AP Statistics
Summer Assignment 2020
American School Foundation

Welcome to Advanced Placement Statistics. This course is like no other mathematics course, as the emphasis is on your ability to **think, reason, explain, and support** your mathematics with writing.

To prepare for the AP Statistics course over the summer, you should:

- Complete the 2 sections attached below (P1: Definitions, P2: Solving)
- Buy a graphing calculator (TI-Nspire recommended, TI-84 and others allowed)
- Understand that AP Stats requires a lot of reading and writing, and that you must be committed to giving your best effort day in and day out.

This summer assignment will NOT be graded; however, you are responsible for all the material within.

In the first two weeks of your AP Statistics course, you will be assessed on these basic pre-requisite skills outlined in this packet. The assessment will count as a full test grade in your first semester average.

If you've read this far and are still interested in this challenging course with college-level material then we wish you best of luck!

Enjoy your summer, get prepared, and see you at the start of school!

Any questions, contact your AP Statistics teacher:

Mr. Rachelson — rachelsong@asf.edu.mx

Mr. Caldwell — us_math1@asf.edu.mx

AP Statistics Summer Assignment:

Directions (Part I): The following are some data sets and scenarios you will use on the problems in the assignment. But first you will have to read through each one and define some basic concepts used in statistics. Provide a simple one or two line definition for each of the *italicized & underlined* terms.

---

1. In a *study* designed to determine the *average* death age reported for the *population* of a major U.S. city, a statistician *randomly selected* 31 obituaries from the city's largest newspaper. The *sample* consists of 14 males and 17 females – their age of death is listed in the table below:

Male	Female		Male	Female		Male	Female
55	75		84	87		70	57
60	77		90	44		93	84
62	82		98	91			79
71	75		61	94			75
74	74		70	60			74
79	89		58	68			

2. We are interested in studying if there is any *correlation* between a person's height and their weight. The following data shows the height (inches) and the weight (pounds) for 15 *individuals*:

Ht	Wt	Ht	Wt	Ht	Wt
62	165	69	211	85	321
58	157	68	188	61	170
71	192	71	244	57	164
81	265	62	215	56	148
74	223	67	199	51	153

3. According to a press release & data on car thefts in 2002, the 1989 Toyota Camry (a 13-year old car in that year) was the most stolen car in 2002. Further, according to the data, the 1990 Camry, 1991 Camry, and 1988 Camry were also all among the top ten most stolen cars that year. The press release claims that the most compelling reason for these cars being stolen is for parts but does not discuss any *confounding variables*.

4. The following data was collected on the average life expectancy and number of residents per television set for a sample of 10 countries:

Country	Life Exp.	# per TV		Country	Life Exp.	# per TV
Angola	44	200		Mexico	72	6.6
Cambodia	49.5	177		Russia	69	3.2
China	70	8		S. Africa	64	11
France	78	2.6		U. States	75.5	1.3
Japan	78	1.8		Vietnam	65	29

The table below provides the summary *descriptive statistics* (*n*, *mean*, *variance*, *standard deviation*, *median*, *range*, *minimum*, *maximum*, *quartile 1* and *quartile 3*):

	N	Mean	Variance	St. Dev.	Median
Life Exp.	10	66.5	133.1116	11.5374	69.5
# per TV	10	44.05	5890.716	76.751	7.3

	Range	Min	Max	Q1	Q3
Life Exp.	34	44	78	64	75.5
# per TV	198.7	1.3	200	2.6	29

5. There is a simple game with the Florida Lottery where you select a three-digit number at the cost of $1. Each night, the winning number is selected by having ping-pong balls numbered 0 through 9 in three separate containers with one digit being selected from each. If your three digit *permutation* (not *combination*) matches, you win $500. The Florida Lottery Commission recently released a report on numbers people like to pick versus what comes up with the following thought:

"People tend to select numbers they feel comfortable with. Many people like to select triples (i.e., 000, 111, 222, etc.) as their pick-3 number. In the past ten years, the triple 000 has come up only four times as the winning number."

Directions (Part II): For each of the following, provide a complete solution to the problem described ("Complete" solutions include explanations / work; not just answers). Do not worry if you cannot **solve** something but do attempt to look up what to do and try – if you are unable to solve a problem, provide thoughts on what you think you should do.

---

1. Refer to the first data set (average age at death):

   (A) Construct one of the following displays of the data for **BOTH** men and women (two pictures):
   - Parallel box-and-whisker plots
   - Histograms
   - Back-to-back stem-and-leaf plots

   (B) Using your displays, describe the general shape & distribution of the data, the center of the data, and the spread of the data.

   (C) A statistician needs to interpret data in order to convey thoughts. Write a paragraph comparing and contrasting the data sets and provide insight into the similarities or differences for age of death for men and women in this particular city.

2. Refer to the second data set (height versus weight):

   (A) Construct a scatter plot of the data with height on the x-axis and weight on the y-axis (Clearly label the axes and provide a scale & title)

   (B) Using a ruler, sketch a line of best fit for the data. Estimate the slope for your line (with appropriate units / context) and estimate the weight of someone who is 65 inches tall.

3. Refer to the third scenario (Car theft):

   (A) The press release claims that these cars are stolen mostly for their parts. Can you think of any OTHER reasons that an older family car would be a prime target for thieves (Think about statistics – provide at least two statistical / mathematical reasons in your answer)?

   (B) Based on the article, someone claims that because there is a high correlation between the age of the car and its theft (i.e., as the car gets older, it becomes more likely that it becomes a target for theft), it's obvious that car thieves prefer older cars. Does the high correlation mean that the age of the car causes the car to become a target? Do car thieves really prefer older cars or is there something else occurring – what are the possible confounding variables?

4. Refer to the fourth data set (Life expectancy & TV):

   (A) What observations can you make about average life expectancy in these countries (Provide the statistics to support your answer)?
   (B) What observations can you make about the average number of people per TV in these countries (Provide the statistics to support your answer)?
   (C) A statistics student constructed a scatter plot of the data by placing the number of residents per TV on the *x*-axis and life expectancy on the *y*-axis. In doing so, there is a definite negative correlation and the student claims that there is a causal relationship – as the number of people per TV decreases, life expectancy increases. Is the student's claim accurate or could there be some underlying reason for this data (i.e., are there any confounding variables to consider)?

5. Refer to the fifth scenario (Lottery):

   (A) What is the probability of winning the pick-3 on a given day? What is the probability of losing the pick-3 on a given day?
   (B) Since you pay $1 to play the pick-3, we can calculate your expected winnings with the following formula:
   E(Winnings on a $1 bet) = (Amount lost x P(loss)) + (Amount won x P(Win))
   It appears that if I select the correct number that I win 500 to 1 but we know the lottery is actually making a profit. Calculate your expected winnings and interpret that value in the context of how much of each dollar the lottery keeps for itself.
   (C) The excerpt from the commission's report mentions that people tend to select numbers they feel comfortable with. If I were to play the pick-3, I would likely select the number 507 (my wife's birthday). How much more or less likely am I to win with 507 as opposed to a triple such as 222?
   (D) The excerpt also mentions that in the past ten years of the daily game, the triple 000 has only come up four times. Explain whether or not this is statistically significant (Hint: Consider the number of **days** in the ten years mentioned in the report).

If you are done with the summer assignment you should have the following completed:

(1) 21 terms and definitions for part I
(2) 14 problems / solutions for part II

Remember, this assignment will NOT be graded, however, you are responsible for the material within. In the first two weeks of your AP Statistics course, you will be assessed on these basic pre-requisite skills outlined in this packet. The assessment will count as a full test grade in your first semester average.