

AP Statistics 2017-2018

Summer Assignment

Welcome to Advanced Placement Statistics. This summer assignment will provide you with an overview of some basic statistics concepts and methods of displaying data. By completing this assignment we will hit the ground running when we return to school in August. This assignment will be due Friday August 18th.

The website www.stattrek.com offers a free tutorial which will help you prepare for the course and complete the summer assignment. Once you arrive on the stattrek homepage, the left side of the page gives a list of topics. Near the bottom, under the Help category, you will see [Statistics Dictionary](#). This dictionary will help you with this summer assignment. The following topics will be useful in completing this assignment.

General Topic: Quantitative Measures	
Subtopics:	Variables
	Central tendency
General Topic: Charts and Graphs	
Subtopics:	Patterns in data
	Dotplots
	Histograms
	Stemplots
	Boxplots
	Cumulative frequency plots
	Comparing distributions
General Topic: Categorical data	
Subtopics:	One – way tables
General Topic: Survey Sampling	
Subtopics:	Survey Sampling Methods

If you need to contact me over the summer, my email is astagner@pasco.k12.fl.us. My phone number is 727-597-3256.

Vocabulary List

Please define each of the following terms from the information on the stattrek website. When asked, provide a unique example of the word not given on the website.

1. Categorical variables (define and include an example)
2. Quantitative variables (define and include an example)
3. Univariate data
4. Bivariate data
5. Median
6. Mean
7. Population (define and include an example)
8. Sample (define and include an example)
9. Measures of center (name 3)
10. Measures of spread (name 3)
11. Symmetry
12. Unimodal and bimodal
13. Skewness (define and draw sketches)
14. Uniform
15. Gaps
16. Outliers
17. Dotplots
18. Difference between bar graph and histogram
19. Stemplot
20. Boxplot
21. Quartiles
22. Range
23. Interquartile range
24. Parallel boxplots
25. Parameter
26. Statistic

CATEGORICAL OR QUANTITATIVE

Determine if the variables listed below are *quantitative (Q)* or *categorical (C)*.

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 medication
12. Jellybean flavors
13. Country of origin
14. Type of meat
15. Number of shoes owned

STATISTICS – 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 homeruns Mark McGuire has hit in each season from 1982 – 2001.

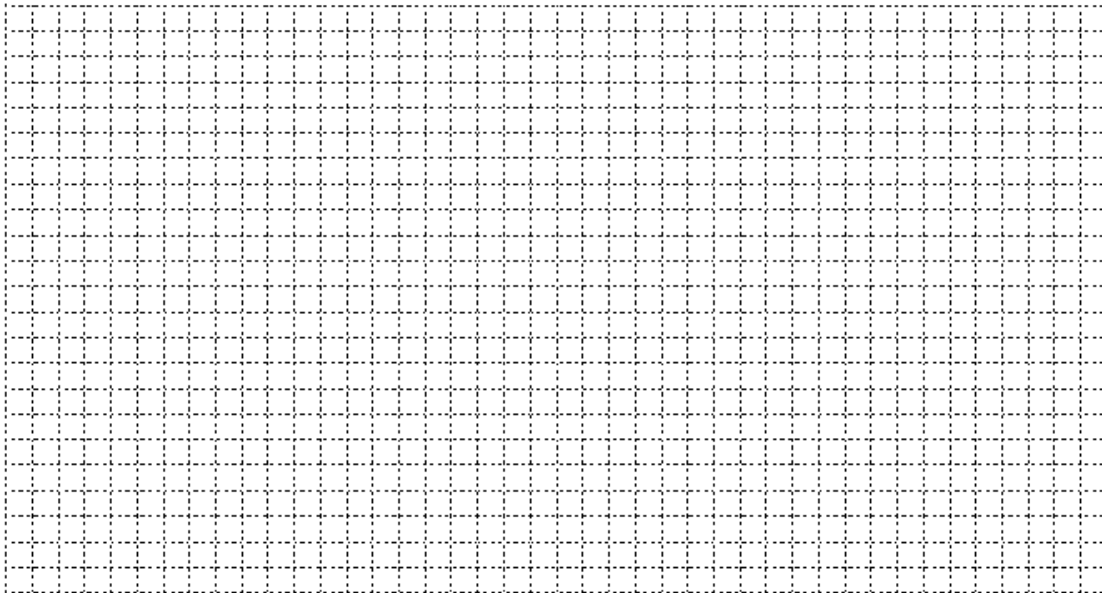
70	52	22	49	3	32	58	39
39	65	42	29	9	32	9	33

Mean	
Minimum	
Maximum	
Median	
Q1	
Q3	
Range	
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.

- 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. NEATLY create a well-labeled **bar graph** of the distribution of causes of accidental deaths. Be sure to include an “other causes” bar.



- d. A pie chart is another graphical display used to show all the categories in a categorical variable relative to each other. Create a pie chart for the accidental death percentages. You may try using a software or Internet source to make one and paste in the space below. (*Microsoft Excel works well*)

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 (round 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.



Key:

- b. Describe the shape, center, and spread of the distribution. Write a few sentences describing the amount of money spent by shoppers at this supermarket.

WHERE DO OLDER FOLKS LIVE?

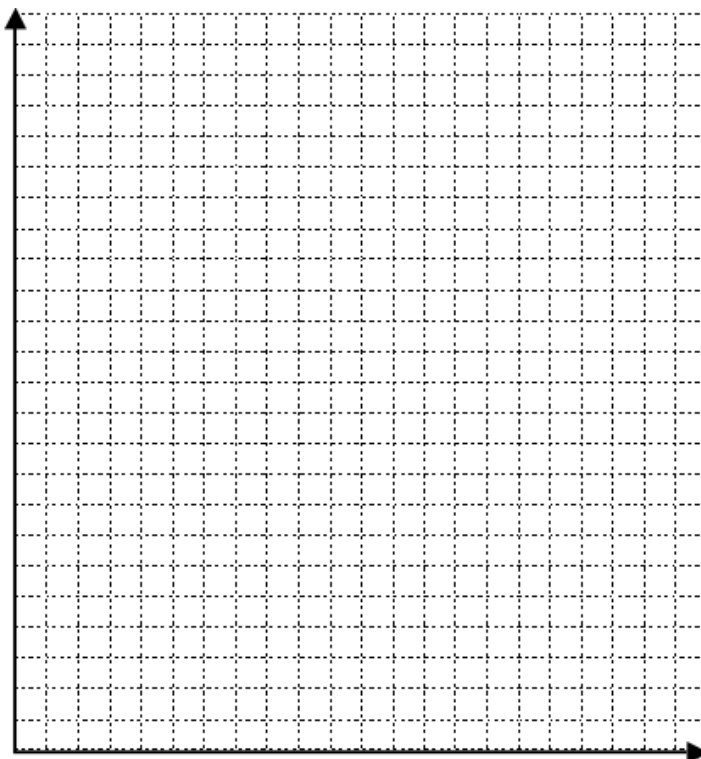
This table gives the percentage of residents aged 65 of older in each of the 50 states.

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

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

Bin Widths	Frequency
4 to < 6	1
6 to < 8	
8 to < 10	



- b. Describe the distribution of people aged 65 and over in the states. Discuss center, shape and spread.

For this first section, you should refer back to your reading on www.stat Trek.com "Survey Sampling Methods".

The 7 types of sampling designs are:

- A. voluntary response B. convenience C. simple random D. stratified
E. cluster F. multistage G. systematic

1. 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 branches 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 branch.
_____ IV. Send out the survey to every member of the Maryland division. Place drop boxes in each WW center. Anyone who returns a 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.

2. What is the population of interest in the WW situation?

Here is a formula that is used often in AP Statistics: $z = \frac{x - \bar{x}}{s}$.

1. If $z = 2.5$, $x = 98$ and $\bar{x} = 100$, what is s ? Show your work.

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

You are expected to have a basic understanding of simple probability.

1. A special lottery is to be held to select the student who will live in the only deluxe room in a 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. $\frac{1}{2}$
2. Which of the following has a probability closest to 0.5?
 - A. The sun will rise tomorrow.
 - B. It will rain tomorrow.
 - C. You will see a dog with only three legs when you leave the room.
 - D. A fair die will come up with a score of 6 four times in a row.
 - E. There will be a plane crash somewhere in the world within the next five minutes.
3. If a coin is tossed twice, what is the probability that on the first toss the coin lands heads and on the second toss the coin lands tails?
 - A. $\frac{1}{6}$
 - B. $\frac{1}{3}$
 - C. $\frac{1}{4}$
 - D. $\frac{1}{2}$
 - E. 1
4. If a coin is tossed twice what is the probability that it will land 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
5. Calculate the following probabilities and arrange them in order from least to greatest.
 - I. The probability that a fair die will produce an even number. _____
 - II. A random digit from 1 to 9 (inclusive) is chosen, with all digits being equally likely. The probability that when it's squared it will end with the digit 1. _____
 - III. The probability that a letter chosen from the alphabet will be a vowel (a,e,i,o,u). _____
 - IV. A random number between 1 and 20 (inclusive) is chosen. The probability that its square root will not be an integer. _____

ORDER: _____, _____, _____, _____