

# Statistics Summer Assignment 2017

Welcome to Statistics future statisticians! The purpose of this assignment is to make you comfortable exploring data analysis and to review statistical concepts. **This assignment is for both Regular Stats and AP Stats classes.** I suggest you print this document out now.

The summer assignment is composed of three 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 the information on the site you will be completing a vocabulary list. To access the site, follow these steps

- Go to [www.stattrek.com](http://www.stattrek.com)
- Click on "AP Statistics", then click on "AP Statistics Tutorial".
- On the left side of the screen is a list of general topics. Under each general topic are lists of subtopics. Please read the following subtopics to complete the vocabulary list.

<b>General Topic:</b>	<b>The basics</b>
<i>Subtopics:</i>	Variables
	Population vs. Sample
	Central Tendency
	Variability
	Position
<b>General Topic:</b>	<b>Charts and Graphs</b>
<i>Subtopics:</i>	Patterns in Data
	Dotplots
	Histograms
	Stemplots
	Boxplots
	Scatterplots
	Comparing Data Sets

## 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 should do so in the spaces provided.

## 3. Graphing Calculator Use

A graphing calculator is a required tool for this class. The TI-83+ or (TI-84 is recommended). 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/](http://www.alcula.com/calculators/statistics/). This has central tendency and dispersion, box plot, scatter-plot, etc., **Graphs for this assignment are to be done by hand with a pencil.**

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 and precisely. This assignment will be graded, and it will count as a quiz grade in the first nine weeks grading period.

I will be checking my g-mail during the summer if you have questions. My g-mail is [jroth@paduafranciscan.net](mailto:jroth@paduafranciscan.net)

### Part One: Vocabulary List

Define each of the following terms from the information on the Stattrek website. Blue words on the website can be clicked on for more information. If a term requires an example or sketch, provide a unique one, not one copied from the website.

#### 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. Symmetry

*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. Differences between Bar Charts and Histograms

31. Stem Plot

32. Box Plot

33. Quartiles

34. Range

35. Interquartile Range (IQR)

36. Four ways to describe data sets (SOCS)

37. Types of Graphs that can be used for comparing Data

### Part Two: Practice Problems

#### **CATEGORICAL OR QUANTITATIVE?**

Determine if the variables listed below are quantitative or categorical.

1. Time it takes to get to school
2. Number of people under 18 years old 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

### 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 homeruns Mark McGuire hit in each season from 1982-2001. The online calculator works for this one. You can also enter the data in you calculator’s list L1, then Stats button, 1-Var Stats.

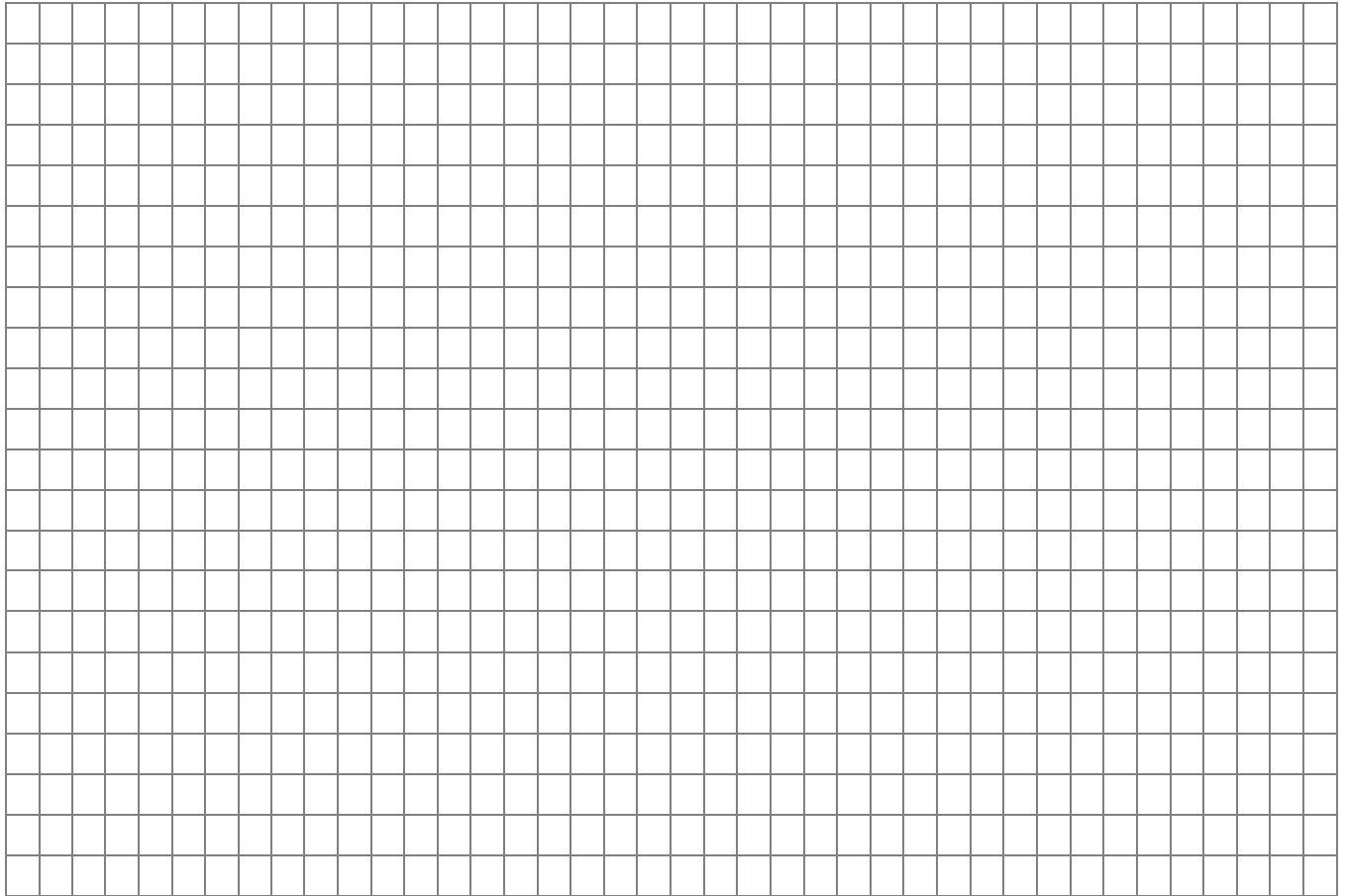
70	52	22	49	3	32	58	39
39	65	42	29	9	32	9	33

Mean	
Minimum	
Maximum	
Median	
Q1	
Q3	
Range	
IQR	Q3-Q1

## ACCIDENTAL DEATHS

In 1997 there were 92353 deaths from accidents in the United States. Among these were 42340 deaths from motor vehicle accidents, 11858 from falls, 10163 from poisoning, 4051 from drowning, and 3601 from fires. The rest were “other” causes.

- Find the percent of accidental deaths from each of these causes, rounded to the nearest whole percent.
- What percent of accidental deaths were from “other” causes?
- Create a neat and well-labeled bar graph of the distributions of causes of accidental deaths below. Be sure to include an “other causes” bar.



- 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. Microsoft Excel works well.

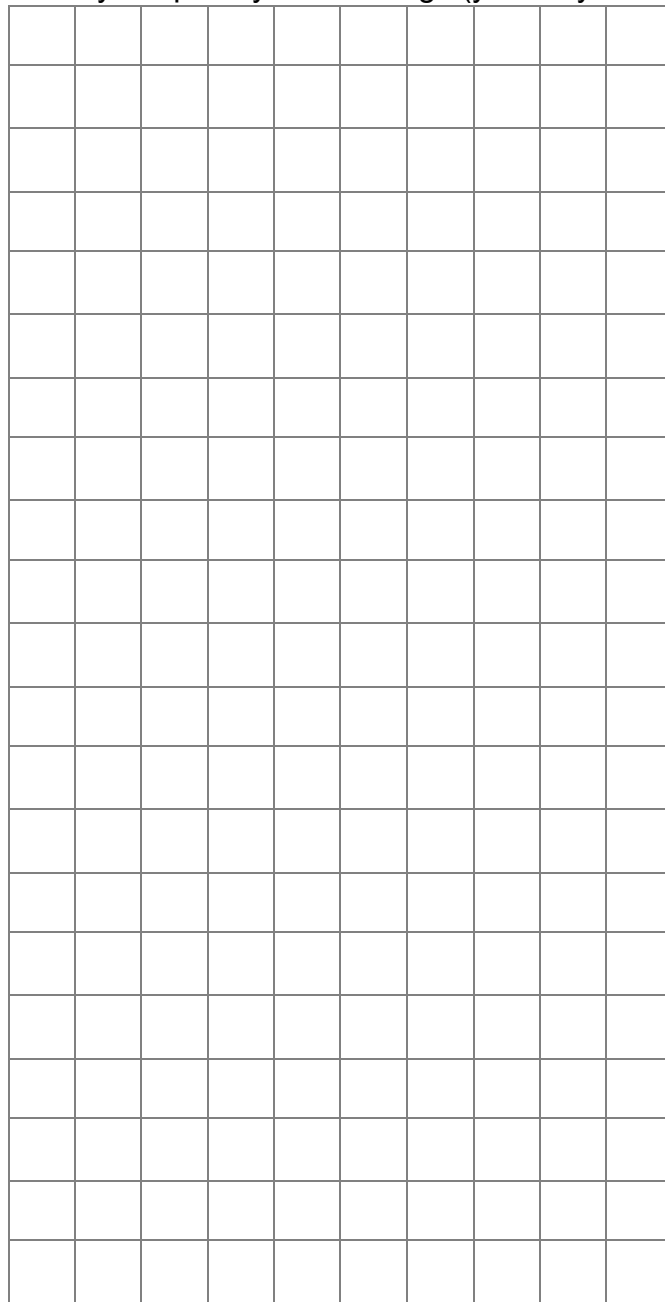


## IT'S A TWISTA

The data below gives the number of hurricanes that happened each year from 1944 through 200 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 dot plot 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.



## 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

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



## WHERE DO OLDER FOLKS LIVE?

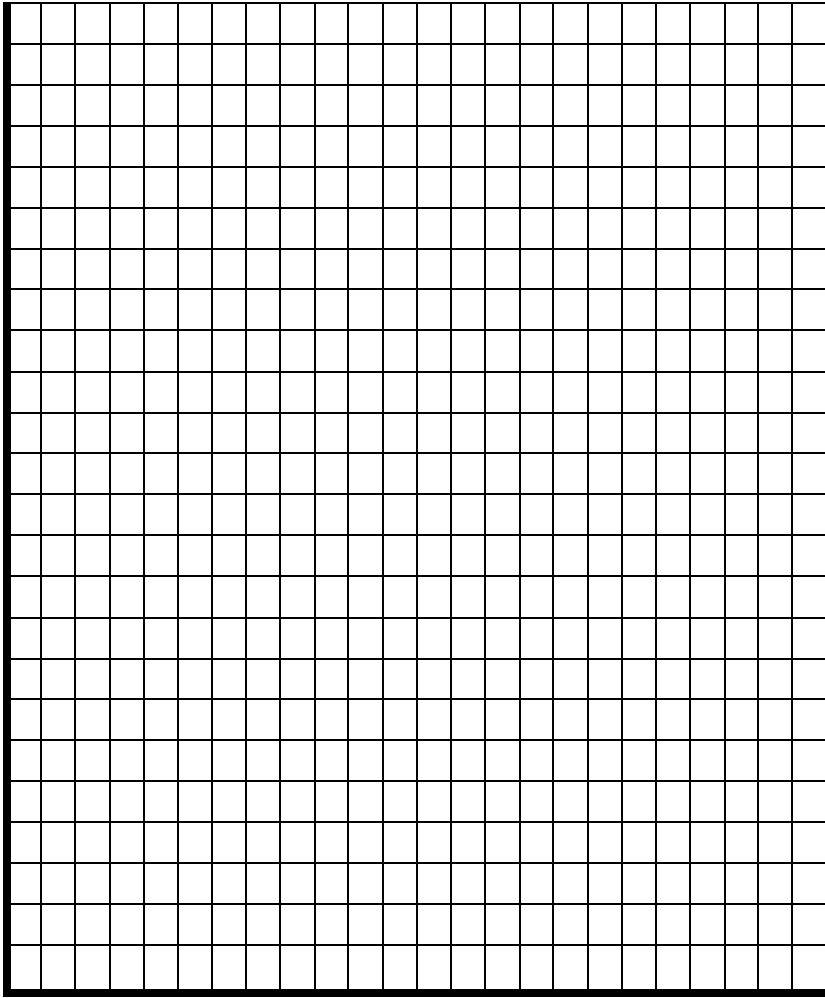
This table gives the percentage of residents aged 65 or older in each of the 50 states.

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 bins (the bars). These bins have the same width and scale and are touching each other 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 the 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	



## 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

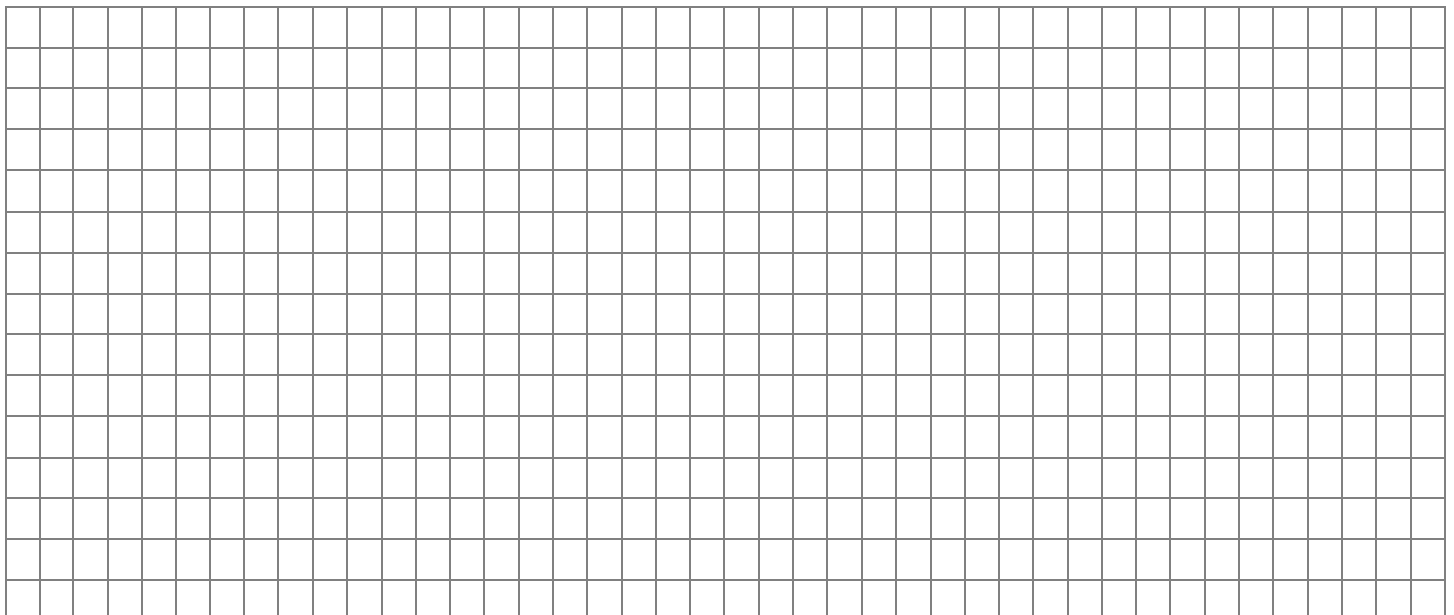
and for 20 first year men:

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 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	
Outliers		Outliers	

b. Using the Minimum, Q1, Median, Q3, Maximum, and any Outliers from each gender, make parallel box plots (two) to compare the distributions. They must share the same scale.



To answer the following, refer to the readings on [www.stat Trek.com](http://www.stat Trek.com) Click on “AP Statistics”, then on “Surveys”, then on “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

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 for 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 20<sup>th</sup> 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 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 members and survey all members with those corresponding numbers.

THANK YOU FOR TAKING TIME OUT OF YOUR SUMMER TO DO a little STATS! Looking forward to working with all of you in the near future. Have a great rest of Summer!