# Welcome to AP Statistics at RHS!!!

*Teacher Contact Information:*

Ms. Mullen

jmullen@wcpss.net

*(at any point this summer that you have questions, please email me from your student account or I cannot respond)*

*I am so glad you have decided to take on the challenge of taking AP Statistics! This course will require you to be fully committed in order to learn the material and be successful in the class. Please consider what type of commitment you are willing to make for the class and speak with a math teacher that you trust to make sure the class is a good fit for you. I plan for the class to interactive and fun but it still will take work! Below is a description of the class:*

A.P. Statistics is equivalent to a one semester, introductory, non-calculus based, college course in statistics.  This course will introduce participants to the major concepts and tools for collecting, analyzing, and drawing conclusions from data.  Participants are exposed to four broad conceptual themes:

1)Exploring Data: Observing patterns and departures from patterns.

2)Planning a Study: Deciding what and how to measure.

3) Anticipating Patterns in Advance: Producing models using probability and simulation.

4) Statistical Inference: Confirming models.

*This summer project will focus on giving you a head start into the Exploring Data portion on the course and completing this project will be imperative to starting the year off smoothly. This project is 15% of your quarter grade (Q1 if taken in the fall and Q3 if taken in the spring).*

*There are two parts to this project. Some of the sections will require separate sheets of paper and some can be written on the given paper. If it has a paper icon at the top you will need a separate sheet of paper. These separate sheets of paper need to be stapled to this packet.*

*#1- Chapter 1 Vocabulary*
*#2- Chapter 1 Assignments*

<u>*Timeline:*</u>

*Due date- Friday, September 2, 2016*
*You will also have your first test this day (covering Chapter 1 and 2).*

<u>***Additional Resources:***</u>

*http://www.apstatsguy.com/*

*http://stattrek.com/tutorials/ap-statistics-tutorial.aspx*

**Chapter 1 Vocabulary**

*For each of the following words, provide meaning in context of statistics (use the internet and/or notes from previous classes. This is to be turned in on a separate sheet of paper or as flashcards upon returning to school)*

Individuals
Variable
Categorical Variable
Quantitative Variable
Binary Variable
Discrete
Continuous
Distribution
Marginal Distribution
Conditional Distribution
Association
Frequency
Relative Frequency
Dotplot
Bar graph
Shape
Center
Spread
Pattern
Mode
Symmetry
Skewed Right
Skewed Left
Unimodal
Bimodal
Multimodal
Stemplot
Back-to-back stemplot
Splitting Stems
Histogram
Outliers
Mean
Median
Range
Interquartile Range
First quartile
Third quartile
1.5xIQR Rule
Five-number summary
Standard Deviation
Variance
Resistant
Boxplot
Variability

QUANTITATIVE VS. QUANTITATIVE

*Exercises: Answer the following questions. Then decide if the data is quantitative or categorical. Then go back and decided whether the quantitative variables are discrete or continuous and the categorical variables are binary or other.*

1.  In what grade did you take Algebra I?


2.  How many pairs of shoes do you own?


3.  How old was your father when you were born?


4.  Choose a random integer between 1 and 20.


5.  How many siblings do you have? (all, whether they live with you or not)


6.  How fast can you run the mile?


7.  How tall are you (inches)?


8.  How many AP classes will you be taking THIS year?


9.  What gender are you?


10. Where did you eat your last meal? (1 = home, 2 = restaurant, 3 = other)


11.  How many years have you lived in this area (Wake County)?


12. How many miles from school do you live?


13. In what year were you born?


14. What was the average temperature the last two days?


15. Can you curl your tongue?

1). {10 15 23 24 18 32 35 19 14 15}
a). Find the range of the data.
b). Find the lower quartile, median, and upper quartile.
c). Find the interquartile range.
d). Does the data contain any outliers? If so, what are they?

2). {352 333 314 301 328 349 336 311}
a). Find the range of the data.
b). Find the lower quartile, median, and upper quartile.
c). Find the interquartile range.
d). Does the data contain any outliers? If so, what are they?

3). {57 45 49 32 82 44 41}
a). Find the range of the data.
b). Find the lower quartile, median, and upper quartile.
c). Find the interquartile range.
d). Does the data contain any outliers? If so, what are they?

*For the data set below, create a dotplot and explain how it was constructed. How do you know that it was constructed properly?*

**Student GPAs**

| 2.4 | 2.6 | 2.4 | 2.8 | 2.9 | 3.2 | 3.4 | 3.4 |
|-----|-----|-----|-----|-----|-----|-----|-----|
| 2.6 | 4 | 2.1 | 2.4 | 2.4 | 3.6 | 3.6 | 3.7 |
| 4.0 | 2.2 | 2.7 | 2.8 | 3.4 | 3.9 | 3.9 | 3.8 |

*For the data below, create a boxplot and explain how it was constructed. How do you know that it was constructed properly*

**Height of Students (in)**

| 44 | 72 | 68 | 64 | 60 | 63 | 73 | 71 | 62 | 63 | 61 |
|----|----|----|----|----|----|----|----|----|----|----|

**\*\*In AP Statistics, a modified boxplot is used. This shows any outliers. The outliers are shown with dots, stars or any other small symbol.**

*For the data above, create a modified boxplot. Be sure to show your calculations for the outliers.*

**Student GPAs**

| 3.1 | 2.1 | 4.0 | 2.3 | 3.3 | 2.7 | 2.8 | 3.5 | 3.1 |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 3.4 | 4.0 | 1.7 | 1.5 | 1.9 | 2.1 | 2.4 | 3.7 | 4.0 |
| 2.6 | 3.2 | 3.0 | 0.7 | 1.0 | 1.1 | 3.5 | 3.6 | 2.3 |

*For the data below, create a frequency table and a histogram and explain how it was constructed. How do you know that it was constructed properly?*

**Points Scored by Winning Team in the Super Bowl**

| 25 | 19 | 9 | 16 | 3 | 21 | 7 | 17 | 4 | 12 |
|----|----|----|----|----|----|----|----|----|----|
| 17 | 5 | 10 | 29 | 22 | 36 | 19 | 32 | 4 | 45 |
| 1 | 13 | 35 | 17 | 23 | 10 | 14 | 7 | 15 | 7 |
| 27 | 3 | 27 | 3 | 11 | 12 | | | | |

| Data Set 1 | 5 | 10 | 4 | 3 | 11 | 18 | 19 | 21 | 23 | 27 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|

| Data Set 2 | 5 | 10 | 4 | 3 | 11 | 18 | 19 | 21 | 23 | 27 | 32 | 30 | 29 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|

| Data Set 3 | 5 | 10 | 4 | 3 | 11 | 18 | 19 | 21 | 23 | 27 | 32 | 30 | 29 | 42 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|

## FIVE-NUMBER SUMMARY

| Data Set 1 | Data Set 2 | Data Set 3 |
|---|---|---|
| Min: | Min: | Min: |
| Q1: | Q1: | Q1: |
| Median: | Median: | Median: |
| Q3: | Q3: | Q3: |
| Max: | Max: | Max: |

## SPREAD

| Data Set 1 | Data Set 2 | Data Set 3 |
|---|---|---|
| Range: | Range: | Range: |
| IQR: | IQR: | IQR: |
| Std Dev: | Std Dev: | Std Dev: |

Change the maximum value in each data set to 1000. Go back and recalculate each of the measurements.

## FIVE-NUMBER SUMMARY

| Data Set 1 | Data Set 2 | Data Set 3 |
|---|---|---|
| Min: | Min: | Min: |
| Q1: | Q1: | Q1: |
| Median: | Median: | Median: |
| Q3: | Q3: | Q3: |
| Max: | Max: | Max: |

## SPREAD

| Data Set 1 | Data Set 2 | Data Set 3 |
|---|---|---|
| Range: | Range: | Range: |
| IQR: | IQR: | IQR: |
| Std Dev: | Std Dev: | Std Dev: |

**Mothers' Ages at Time of 1st Child's Birth**

| 50 | 26 | 23 | 30 | 28 | 33 | 23 | 32 | 38 |
|----|----|----|----|----|----|----|----|----|
| 23 | 35 | 24 | 24 | 33 | 24 | 23 | 24 | 32 |
| 23 | 30 | 24 | 29 | 34 | 35 | 26 | 31 | 28 |

*Exercise:* Create the following graphs using the data from "Mothers' Ages at Time of 1st Child's Birth". Make sure to include labels for your number line and a title.

## Dot plot

## Modified Boxplot

## Stem plot

## Split Stem plot

**DESCRIBING DISTRIBUTIONS**

*When describing distributions in Statistics, several things must be addressed. Below discuss your options when addressing shape, center, spread, and outliers.*

*Shape:*




*Center:*




*Spread:*




*Outliers:*




Why is context important when describing a distribution?

Compare the following two distributions by construction a side-to-side modified boxplot and a back-to-back stem plot. Comment on that data's distribution. Make sure you address shape, center, spread, clusters or gaps, and any outliers in *context* of the data.

**Number of Babe Ruth Home Runs 1920—1934**

| 54 | 59 | 35 | 41 | 46 | 25 | 47 | 60 | 54 | 46 |
|----|----|----|----|----|----|----|----|----|----|
| 49 | 46 | 41 | 34 | 22 |    |    |    |    |    |

**Number of Roger Maris Hone Runs, 10 years in AL**

| 13 | 23 | 26 | 16 | 33 | 61 | 28 | 39 | 14 | 8 |
|----|----|----|----|----|----|----|----|----|----|

**Comments:**

**Graphs**