

POSNACK

S C H O O L

AP Statistics Summer Assignment

Dear Parents/Students,

Welcome to Advanced Placement Statistics. This course is like no other mathematics course in that the emphasis is placed on your ability to **think, reason, explain, and support** as opposed to performing rudimentary computations. You should be competent in basic algebra and will need to familiarize yourself with the following topics:

- **Descriptive Statistics:**
Mean, median, mode, variance, standard deviation, range, quartile 1, quartile 3, minimum, and maximum
- **Statistical Displays:**
Box-and-whisker plot, scatter plot, bar graph, histogram, circle (pie) graph, and stem-and-leaf plot
- **Elementary Probability & Logic**

1. Videos:

Summer video 1:

https://www.youtube.com/watch?feature=player_detailpage&v=XPmTISOdPJs

Summer video 2:

https://www.youtube.com/watch?feature=player_detailpage&v=j_Y_0eh-FCQ

Summer video 3:

https://www.youtube.com/watch?feature=player_detailpage&v=fbckNjLK7mw

Summer video 4:

https://www.youtube.com/watch?feature=player_detailpage&v=KzVvo0u__-o

2. Reading and Vocabulary: You will use a free online Statistical tutoring site that will give you information on variable and data displays. While reviewing the information on the site you will be completing a vocabulary list (See pages 3 – 6). *Follow the steps below:*

- Go to www.stattrek.com
- Click on “AP Statistics” then “AP Tutorial”
- On the left side of the screen is a list of general topics. Under each general topic are a list of subtopics. You will read the following subtopics to complete the vocabulary list.

General Topic: Exploring Data	
Subtopics:	Variables
	Population Vs. Sample
	Central Tendency
	Variability
	Position
General Topic: Charts and Graphs	
Subtopics:	Charts and Graphs
	Patterns in data
	Dotplots
	Histograms
	Stemplots
	Boxplots
	Scatterplots
	Comparing Data Sets

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

A TI-84 Plus graphing calculator is a required tool for this course, and you **MUST** bring it to class every day. As you complete the practice problems refer to the TI Guidebook to become familiar with the list and statistical functions. For an online calculator go to www.alcula.com/calculators/statistics

This packet should be completed by your return to school in August. You are expected to complete each part of each problem and to construct all data displays neatly.

Statistics AP Summer Assignment

Part 2: Vocabulary List

Please define each of the following terms from the information on the www.stat Trek.com website. When asked, provide a UNIQUE example or sketch of the word... One NOT given on the website and NOT the one your friends use.

1. Categorical Variables

Example:

2. Quantitative Variables

Example:

3. Discrete Variables

4. Continuous Variables

5. Univariate Data

6. Bivariate Data

7. Population

Example:

8. Sample

Example:

9. Median

10. Mean

Formula:

11. Outlier

12. Parameter

13. Statistics

14. Range

15. Standard Score (z-score)

Formula:

16. Center

17. Spread

18. Variance:

Formula:

19. Standard Deviation

Formula:

20. Symmetric

Sketch:

21. Unimodal

Sketch:

22. Bimodal

Sketch:

23. Skewness

Sketch Skewed left:

Sketch Skewed right:

24. Uniform

Sketch:

25. Gaps

Sketch:

26. Outliers

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 be used for comparing data

PART 3 – Categorical or Quantitative?

Determine if the variables listed below are quantitative or categorical. Place an “X” in the appropriate column.

Variable	Categorical	Quantitative
Time it takes to get to school		
Number of minors living in a household		
Hair color		
Temperature of a cup of coffee		
Salary		
Gender		
Smoking status		
Height		
Amount of milk spilled		
Age of Oscar winners		
Type of shoes worn		
Jelly bean flavor		
Type of cheese		
Number of shoes owned		

Part 4: Making sense of statistical data with graphs and charts

1. Statistic – What is that?

A statistic is a number calculated from data that gives us insight about trends, preferences, and relationships among the data. Quantitative data has many different statistics that can be calculated. Determine the given statistics from the data below on the number of home runs Mark McGuire hit in each season from 1982 to 2001.

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

- a. Mean:
- b. Minimum:
- c. Maximum:
- d. Median
- e. Q_1 :
- f. Q_3 :
- g. Range:
- h. IQR:

2. 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.

- a. Find the percent of accidental deaths from each of these causes, rounded to the nearest percent.
- b. What percent of accidental deaths were from "other" causes?
- c. A pie chart is a graphical display used to show all the categories in a categorical variable relative to each other. Create a well-labeled pie chart for the accidental death percentages. Be sure to include a category for “other causes.” You may try using a software or internet source to make one and paste it below. Try Microsoft Excel. It is an excellent graph-making software that we will be using in class.

3. 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																

Make a dotplot to display these data. Make sure you include appropriate labels, title and scale. The graph paper should help ensure you space your markings (you may use x's or dots consistently).



4. A marketing consultant observed 50 shoppers at a supermarket. One variable of interest was how much each shopper spent at 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

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



5. The table below shows the percentage of residents age 65 or older in each of the 50 states.

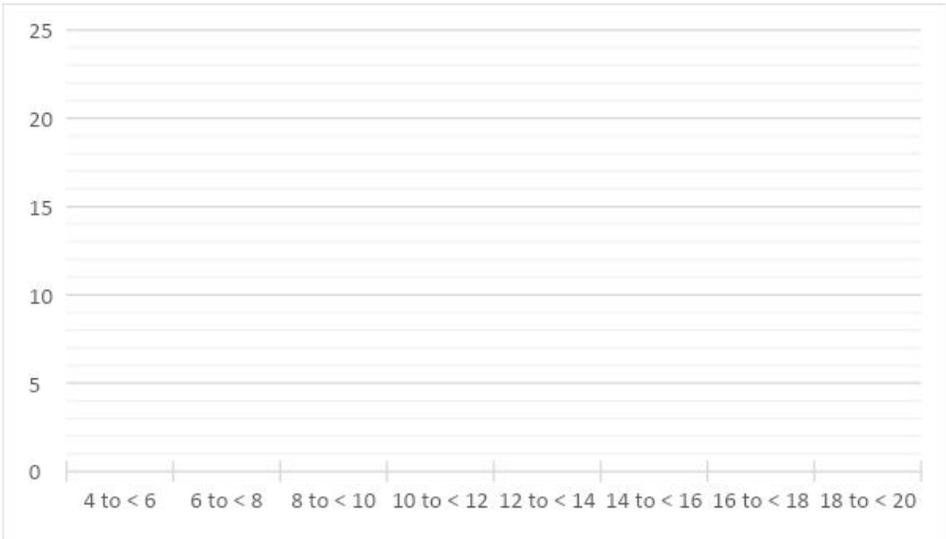
State	Percent	State	Percent	State	Percent
Alabama	13.1	Louisiana	11.5	Ohio	13.4
Alaska	5.5	Maine	14.1	Oklahoma	13.4
Arizona	13.2	Maryland	11.5	Oregon	13.2
Arkansas	14.3	Massachusetts	14.0	Pennsylvania	15.9
California	11.1	Michigan	12.5	Rhode Island	15.6
Colorado	10.1	Minnesota	12.3	South Carolina	12.2
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 (also called “bins”) which are represented by bars. The height of each bar represents the frequency, or number of data values in each class. The classes have the same width and the bars must touch each other because the number line is continuous. To make a histogram, we must first decide on an appropriate bin width and the number of classes. Usually, 5 to 10 is a good number of classes, and the bin width will be determined by the number of classes you choose to efficiently display the data. We then count the number of data values in each class, which will determine the height of each bar. The classes for the percentage of residents 65 or older have been started for you.

a. First, complete the table with the class widths and the frequencies of data values in each class.

Class Widths	Frequency
4 to 5.9	1
6 to 7.9	0
8 to 9.9	2

b. Construct a histogram on the grid below. Make sure to include a chart title and axis labels in the context of the problem. The scale has already been added.



6. SSHA Scores

Below are the scores on the Survey of Study Habits and Attitudes for 18 college freshman women:

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

and for 20 college freshman men:

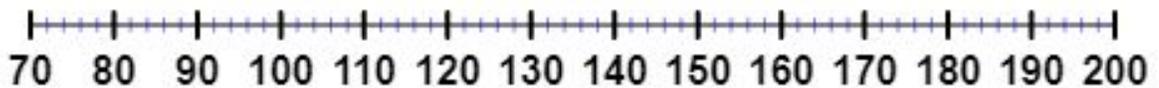
108 140 114 91 180 115 126 92 169 146 109 132 75 88 113 151 70 115 187 104

- a. Put the data values in ascending order for each gender. Then compute the numerical summaries for each gender.

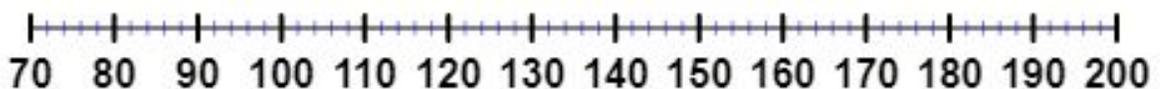
	Women	Men
Mean		
Minimum		
Q_1		
Median		
Q_3		
Maximum		
Range		
IQR		

- b. Using the Minimum, Q_1 , Median, Q_3 , and Maximum from each gender, construct parallel boxplots to compare the distributions. Note that the scales are exactly the same to get a better comparison of the distributions

Men:



Women:



To answer the following, refer to the readings on www.stattrek.com "Survey Sampling Methods."

7. The 7 types of sampling designs are:

- A. Voluntary response B. Convenience C. Simple random D. Stratified E. Cluster
F. Multistage G. Systematic

- a. The Maryland division of Weight Watchers is doing research to determine how many people on the Weight Watchers diet cheat at least once a week. They decide that anonymous surveys will give them an accurate representation but do not have time to get responses from **all** the Maryland Weight Watchers people.

Read the scenarios below and determine which of the 7 sampling methods best describes it

- _____ I. Randomly select 10 members from each of the WW centers in the Maryland division.
- _____ II. Use an alphabetical listing of all Maryland division members. Randomly choose a starting person on the list. Then select every 20th person thereafter.
- _____ III. Randomly select 2 or 3 branches of the Maryland division and survey every member of that center.
- _____ IV. Send out the survey to every member of the Maryland division. Place drop boxes in each WW center. Anyone who returns the survey will be in the sample.
- _____ V. The Maryland regional office is in Baltimore, so they survey members at the WW center in Baltimore.
- _____ VI. From a numbered list of all Maryland division members, use a computer to randomly select 100 numbers and survey all members with those corresponding numbers.

- b. What is the population of interest in the WW situation?

Part 5-Algebra Skills

8. The prerequisite for AP Statistics is Algebra II. You will not find very much equation-solving in this course, but some quick review of Algebra I and Algebra II content will be helpful

Here is a formula that is used often in AP Statistics: $z = \frac{x - \bar{x}}{s}$. Use your algebra skills to find the following:

- a. If $z = 2.5$, $x = 102$, and $\bar{x} = 100$, what is s ? Show your work.

- b. If $z = -3.35$, $x = 60$, and $s = 4$, what is \bar{x} ? Show your work.

Part 6-Regression

It is expected that you have a thorough understanding of linear functions and scatter plots.

9. 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.
- If 1990 could be considered “year 0”, which of the above numbers represents the slope and which represents the y-intercept?
 - What is the equation of the line of best fit using the slope and y-intercept above?
 - Predict the average consumption of sweeteners per person for the year 2005.
 - 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).

- What does the slope represent in this problem? Interpret it in the context of this problem/situation.
- What does the y – intercept represent in this problem? Interpret it in context.
- 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 (in inches):

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

- Construct a scatter plot of the data. You must include labels.

- Describe the association between the heights of women and the men they date.

Part 7-Probability

You are expected to have a basic understanding of simple probability. If you have any problems with these calculations, there are numerous sites available online that provide basic probability explanations.

- f. A special lottery is to be held to select the student who will live in the only “deluxe” room in the 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. $\frac{1}{8}$ B. $\frac{2}{9}$ C. $\frac{2}{7}$ D. $\frac{3}{8}$ E. $\frac{1}{2}$
- g. Which of the following probabilities is probably closest to 0.5?
- The sun will rise tomorrow
 - It will rain tomorrow
 - You will see a 3-legged dog when you leave the room
 - A fair die is tossed and a 6 appears four times in a row
 - There will be a plane crash somewhere in the world in the next five minutes
- h. If a coin is tossed twice, what is the probability that on the first toss the coin lands on heads, and on the second toss the coin lands on tails?
- A. $\frac{1}{6}$ B. $\frac{1}{3}$ C. $\frac{1}{4}$ D. $\frac{1}{2}$ E. 1
- i. If a coin is tossed twice, what is the probability that it will land on either heads both times or tails both times?
- A. $\frac{1}{8}$ B. $\frac{1}{6}$ C. $\frac{1}{4}$ D. $\frac{1}{2}$ E. 1
- j. Calculate the following probabilities and arrange them in order from least to greatest:
- The probability that a fair die will produce an even number. _____
 - A random digit from 1 to 9 (inclusive) is chosen with all digits being equally likely. The probability that when the digit is squared, the answer will contain the digit 1. _____
 - The probability that a letter chosen from the alphabet will be a vowel. _____
 - A random number between 1 and 20 (inclusive) is chosen. The probability that its square root will not be an integer. _____

Order: _____, _____, _____, _____