem? Interpret it in context.

   b. What does the y-intercept represent in this problem? Interpret it in context.

   c. Using the equation, predict a boy's height when he is 12 years old.

III. **You are expected to have a basic understanding of <u>probability</u>:**

1. A special lottery is to be held to select the student who will live in the only deluxe room in a dormitory. There are 100 seniors, 150 juniors, and 200 sophomores who applied. Each senior's name is placed in the lottery 3 times; each junior's name, 2 times; and each sophomore's name, 1 time. What is the probability that a senior's name will be chosen?

   A. $\frac{1}{8}$ B. $\frac{2}{9}$ C. $\frac{2}{7}$ D. $\frac{3}{8}$ E. $\frac{1}{2}$

2. Which of the following has a probability closest to 0.5?
   A. The sun will rise tomorrow.
   B. It will rain tomorrow.
   C. You will see a dog with only three legs when you leave the room.
   D. A fair die will come up with a score of 6 four times in a row.
   E. There will be a plane crash somewhere in the world within the next five minutes.

3. If a coin is tossed twice, what is the probability that on the first toss the coin lands heads and on the second toss the coin lands tails?

   A. 1/6 B. 1/3 C. ¼ D. ½ E. 1

4. If a coin is tossed twice what is the probability that it will land either heads both times or tails both times?

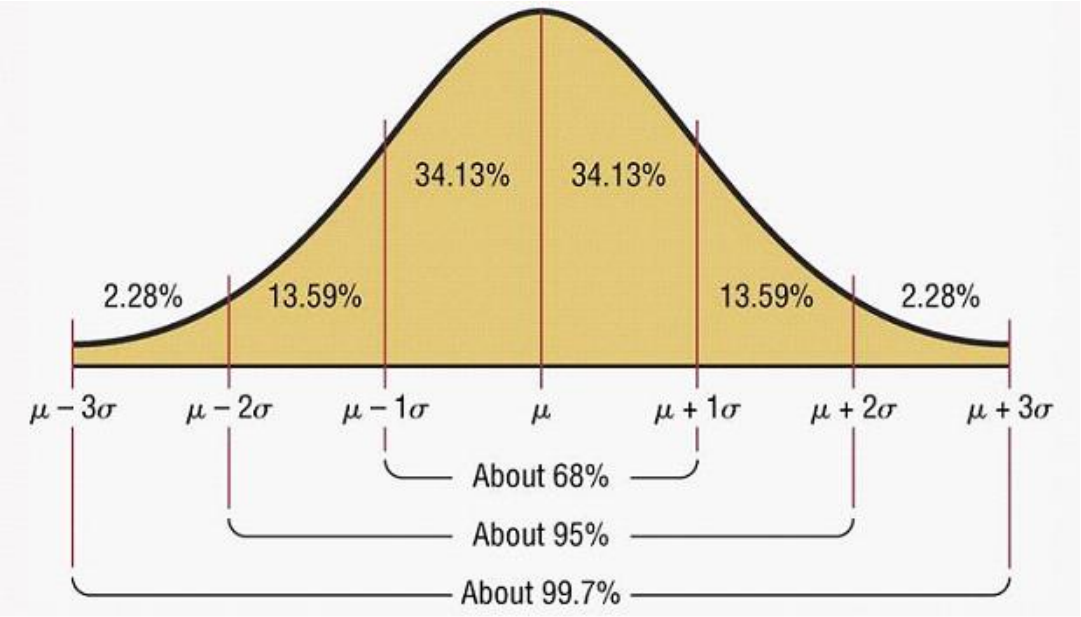
   A. 1/8 B. 1/6 C. 1/4 D. 1/2 E. 1

5. Calculate the following probabilities and arrange them in order from least to greatest.
   a. The probability that a fair die will produce an even number. _____
   b. A random digit from 1 to 9 (inclusive) is chosen, with all digits being equally likely. The probability that when it's squared it will end with the digit 1. _____
   c. The probability that a letter chosen from the alphabet will be a vowel. _____
   d. A random number between 1 and 20 (inclusive) is chosen. The probability that its square root will not be an integer. _____

   ORDER: _____, _____, _____, _____

IV. **You are expected to have a basic understanding of <u>The Normal Distribution</u>:**

To complete these problems, use the graphic below.



The life span of a certain light bulb is normally distributed with a mean of 670 hours and a standard deviation of 65 hours.

1. Between what two life spans will 68 % of the bulbs fall?

2. Between what two life spans will 95 % of the bulbs fall?

3. Between what two life spans will 99.7 % of the bulbs fall?

4. Estimate the probability that a randomly selected bulb will last between 605 and 735 hours.

5. What percentage of light bulbs last between 605 and 800 hours?

6. What percentage of light bulbs last between 475 and 800 hours?

7. What percentage of light bulbs last between 670 and 865 hours?

8. What percentage of light bulbs last less than 475 hours and more than 865 hours?