SAS Project 2 Spring 2020

1. The ultimate goal of your project is to see if two or more variables are correlated and give a linear equations which reflects the relationship.
2. You should use the data you collected in Project one. Your data may not correlate, but you will know for sure after you do a Proc Reg. (Remember, usually we do the scatterplot and PROC CORR to make sure there is a linear relationship. We are skipping this crucial part of the procedure. It will show some of you the importance of a good data analysis.)
3. Do a simple linear regression using only two of your quantitative variables. (points written in blue)
4. (Optional, points written in red )Do a multiple linear regression using all your variables (including the qualitative – making dummy variables).
5. Do a brief write up of what you found.
   1. First Page is a Title Page with name, and STAT 430
   2. Subsequent page(s) should be your summary
      1. Do a write up on the regression of your two quantitative variables of interest.
         1. Anova results from the PROC REG and what it means (p-value and significance) (P-value stated, 5 points 10points. Explanation 5 points 10points)
         2. The r-squared value and what it means. (5 10points)
         3. Regression results and what it means (p-values on variables, significance, residuals)

(You must give the following:

* The regression equation (5 10points)
* The p-value for the independent variable(s) (5 10points)
* Explain what the p-values mean (5 10points)
* The residual analysis, look at the normal probability plot and analyze (5 10points) and the residuals by regression for dependent variable (510points)
  + - * + Make sure all assumptions are met. Explain what this does to your results. (5points 10points.)
    1. Do a write up on the regression of all your variables.
       1. Anova results from the PROC REG and what it means (p-value and significance) (P-value stated, 5 points. Explanation 5 points)
       2. The r-squared value and what it means. (5)
       3. Regression results and what it means (p-values on variables, significance, residuals)

(You must give the following:

* The regression equation (5)
* The p-value for the independent variable(s) (5)
* Explain what the p-values mean (5)
* The residual analysis, look at the normal probability plot and analyze (5) and the residuals by regression for dependent variable (5)
* Make sure all assumptions are met. Explain what this does to your results.(5)

1. Attach all supporting documentation: (10 10points)
   1. SAS code
   2. SAS output including ALL graphs
   3. Survey if applicable