AP Statistics Summer Assignment 2020

Welcome to AP Statistics!! The purpose of this assignment is to get you familiar with some terminology and make you comfortable exploring and analyzing data. It will take several hours to properly complete, so please allow yourself a few days to do it and don't wait until the last minute! Successful completion of this assignment will help us "hit the ground running" and ensure adequate time in 2nd semester for proper review leading up to the AP Exam.

This Summer Assignment will consist of 3 parts: Videos, Reading/Vocabulary, and Calculator Practice (sent/posted separately).

This assignment will be due on the 1st day of class. It consists of 3 parts

1. Videos: Please go to www.apstatsguy.com and watch the videos named "Summer Video One", "Summer Video Two", "Summer Video Three", and "Summer Video Four". You will need to take notes or write a BRIEF summary for these videos. These will be due on the first day of school and will be worth 20% of your summer assignment grade.
2. Vocabulary: There will be a quiz on vocabulary (and possibly information from the video) on the first day as well. I recommend making flashcards, but if you don't like flashcards, click on "Summer Vocab" link on www.apstatsguy.com, and download the BRAINSCAPE app. Upload the AP STATS VOCAB and use the app to help you study. This app is the most efficient flashcard app available. On the first day of school, you can either turn in your flashcards or show me your mastery level on the BRAINSCAPE app. This will be worth the remaining 80% of your summer assignment grade.
3. The third part of the summer assignment is not for a grade but is HIGHLY recommended. You need to become familiar with your calculator. A TI-84 is preferred and is an investment you will not regret. This calculator will be your BFF in AP Statistics. Please take the time to go though the calculator practice in a separate document.

Also, there are a few specific supplies you will need for class on Day 1:
- Graphing calculator (TI-84 is preferred).
- One-inch (or bigger) 3 ring binder
- Index cards for vocabulary
- Colored pens/pencils

I will check emails periodically throughout the summer. If you have questions, please email, but beware that it may take several days for me to get back to you. ksumner@williamANDreed.com

Have a great summer!!!

Mrs. Sumner

Vocabulary:

1	What is Statistics?	The study of variability
2	What is variability?	Differences, how things differ. There is variability everywhere. We all look different, act different, have different preferences. Statisticians look at these differences.
3	What are 2 branches of AP STATS?	Inferential and Descriptive
4	What are DESCRIPTIVE STATS?	Tell me what you got! Describe to me the data that you collected, use pictures or summaries like mean, median, range, etc...
5	What are INFERENTIAL STATS?	Look at your data, and use that to say stuff about the BIG PICTURE.., like tasting soup... a little sample can tell you a lot about the big pot of soup (the population)
6	Compare Descriptive and Inferential STATS	Descriptive explains about the data that you have; inference uses that data - you have to try to say something about an entire population
7	What is data?	Any collected information. Generally, each little measurement... Like, if it is a survey about liking porridge.., the data might be "yes, yes, no, yes, yes" if it is the number of saltines someone can eat in 30 seconds, the data might be "3, 1, 2, 1, 4,3, 3, 4"
8	What is a population?	the group you're interested in. Sometimes it's big, like "all teenagers in the US" other times it is small, like "all AP Stats students in my school"
9	What is a sample?	A subset of a population, often taken to make inferences about the population. We calculate statistics from samples.
10	Compare population to sample	Populations are generally large, and samples are small subsets of these. We take samples to make inferences about populations. We use statistics to estimate parameters.
11	Compare data to statistics	Data is each little bit of information collected from the subjects.... They are the INDIVIDUAL little things we collect. We summarize them by, for example, finding the mean of a group of data. If it is a sample, then we call that mean a "statistic" if we have data from each member of population, then that mean is called a "parameter"
12	Compare data to parameters	Data is each little bit of information collected from the subjects.... They are the INDIVIDUAL little things we collect... we summarize them by, for example, finding the mean of a group of data. If it is a sample, then we call that mean a "statistic" if we have data from each member of population, then that mean is called a "parameter"
13	What is a parameter?	A numerical summary of a population. Like a mean, median, range... of a population
14	What is a statistic?	A numerical summary of a sample. Like a mean, median, range... of a sample.

15	We are curious about the average wait time at a Dunkin Donuts drive through in your neighborhood. You randomly sample cars one afternoon and find the average wait time is 3.2 minutes. What is the population parameter? What is the statistic? What is the parameter of interest? What is the data?	The parameter is the true average wait time at that Dunkin Donuts. This is a number you don't have and will never know. The statistic is "3.2 minutes." It is the average of the data you collected. The parameter of interest is the same thing as the population parameter. In this case, it is the true average wait time of all cars. The data is the wait time of each individual car, so that would be like "3.8 min, 2.2 min, .8 min, 3 min". You take that data and find the average, that average is called a "statistic," and you use that to make an inference about the true parameter.
16	Compare DATA-STATISTIC-PARAMETER using categorical example	Data are individual measures... like meal preference: "taco, taco, pasta, taco, burger, burger, taco"... Statistics and Parameters are summaries. A statistic would be "42% of sample preferred tacos" and a parameter would be "42% of population preferred tacos."
17	Compare DATA-STATISTIC-PARAMETER using quantitative example	Data are individual measures, like how long a person can hold their breath: "45 sec, 64 sec, 32 sec, 68 sec." That is the raw data. Statistics and parameters are 17 summaries like "the average breath holding time in the sample was 52.4 seconds" and a parameter would be "the average breath holding time in the population was 52.4 seconds"
18	What is a census?	Like a sample of the entire population, you get information from every member of the population
19	Does a census make sense?	A census is ok for small populations (like Mrs. Sumner's students) but impossible if you want to survey "all US teens"
20	What is the difference between a parameter and a statistic?	BOTH ARE A SINGLE NUMBER SUMMARIZING A LARGER GROUP OF NUMBERS....But pppp parameters come from pppp populations... sss statistics come from ssss statistics.
21	If I take a random sample of 20 hamburgers from FIVE GUYS and count the number of pickles on a bunch of them... and one of them had 9 pickles, then the number 9 from that burger would be called _____?	a datum, or a data value.
22	If I take a random sample 20 hamburgers from FIVE GUYS and count the number of pickles on a bunch of them and the average number of pickles was 9.5, then 9.5 is considered a _____?	statistic. (it is a summary of a sample.)
23	If I take a random sample of 20 hamburgers from FIVE GUYS and count the number of pickles on a bunch of them... and I do this because I want to know the true average number of pickles on a burger at FIVE GUYS, the true average number of pickles is considered a _____?	parameter, a one number summary of the population. The truth. AKA the parameter of interest.
24	What is the difference between a sample and a census?	With a sample, you get information from a small part of the population. In a census, you get info from the entire population. You can get a parameter from a census, but only a statistic from a sample.

25	Use the following words in one sentence: population, parameter, census, sample, data, statistics, inference, population of interest	I was curious about a population parameter, but a census was too costly so I decided to choose a sample, collect some data, calculate a statistic and use that statistic to make an inference about the population parameter (aka the parameter of interest).
26	If you are tasting soup. Then the flavor of each individual thing in the spoon is the _____, the entire spoon is a _____. The flavor of all of that stuff together is like the _____ and you use that to _____ about the flavor of the entire pot of soup, which would be the _____	If you are tasting soup. Then the flavor of each individual thing in the spoon is DATA, the entire spoon is a SAMPLE. The flavor of all of that stuff together is like the STATISTIC, and you use that to MAKE AN INFERENCE about the flavor of the entire pot of soup, which would be the PARAMETER. Notice you are interested in the parameter to begin with… that is why you took a sample.
27	What are random variables?	If you randomly choose people from a list, then their hair color, height, weight and any other data collected from them can be considered random variables.
28	What is the difference between quantitative and categorical variables?	Quantitative variables are numerical measures, like height and IQ. Categorical are categories, like eye color and music preference
29	What is the difference between quantitative and categorical data?	The data is the actual gathered measurements. So, if it is eye color, then the data would look like this "blue, brown, brown, brown, blue, green, blue, brown... etc." The data from categorical variables are usually words, often it is simply "YES, YES, YES, NO, YES, NO" If it was weight, then the data would be quantitative like "125, 155, 223, 178, 222, etc.." The data from quantitative variables are numbers.
30	What is the difference between discrete and continuous variables?	Discrete can be counted, like "number of cars sold" they are generally integers (you wouldn't sell 9.3 cars), while continuous would be something like weight of a mouse... 4.344 oz.
31	What is a quantitative variable?	Quantitative variables are numeric like: Height, age, number of cars sold, SAT score
32	What is a categorical variable?	Qualitative variables are like categories: Blonde, Listens to Hip Hop, Female, yes, no... etc.
33	What do we sometimes call a categorical variable?	qualitative
34	What is quantitative data?	The actual numbers gathered from each subject. 211 pounds. 67 beats per minute.
35	What is categorical data?	The actual individual category from a subject, like "blue" or "female" or "sophomore"
36	What is a random sample?	When you choose a sample by rolling dice, choosing names from a hat, or other REAL RANDOMLY generated sample. Humans can't really do this well without the help of a calculator, cards, dice, or slips of paper.
37	What is frequency?	How often something comes up
38	data or datum?	datum is singular. Like "hey dude, come see this datum I got from this rat!" data is the plural. "hey look at all that data Edgar got from those chipmunks over there!"

39	What is a frequency distribution?	A table, or a chart, that shows how often certain values or categories occur in a data set.
40	What is meant by relative frequency?	The PERCENT of time something comes up (frequency/total)
41	How do you find relative frequency?	just divide frequency by TOTAL....
42	What is meant by cumulative frequency?	ADD up the frequencies as you go. Suppose you are selling 25 pieces of candy. You sell 10 the first hour, 5 the second, 3 the third and 7 in the last hour, the cumulative frequency would be 10, 15, 18, 25
43	Make a guess as to what relative cumulative frequency is...	It is the ADDED up PERCENTAGES. An example is selling candy, 25 pieces sold overall..., with 10 the first hour, 5 the second, 3 the third, and 7 the fourth hour, we'd take the cumulative frequencies, 10, 15, 18 and 25 and divide by the total giving cumulative percentages... .40, .60, .64, and 1.00. Relative cumulative frequencies always end at 100 percent.
44	What is the difference between a bar chart and a histogram?	bar charts are for categorical data (bars don't touch) and histograms are for quantitative data (bars touch)
45	What is the mean?	the old average we used to calculate. It is the balancing point of the histogram
46	What is the difference between a population mean and a sample mean?	Population mean is the mean of a population, it is a parameter, sample mean is a mean of a sample, so it is a statistic. We use sample statistics to make inferences about population parameters.
47	What symbols do we use for population mean and sample mean?	Mu (μ) for population mean (parameter), x-bar ($\bar{x}$) for sample mean (statistic)
48	How can you think about the mean and median to remember the difference when looking at a histogram?	mean is balancing point of histogram, median splits the area of the histogram in half.
49	What is the median?	the middlest number, it splits area in half (always in the POSITION (n+1)/2)
50	What is the mode?	the most common, or the peaks of a histogram. We often use mode with categorical data
51	When do we often use mode?	With categorical variables. For instance, to describe the average teenager's preference, we often speak of what "most" students chose, which is the mode. It is also tells the number of bumps in a histogram for quantitative data (unimodal, bimodal, etc...).
52	Why don't we always use the mean, we've been calculating it all of our life?	It is not RESILIENT, it is impacted by skewness and outliers
53	When we say "the average teenager" are we talking about mean, median or mode?	It depends, if we are talking height, it might be the mean, if we are talking about parental income, we'd probably use the median, if we were talking about music preference, we'd probably use the mode to talk about the average teenager.

54	What is a clear example of where the mean would change but median wouldn't? (this would show its resilience)	Imagine if we asked eight people how much money they had in their wallet. We found they had {1, 2, 2, 5, 5, 8, 8, 9}. The mean of this set is 5, and the median is also 5. You might say the average person in this group had 5 bucks." But imagine if one of them just got back from the casino, and instead it was (1, 2, 2, 5, 5, 8, 8, 90001, in this case, the median would still be 5, but the mean goes up to over 1000. Which number better describes the amount of money the average person in the group carries, 5 bucks or 1000 bucks? I think 5 is a better description of the average person in this group and the 9000 is simply an outlier.
55	How are mean, median and mode positioned in a skewed left histogram?	goes in that order from left to right. Mean-median-mode
56	How are mean, median and mode positioned in a skewed right histogram?	goes in the opposite order. Mode-median-mean
57	Who chases the tail?	The mean chases the tail, the mean chases the tail, high-ho the derry-oh the mean chases the tail.., and outliers
58	Is there a way to study these efficiently instead of just rereading them?	YES.. Go to APSTATSGUY.COM and click on the SUMMER VOCAB FLASHCARDS link. Make sure to open account at BRAINSCAPE.COM and then add this deck to your library. Follow the directions. RATE THE CARDS HONESTLY FOR SUPER RESULTS!!