2021 – 2022 AP Statistics Summer Assignment

All questions or concerns related to this assignment should be directed to Mrs. Barragato on or before Tuesday June 22nd, 2021.

If any concerns should arise over the summer, please email both the teacher and math supervisor listed below:

Teacher: Mrs. Barragato – dbronder@dumontnj.org
Math Supervisor: Ms. Warnock – dwarnock@dumontnj.org

Due Date: Friday September 10th, 2021 (it is highly suggested to complete the assignment by the first day of school)

Weight of Assignment: 30-point homework assignment

What is expected:
- Join the "21 – 22 AP Stats Summer Assignment" Google Classroom page (use class code: 2v5v6q7)
- Review previously learned statistical concepts
- Practice previously learned statistical concepts

Resources and Supplies Needed:
- Textbook (you may pick one up if you want one, just email me to let me know; otherwise, textbook sections are posted in the Google Classroom)
- Online Resources: (all links are in the Google Classroom)
  - https://www.khanacademy.org/math/ap-statistics
  - https://www.varsitytutors.com/ap_statistics-help
  - https://stattrek.com/tutorials/ap-statistics-tutorial.aspx
- Graphing Calculator (a link to an online calculator is on the Google Classroom Page, but it HIGHLY SUGGESTED that you have one of your own to use)

Overview:
In this summer assignment, you will be working on completing the material in chapter 1 on your own. This chapter is a review of previously learned statistics topics. Completing this chapter early will save us time during the year for AP Exam Review. Chapter 1 has an Introduction and 3 Sections. There are guided notes on the Google Classroom for your use, but it is NOT required that you fill them out. After each section there will be a set of assigned homework problems to complete on a separate piece of paper. We will review the material before taking a quiz the first full week of school. Each section's assignment is worth 10 points towards the grade on this assignment.

Have a great summer and I am looking forward to a great year!

~Mrs. Barragato

Introduction Homework Assignment:

1. How can we help wood surfaces resist weathering, especially when restoring historic wooden buildings? In a study of this question, researchers prepared wooden panels and then exposed them to the weather. Here are some of the variables recorded: type of wood (yellow poplar, pine, cedar); type of water repellent (solvent-based, water-based); paint thickness (millimeters); paint color (white, gray, light blue); weathering time (months). Identify each variable as categorical or quantitative.

2. Here is a small part of the data set that describes the students in an AP® Statistics class. The data come from anonymous responses to a questionnaire filled out on the first day of class.

Gender	Hand	Height (in.)	Homework time (min)	Favorite music	Pocket change (cents)
F	L	65	200	Hip-hop	50
M	L	72	30	Country	35
M	R	62	95	Rock	35
F	L	64	120	Alternative	0
M	R	63	220	Hip-hop	0
F	R	58	60	Alternative	76
F	R	67	150	Rock	215

   a. What individuals does this data set describe?
   b. What variables were measured? Identify each as categorical or quantitative.
   c. Describe the individual in the highlighted row.

For #'s 3 and 4 refer to the following: At the Census Bureau Web site **www.census.gov**, you can view detailed data collected by the American Community Survey. The following table includes data for 10 people chosen at random from the more than 1 million people in households contacted by the survey. "School" gives the highest level of education completed.

Weight (lb)	Age (yr)	Travel to work (min)	School	Gender	Income last year ($)
187	66	0	Ninth grade	1	24,000
158	66	n/a	High school grad	2	0
176	54	10	Assoc. degree	2	11,900
339	37	10	Assoc. degree	1	6000
91	27	10	Some college	2	30,000
155	18	n/a	High school grad	2	0
213	38	15	Master's degree	2	125,000
194	40	0	High school grad	1	800
221	18	20	High school grad	1	2500
193	11	n/a	Fifth grade	1	0

3. The individuals in this data set are:
   a. Households
   b. People
   c. Adults
   d. 120 variables
   e. Columns

4. This data set contains
   a. 7 variables, 2 of which are categorical
   b. 7 variables, 1 of which is categorical
   c. 6 variables, 2 of which are categorical
   d. 6 variables, 1 of which is categorical
   e. None of these

Section 1.1 Homework Assignment

1. Births are not evenly distributed across the days of the week. Here are the average numbers of babies born on each day of the week in the United States in a recent year

Day	Births
Sunday	7374
Monday	11,704
Tuesday	13,169
Wednesday	13,038
Thursday	13,013
Friday	12,664
Saturday	8459

   a. Present these data in a well-labeled bar graph. Would it also be correct to make a pie chart?
   b. Suggest some possible reason why there are fewer births on weekends

2. Young people are more likely than older folk to buy music online. Here are the percents of people in several age groups who bought music online in a recent year

Age group	Bought music online
12 to 17 years	24%
18 to 24 years	21%
25 to 34 years	20%
35 to 44 years	16%
45 to 54 years	10%
55 to 64 years	3%
65 years and over	1%

   a. Explain why it is *not* correct to use a pie chart to display these data.
   b. Make a bar graph of the data. Be sure to label your axes.

3. Students in a high school statistics class were given data about the main method of transportation to school for a group of 30 students. They produced the pictograph shown.

Key:
- 🚴 = 1 Cycler
- 🚗 = 2 Cars
- 🚌 = 7 Bus takers
- 🚶 = 2 Walkers
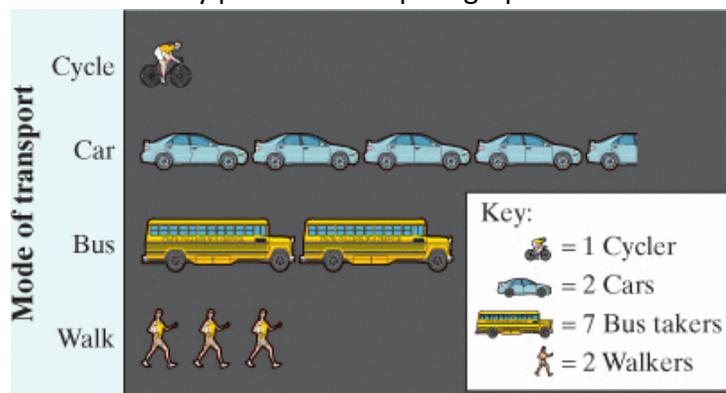
a. How is this graph misleading?
b. Make a new graph that isn't misleading.
4. Recycling is supposed to save resources. Some people think recycled products are lower in quality than other products, a fact that makes recycling less practical. People who use a recycled product may have different opinions from those who don't use it. Here are data on attitudes toward coffee filters made of recycled paper from a sample of people who do and don't buy these filters

Think quality is	Buy recycled filters?	
	Yes	No
Higher	20	29
The same	7	25
Lower	9	43

a. How many people does this table describe? How many of these were buyers of coffee filters made of recycled paper?
b. Give the marginal distribution (in percents) of opinion about the quality of recycled filters. What percent of the people in the sample think the quality of the recycled product is the same or higher than the quality of other filters?
c. To see the relationship between opinion and experience with the product, find the conditional distributions of opinion (the response variable) for buyers and nonbuyers. What do you conclude?
5. Yellowstone National Park surveyed a random sample of 1526 winter visitors to the park. They asked each person whether they owned, rented, or had never used a snowmobile. Respondents were also asked whether they belonged to an environmental organization (like the Sierra Club). The two-way table summarizes the survey responses.

|	Environmental Club		
	No	Yes	Total
Never used	445	212	657
Snowmobile renter	497	77	574
Snowmobile owner	279	16	295
Total	1221	305	1526

Do these data suggest that there is an association between environmental club membership and snowmobile use among visitors to Yellowstone National Park? Give appropriate evidence to support your answer.

For #'s 6-10 refer to the following information: The National Survey of Adolescent Health interviewed several thousand teens (grades 7 to 12). One question asked was "What do you think are the chances you will be married in the next ten years?" Here is a two-way table of the responses by gender

	Female	Male
Almost no chance	119	103
Some chance, but probably not	150	171
A 50-50 chance	447	512
A good chance	735	710
Almost certain	1174	756

6. The percent of females among the respondents was
   a. 2625
   b. 4877
   c. About 46%
   d. About 54%
   e. None of these

7. Your percent from the previous exercise is part of
   a. the marginal distribution of females.
   b. the marginal distribution of gender.
   c. the marginal distribution of opinion about marriage.
   d. the conditional distribution of gender among adolescents with a given opinion.
   e. the conditional distribution of opinion among adolescents of a given gender.

8. What percent of females thought that they were almost certain to be married in the next ten years?
   a. About 16%
   b. About 24%
   c. About 40%
   d. About 45%
   e. About 61%

9. Your percent from the previous exercise is part of
   a. the marginal distribution of gender.
   b. the marginal distribution of opinion about marriage.
   c. the conditional distribution of gender among adolescents with a given opinion.
   d. the conditional distribution of opinion among adolescents of a given gender.
   e. the conditional distribution of "Almost certain" among females.

10. The following bar graph shows the distribution of favorite subjects for a sample of 1000 students. What is the most serious problem with the graph?
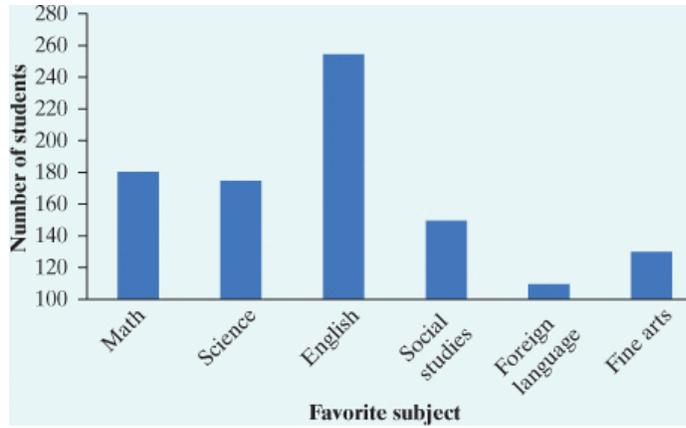
a. The subjects are not listed in the correct order.
b. This distribution should be displayed with a pie chart.
c. The vertical axis should show the percent of students.
d. The vertical axis should start at 0 rather than 100.
e. The foreign language bar should be broken up by language.

Section 1.2 Homework Assignment
1. Students in a college statistics class responded to a survey designed by their teacher. One of the survey questions was "How much sleep did you get last night?" Here are the data (in hours):

9	6	8	6	8	8	6	6.5	6	7	9	4	3	4
5	6	11	6	3	6	6	10	7	8	4.5	9	7	7

   a. Make a dotplot to display the data.
   b. Describe the overall pattern of the distribution and any departures from that pattern.

2. Below is a stemplot of the percent of residents aged 25 to 34 in each of the 50 states. As in the stemplot for older residents (**page 33**), the stems are whole percents, and the leaves are tenths of a percent. This time, each stem has been split in two, with values having leaves 0 through 4 placed on one stem, and values ending in 5 through 9 placed on another stem.

```
11 | 44
11 | 66778
12 | 0134
12 | 666778888
13 | 0000001111444
13 | 7788999
14 | 0044
14 | 567
15 | 11
15 |
16 | 0
```

   a. Why did we split stems?
   b. Give an appropriate key for this stemplot.
   c. Describe the shape, center, and spread of the distribution. Are there any outliers?

3. We asked the students in a large first-year college class how many minutes they studied on a typical weeknight. Here are the responses of random samples of 30 women and 30 men from the class:

Women					Men				
180	120	180	360	240	90	120	30	90	200
120	180	120	240	170	90	45	30	120	75
150	120	180	180	150	150	120	60	240	300
200	150	180	150	180	240	60	120	60	30
120	60	120	180	180	30	230	120	95	150
90	240	180	115	120	0	200	120	120	180

   a. Examine the data. Why are you not surprised that most responses are multiples of 10 minutes? Are there any responses you consider suspicious?
   b. Make a back-to-back stemplot to compare the two samples. Does it appear that women study more than men (or at least claim that they do)? Justify your answer.

4. How long do people travel each day to get to work? The following table gives the average travel times to work (in minutes) for workers in each state and the District of Columbia who are at least 16 years old and don't work at home

|					
AL	23.6	LA	25.1	OH	22.1
AK	17.7	ME	22.3	OK	20.0
AZ	25.0	MD	30.6	OR	21.8
AR	20.7	MA	26.6	PA	25.0
CA	26.8	MI	23.4	RI	22.3
CO	23.9	MN	22.0	SC	22.9
CT	24.1	MS	24.0	SD	15.9
DE	23.6	MO	22.9	TN	23.5
FL	25.9	MT	17.6	TX	24.6
GA	27.3	NE	17.7	UT	20.8
HI	25.5	NV	24.2	VT	21.2
ID	20.1	NH	24.6	VA	26.9
IL	27.9	NJ	29.1	WA	25.2
IN	22.3	NM	20.9	WV	25.6
IA	18.2	NY	30.9	WI	20.8
KS	18.5	NC	23.4	WY	17.9
KY	22.4	ND	15.5	DC	29.2

a. Make a histogram of the travel times using classes of width 2 minutes, starting at 14 minutes. That is, the first class is 14 to 16 minutes, the second is 16 to 18 minutes, and so on.
b. The shape of the distribution is a bit irregular. Is it closer to symmetric or skewed? Describe the center and spread of the distribution. Are there any outliers?

5. Here are the amounts of money (cents) in coins carried by 10 students in a statistics class: 50, 35, 0, 97, 76, 0, 0, 87, 23, 65. To make a stemplot of these data, you would use stems
   a. 0, 1, 2, 3, 4, 5, 6, 7, 8, 9.
   b. 0, 2, 3, 5, 6, 7, 8, 9.
   c. 0, 3, 5, 6, 7.
   d. 00, 10, 20, 30, 40, 50, 60, 70, 80, 90.
   e. None of these.

6. The histogram below shows the heights of 300 randomly selected high school students. Which of the following is the best description of the shape of the distribution of heights?
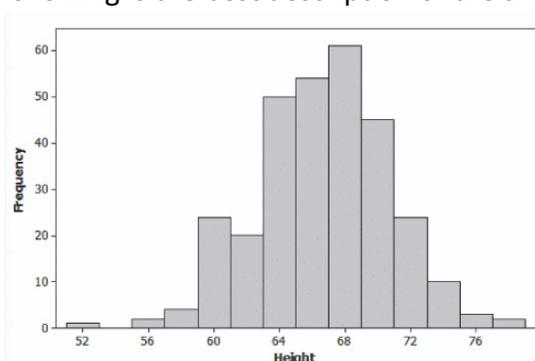
   a. Roughly symmetric and unimodal
   b. Roughly symmetric and bimodal
   c. Roughly symmetric and multimodal
   d. Skewed to the left

    e. Skewed to the right

7. You look at real estate ads for houses in Naples, Florida. There are many houses ranging from $200,000 to $500,000 in price. The few houses on the water, however, have prices up to $15 million. The distribution of house prices will be
    a. skewed to the left.
    b. roughly symmetric.
    c. skewed to the right.
    d. unimodal.
    e. too high.

8. The following histogram shows the distribution of the percents of women aged 15 and over who have never married in each of the 50 states and the District of Columbia. Which of the following statements about the histogram is correct?
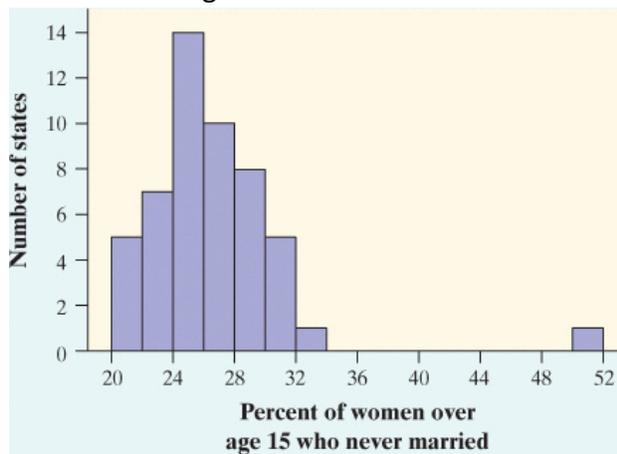
    a. The center of the distribution is about 36%.
    b. There are more states with percents above 32 than there are states with percents less than 24.
    c. It would be better if the values from 34 to 50 were deleted on the horizontal axis so there wouldn't be a large gap.
    d. There was one state with a value of exactly 33%.
    e. About half of the states had percents between 24% and 28%.

9. When comparing two distributions, it would be best to use relative frequency histograms rather than frequency histograms when
    a. the distributions have different shapes.
    b. the distributions have different spreads.
    c. the distributions have different centers.
    d. the distributions have different numbers of observations.
    e. at least one of the distributions has outliers.

10. Which of the following is the best reason for choosing a stemplot rather than a histogram to display the distribution of a quantitative variable?
    a. Stemplots allow you to split stems; histograms don't.
    b. Stemplots allow you to see the values of individual observations.
    c. Stemplots are better for displaying very large sets of data.
    d. Stemplots never require rounding of values.
    e. Stemplots make it easier to determine the shape of a distribution.