You are a researcher at a marketing company responsible for doing analysis. You will have to interpret your results in layman’s terms for your boss. Your company has already hired an external agency to go and collect data. Now, it’s your job to analyze this dataset and find insights.

Extra credit: 20 (in Excel) or 30 (in R)

The dataset contains the following variables:

• Satisfaction: Customers indicate their satisfaction from the scale of 0-100

• Age: Customer’s age

• Gender: 1 = female, 0 = male

• Previous purchase: number of products purchased previously

• Social media: whether the customer has a social media account or not.

Questions:

1. Get the mode for the variable gender, social media. Interpret what these numbers mean

**Mode of Gender:** is 1. Interpreting what this number means is that the majority of the people who took this survey were female.

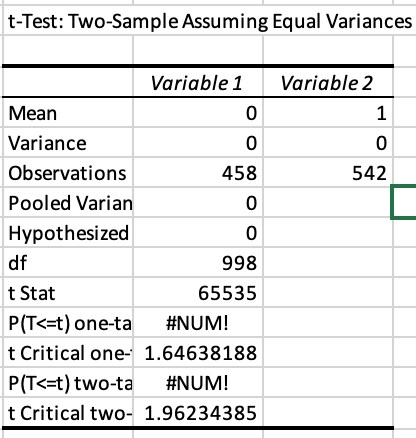
**Mode of Social Media:** Is 0. This means that a majority of the people who use social media are males. Or it means that a majority does not have social media. The instructions are not too clear what 1 aand 0 mean in terms of the social media data set.

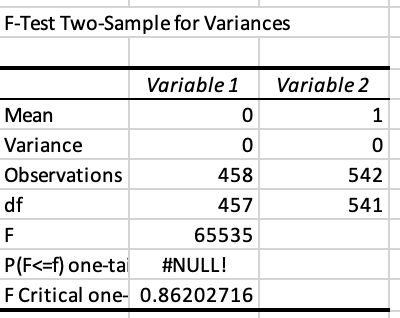
2. Get the mean for the variable satisfaction, age. Interpret what these numbers mean

Mean of Satisfaction: 53.87. This means that the ooverall satuisfaciton from this survey for this company

Mean of Age: 29.87. This means that the overall average age of the survey takers and customers were about 29-30 years old.

3. Is there a difference between male and females regarding their levels of satisfaction?

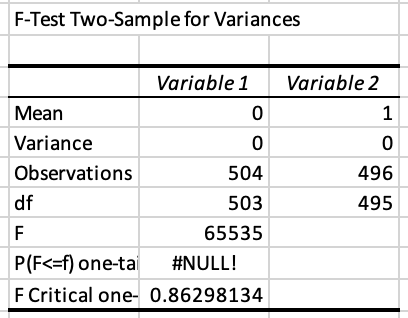
Conduct F-test, t-test, and interpret 

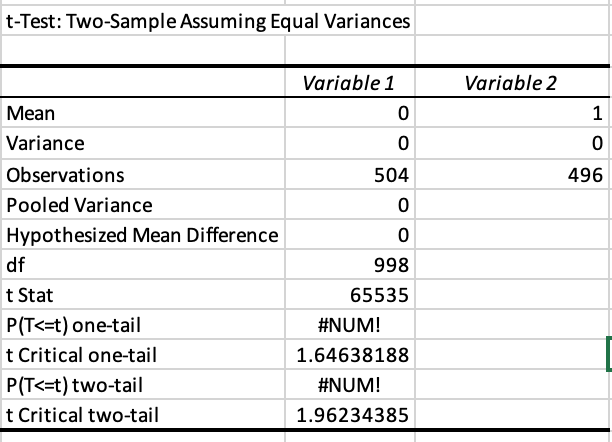
F-test: 

T-Test: Interpreting the T-Test, I see that there were 458 males taking the survey and 542 females took the survey. The variance was 0 for both variables. We reject the null.

The F-Test: The F-Test told me that the F was 65535 and the F critical was 0.862

4. Is there a difference between those who have social media accounts and those who don’t regarding their levels of satisfaction? Conduct F-test, t-test, and interpret

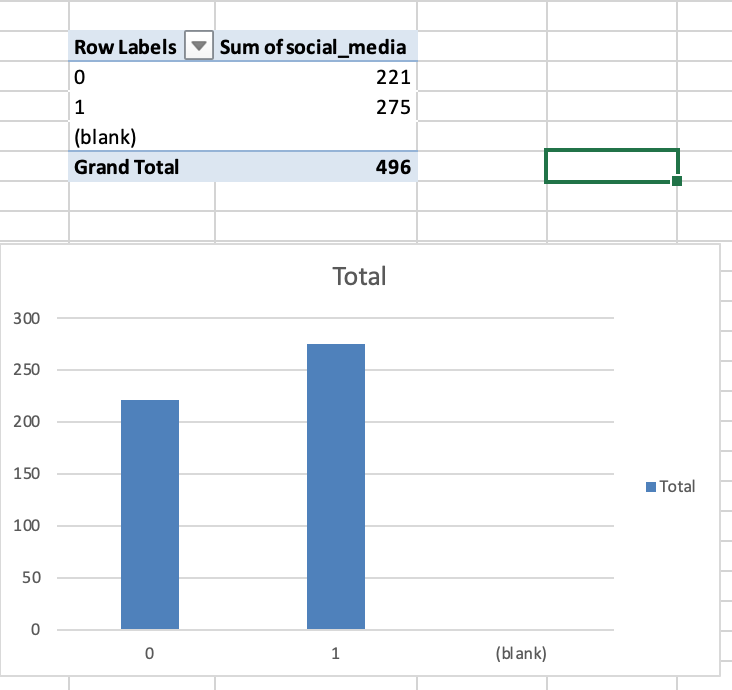
F Test: Again, the instructions were confusing. I am not sure if 0 is for male and 1 is for females for social media. This would not make sense because it would not answer if they have social media or not. I will take this as 0 = No and 1=yes. This means that I observed a majority of the survey takers do not have social media. This also tells me that a majority of the survey takers hhad a smaller satisfaction. The F critical was 0.863.

T-Test: 

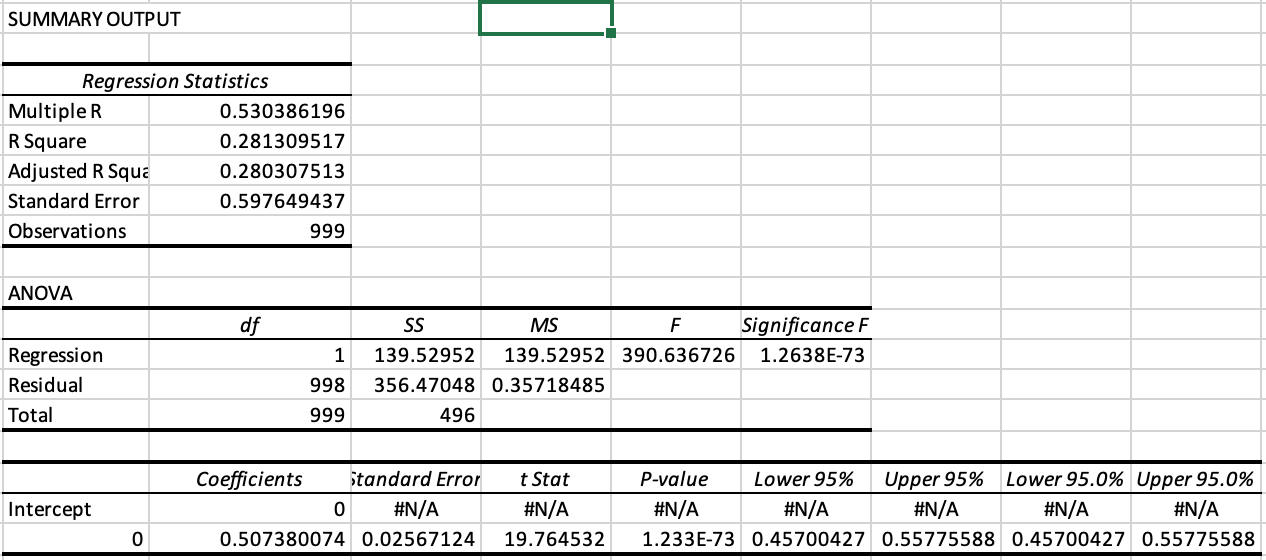
The T-Test was the same for the last T-Test. This shows that there is a difference between the people who have social media and ones who do not. The difference is that there are more

There is a difference between people who have sociaal media accounts and those who do not. Those who have social media accounts have a higher level of satisfaction. We reject the null.

5. Create a cross-tabulation table with the variables social media and gender (each cell contains the number of customers belonging to that category).



6. Conduct regression analysis and report and interpret the regression equation



As we can see, the

Correlation coefficient is 0.53. This tells us that the linear relationship is not too strong.

The coefficient of determination is 0.28. This tells us that the point barely falls on the regression line.

Standard error of regression: 0.598.

This is a photo showing that I completed this bonus problem before May 6th.

