# Lab Exercise 3: Basic Statistics, Visualization, and Hypothesis Tests Part 1 - Basic Statistics and Visualization Using R

## Step 2

### Part 2

The mean is 67200.

### Part 3

The median is 50300.

### Part 4

The standard deviation is 68178.

### Step 3

The mean is 5.63.  
The median is 6.00.  
The standard deviation is 1.99.

## Step 4

### Part 1

There is a household income of 4 to 1620560 which makes sense as there will always be people on either extreme.

### Part 3

After trimming 10% off each end of the data, the new mean is 55347.4970339335. This is lower than the mean previously found.

### Part 5

The mean for income between 10000 and 1000000 is 72018. This is higher than previous means, meaning we trimmed more lower data.

### Part 6

This makes similar sense to Part 3. We trim out the extreme values and only looking at a subsection of data.

### Part 7

It depends what the question is being asked. In general, I believe trimming a percent of data from both sides makes the most sense to me.

# Part 2 - Graphics Package Plots and Hypothesis Tests

## Step 5

### Part 2

The p-value on the F-stat is 6.336e-13. We can reject the null hypothesis as the p-value is small, and it is the probability that the “qualified” sample could be produced randomly.

### Part 3

The p-values on the following coefficients are

log10(purchase\_amt | no offer) aka m0 = < 2e-16  
mean(log10(purchase\_amt |offer1)) – m0 = 2.68e-09  
mean(log10(purchase\_amt |offer2)) – m0 = 4.02e-12

### Part 4

Yes, we can reject the null hypothesis that the mean purchase amount for offer 1 was different from no offer because of the highly significant p-value. Similarly, we can reject the null hypothesis that the mean for offer 2 was different from no offer.

## Step 6

### Part 1

Because the p-value between offers was 0.4602617, we cannot tell whether offer 1 was different than offer 2.

### Part 2

Having an offer definitely affects purchase amount, but there is no significant effect between either offers. My recommendation is to continue the offer that affects the bottom line the least (the cheapest option) if it is cost effective.

## Step 16

### Part 3

The means are different because p < 0.05 at 0.001297255.

## Step 17

### Part 2

t.test() gives the same result of p-value = 0.001297.