Dutchtown High School
AP Statistics Summer Assignment

Welcome to AP Statistics!

The purpose of this summer assignment is to:
   1. Refresh your knowledge on statistics
   2. Demonstrate your abilities to analyze data and determine conclusions

All of your work is attached with all instructions. Be sure to finish all work by the beginning of school in August as it will be collected and graded. There will also be an assessment on the materials in this packet within the first month of returning to school.

Assignment
   1. Reading and Vocabulary
         a. Go to www.stattrek.com
         b. Click on "AP Statistics" then "APtutorial"
         c. On the left side of the screen is a list of general topics. Under each general topic is a list of subtopics. You will read the following subtopics to complete the vocabulary list
               i. General Topic: Exploring Data
                     1. Subtopic: Variables
                     2. Subtopic: Population vs. Sample
                     3. Subtopic: Central Tendency
                     4. Subtopic: Variability
                     5. Subtopic: Position
               ii. General Topic: Charts and Graphs
                     1. Subtopic: Charts and Graphs
                     2. Patterns in Data
                     3. Dotplots
                     4. Histograms
                     5. Stemplots 6.
                     Boxplots
                     7. Scatterplots
                     8. Comparing Data sets
   2. Practice Problems 3.
   Review Problems

. If you have any questions or clarifications, email Mrs. Gilmore.

Part One: Reading and Vocabulary

Define each of the following terms. I encourage you to rephrase the definition in your terms in order to challenge your learning. When asked to provide a unique example or sketch of the word, do not provide one that is given from the website.

1. Categorical Variables

    a. Example:

2. Quantitative Variables

    a. Example:

3. Discrete Variables

4. Continuous

5. Univariate Data

6. Bivariate Data

7. Population

    a. Example:

8. Sample

    a. Example:

9. Median

10. Mean

    a. Formula:

11. Outlier

12. Parameter

13. Statistics

14. Range

15. Standard Score (z-score)

    a. Formula:

16. Center

17. Spread

18. Variance

    a. Formula:

19. Standard Deviation

    a. Formula:

20. Symmetry

    a. Sketch:

21. Unimodal

    a. Sketch:

22. Bimodal

    a. Sketch:

23. Skewness

    a. Sketch Skewed left:

    b. Sketch Skewed right:

24. Uniform

    a. Sketch:

25. Gaps

    a. Sketch:

26. Outliers

    a. Sketch:

27. Dot plots

28. Bar Chart

29. Histogram

30. Difference between bar chart and histogram

31. Stemplots

32. Boxplots

33. Quartiles

34. Range

35. Interquartile Range

36. Four Ways to Describe Data Sets

37. Types of Graphs that can used for comparing data

Part Two: Practice Problems

1. Categorical or Quantitative: Determine if the variables listed below are quantitative or categorical
   a. Time it takes to get to school

   b. Number of people under 18 living in a household

   c. Hair color

   d. Temperature of a cup of coffee

   e. Teacher salaries

   f. Gender

   g. Smoking

   h. Height

   i. Amount of oil spilled

   j. Age of Oscar winners

   k. Type of depression medication

   l. Jellybean flavors

   m. Country of origin

   n. Type of meat

   o. Number of shoes owned

2. Accidental Deaths
   a. In 1997, there were 92,353 deaths from accidents in the United States. Amount these were 42,340 deaths 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.
      i. Find the percent of accidental deaths from each of these causes, rounded to the nearest percent.

      ii. What percent of accidental deaths were from "other causes"?

iii. Neatly create a well-labeled bar graph of the distribution of causes of accidental deaths. Be sure to include an "other causes"bar.
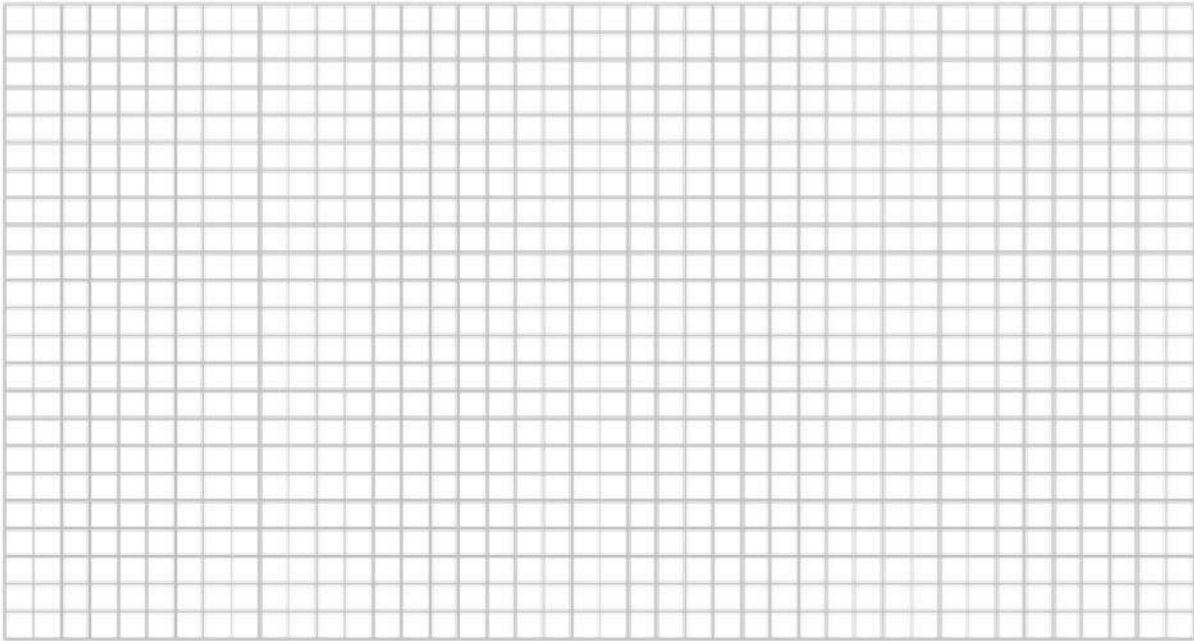
iv. A pie chart is another graphical display used to show all categories in a categorical variable relative to each other. Create a pie chart for the accidental death percentages. You may try using a software or internet source to make one and paste in the space below.

3. Shopping Spree
   a. A marketing consultant observed 50 consecutive shoppers at a supermarket. One variable of interest was how much each shopper spent in the store. Here is the data (round to the nearest dollar), arranged in increasing order.

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

   b. Make a stemplot using tens of dollars as the stem and dollars as the leaves. Make sure you include appropriate labels.

4. Where do older folks live?

a. 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		

b. 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 54 or order have been started below for you.

c. Finish the chart of bin widths and then create a histogram using those bins on the grid below. Be sure to include appropriate labels.

Bin Widths	Frequency
4 to < 6	1
6 to < 8	
8 to < 10	
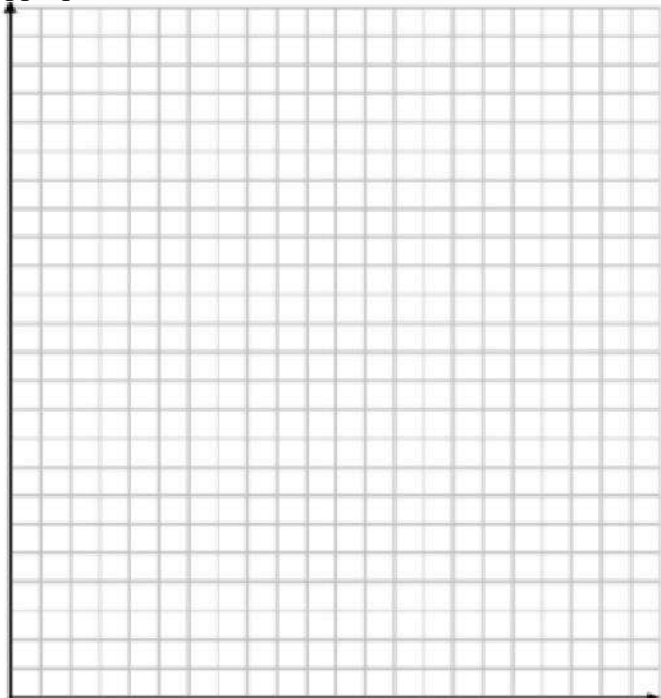
Part Three: Review Problems

1. Here is a formula that is used often in AP statistics: $z = \dfrac{x - \bar{x}}{s}$. Use your algebra skills...
   a. If z = 2.5, x = 102, and x = 100, what is s? Show your work

   b. If z = -3.35, x = 60, and s = 4, what is x? Show your work.

2. 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. If 1990 could be considered "year 0", which of the above numbers represents the slope and which represents the y-intercept?

   b. What is the equation of the line of best fit using the slope and y-intercept above?

   c. Predict the average consumption of sweeteners per person for the year 2005.

3. 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 the context of this problem/situation.

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

4. Hilary wonders if people of similar heights tend to date each other. She measures herself, her dormitory roommate, and the women in the adjoining rooms; then she measures the next man each woman dates. Here are the data (heights in inches):

Women	66	64	66	65	70	65
Men	72	68	70	68	74	69

a. Construct a scatterplot of the data.
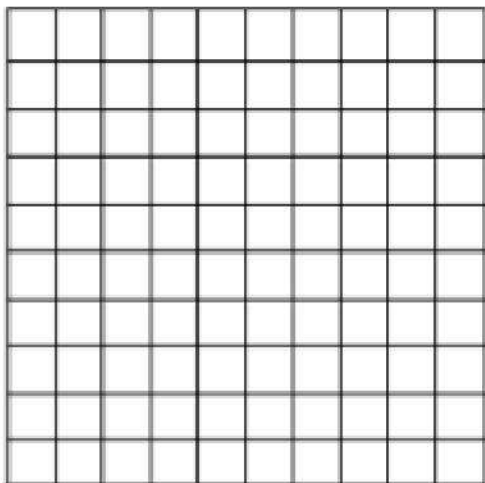
b. Describe the association between the heights of the women and the men they date.

5. 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. 1/8
   b. 2/9
   c. 2/7
   d. 3/8
   e. 1/2

6. 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.

7. 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? (Hint: What are the possible outcomes when you toss a coin twice?)

    a. A. 1/6
    b. B. 1/3
    c. C. ¼
    d. D. ½
    e. E. 1

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

    a. A. 1/8
    b. B. 1/6
    c. C. ¼
    d. D. 1/2
    e. E. 1

9. 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 the answer will contain 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. _____

    e.   **ORDER:____,_____,_____,____**