

SAS Project 2

Due Dec 4th, 2018

1. The ultimate goal of your project is to see if two or more variables are correlated and give a linear equation which reflects the relationship.
2. You should use the data you collected in Project one. (If you do not like the data you obtained for Project 1, please take another survey and collect new data.) If your data does not correlate, please contact me as soon as possible. (No later than November 20th, 2018).
3. Do an analysis on the data –statistical summary, correlation and regression.
4. Do a brief write up of what you found.
 - a. First Page is a Title Page with name, and STAT 430
 - b. Second Page is a brief introduction to your problem. It should include an explanation of your variables, how you collected them, and the measures of center you found in the Project 1. It does not need to be long, but should be enough background so that anyone reading it has a general idea of what you have done.
 - c. Subsequent page(s) should be your statistical summary
 - i. Do a write up on the correlation and regression of your variables of interest. It should include:
 1. The Correlation Coefficient, state it and explain what it means
 2. Anova results from the PROC REG and what they mean (p-value and significance)
 3. Regression results and what they mean (p-values on variables, significance, residuals)

You must give the following:

 - The regression equation
 - The p-value for the independent variable(s)
 - Explain what the p-values mean
 - The residual analysis, look at the normal probability plot and analyze and the residuals by regression for dependent variable
 - ii. Make sure all assumptions are met. Explain what this does to your results.
 - iii. Explain why your sample is good or not good.
5. Attach all supporting documentation:
 - a. SAS code fully commented
 - b. SAS output including ALL graphs (This can either be part of your write up or at the end)
 - c. Survey if you are not using the same one as project 1.