Welcome to AP Statistics!

In order to be ready to hit the ground running in September, you have some preliminary work to do to become familiar with descriptive statistics, types of data, and graphical displays.

In this class you will learn to describe and analyze sets of data and use that analysis to draw conclusions in context about the situation that gave the data. This course is not like any other math class you have ever taken. It is comprised of a combination of Math, English, and Science.

Communication skills are *essential*, and there is much more reading and writing than what you are used to in a math class. It is a very rewarding course and a very important one but can be quite difficult at times. Since it is an AP course, it is considered to be college-level. The mathematics required for this course may not be as difficult as in other advanced math courses, but some of the concepts can be very confusing. In addition, there is a great deal of material that we are expected to cover by the end of April, so you need to be committed to giving it your absolute best effort.

Preliminaries:

1. You must be competent with algebra concepts and skills.
2. You should be familiar with basic statistical plots such as box-and-whiskers plots, bar graphs, histograms, circle (pie graphs), and stem-and-leaf plots.
3. You must be willing to explain your answers. Just getting the correct number will not be sufficient for this course.
4. You must own your own calculator and bring it to class everyday! A TI-83/84 is an essential tool for this course, as it has many statistical features we can take advantage of.

For this summer assignment:

1. The vocabulary definitions for Part I must be typed up and submitted via email to jlee@mppanthers.org by 11:59 p.m. before the first day of school. This part will count as a quiz grade for the first marking period. There will also be a separate vocabulary quiz within the first week of school.
2. Part 2 of this assignment is due on the first day of class and will count as a test grade. Print out pages 4-10 of this packet and show your work on the pages. If you need extra sheets of paper, staple them to the back of the packet before submission. Each day late will result in a 10 point penalty. Assignments will not be accepted after 3 days.

I'm looking forward to an enjoyable 2023-2024 school year.

Have a great summer and I'll see you in September.

Ms. Lee

Name: AP Statistics - Lee

The summer assignment is composed of two parts.

1. Reading and Vocabulary: You will use a free online Statistical tutoring site that will give you information on variables and data displays. While reviewing information on the site you will be completing a vocabulary list (on preceding pages). Follow the steps below:
    - Go to www.stattrek.com
    - Click on "Tutorials"
    - Click on the AP Statistics tutorial.
    - A list of general topics will appear on the left hand side of the screen. When you click on the general topic a list of subtopics will appear. You will read or watch the video on the following subtopics to complete the vocabulary list.

General Topic: The Basics	
Subtopics:	Variables
	Populations vs. Samples
	Central tendency
General Topic: Charts and Graphs	
Subtopics:	Patterns in data
	Dotplots
	Histograms
	Stemplots
	Boxplots
	Cumulative plots
	Comparing data sets
General Topic: Categorical Data	
Subtopics:	One-way tables
	Two-way tables

2. Practice Problems. After reading all the material above, you should be able to complete the questions in the remaining pages of this packet. You may do so in the spaces provided.

Part I: Vocabulary List

Please define each of the following terms.

1. Categorical Variables
2. Quantitative Variables
3. Univariate Data
4. Bivariate Data
5. Median
6. Mean
7. Population
8. Sample
9. Center
10. Spread
11. Symmetry
12. Unimodal and Bimodal
13. Skewness
14. Uniform
15. Gaps
16. Outliers
17. Dotplots
18. Stemplots
19. Boxplots
20. Parallel Boxplots
21. Quartiles
22. Range
23. Interquartile Range
24. Difference between a frequency table and relative frequency table
25. Parameter
26. Statistic
27. Marginal Distribution
28. Conditional Distribution
29. What are the W's of data

Name: AP Statistics - Lee

Part 2: Practice Problems

Categorical or Quantitative
Determine if the variables listed below are categorical or quantitative.

1. Time it takes to get to school
2. Number of people under 18 living in a household
3. Hair color
4. Temperature of a cup of coffee
5. Teacher salaries
6. Gender
7. Smoking
8. Height
9. Amount of oil spilled
10. Age of Oscar winners
11. Type of depression medicine
12. Jellybean flavors
13. Country of origin
14. Type of meat
15. Number of shoes owned

Statistic – What is that?
A statistic is a number calculated from data. Quantitative data has many different statistics that can be calculated. Determine the given statistics from the data below on the number of homerooms Mark McGuire has hit in each season from 1982-2001.

70	52	22	49	3	32	58	39
39	65	42	29	9	32	9	33

16. Mean
17. Minimum
18. Maximum
19. Median
20. Q1
21. Q3
22. Range
23. IQR

Accidental Deaths

In 1997 there were 92,353 deaths from accidents in the United States. Among 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.

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

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

26. Neatly create a well-labeled bar graph of the distribution of causes of accidental deaths. Be sure to include an "other causes" bar.

Name: AP Statistics - Lee

It's a Twista

The data below gives the number of hurricanes that happened each year from 1944 through 2000 as reported by *Science* magazine.

3	2	1	4	3	7	2	3	3	2
5	2	2	4	2	2	6	0	2	5
1	3	1	0	3	2	1	0	1	2
3	2	1	2	2	2	3	1	1	1
3	0	1	3	2	1	2	1	1	0
5	6	1	3	5	3				

27. Make a dotplot to display the data. Make sure you include appropriate title, labels, and scale.

Shopping Spree!

A marketing consultant observed 50 consecutive shoppers at a supermarket. One variable of interest was how much each shopper spent in the store. Here are the data (rounded 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

28. Make a stemplot using tens of dollars as the stem and dollars as the leaves. Make sure you include appropriate title, labels, and key.

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 classes (the bars). These classes 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 class width and count how many observations are in each class. The classes for percentage of residents ages 65 or older have been started below for you.

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

Class Widths	Frequency
4 to <6	
6 to <8	
8 to <10	

Name: AP Statistics - Lee

SSHA Scores

Here are the scores on the Survey of Study Habits and Attitudes (SSHA) for 18 first-year college women:

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

and for 20 first-year college men:

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

30. Put the data in order for each gender. Compute numerical summaries for each gender.

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

31. Using the minimum, Q1, median, Q3, and maximum from each gender, make parallel boxplots to compare the distributions.

Read the following…

"Teen Automobile Crash Rates Are Higher When School Starts Earlier"

ScienceDaily (June 10, 2010) — Earlier school start times are associated with increased teenage car crash rates, according to a research abstract presented June 9, 2010, in San Antonio, Texas, at SLEEP 2010, the 24th annual meeting of the Associated Professional Sleep Societies LLC.

Results indicate that in 2008 the teen crash rate was about 41 percent higher in Virginia Beach, Va., where high school classes began at 7:20 a.m., than in adjacent Chesapeake, Va., where classes started more than an hour later at 8:40 a.m. There were 65.4 automobile crashes for every 1,000 teen drivers in Virginia Beach, and 46.2 crashes for every 1,000 teen drivers in Chesapeake.

"We were concerned that Virginia Beach teens might be sleep restricted due to their early rise times and that this could eventuate in an increased crash rate," said lead author Robert Vorona, MD, associate professor of internal medicine at Eastern Virginia Medical School in Norfolk, Va. "The study supported our hypothesis, but it is important to note that this is an association study and does not prove cause and effect."

The study involved data provided by the Virginia Department of Motor Vehicles. In Virginia Beach there were 12,916 drivers between 16 and 18 years of age in 2008, and these teen drivers were involved in 850 crashes. In Chesapeake there were 8,459 teen drivers and 394 automobile accidents. The researchers report that the two adjoining cities have similar demographics, including racial composition and per-capita income.

Answer the following questions regarding the above excerpt:

32. *Who* is being studied?

33. *What* about those individuals is being recorded / analyzed (i.e. what are the variables?)? Do you think the variables are categorical or quantitative in nature?

34. *When* was the data collected?

35. *Where* was the data collected (more accurately: what geographical area is associated with the data)?

36. *Why* do you think this data was collected and analyzed?

37. *How* was the data collected and analyzed? In other words, what methods were used?

38. Why do you think the authors of the study mentioned, "it is important to note that this is an association study and does not prove cause and effect?"