Assignment #2: Single Variable Regression Model (0 points)

Data Directory: Data can be accessed on the SAS OnDemand server using this libname statement.

```
libname mydata '/courses/u_northwestern.edu1/i_833463/c_3505/SAS_Data/' access=readonly;
```

Data Set: mydata.building_prices

Data Description: See the data dictionary or pp. 328-329 of *Regression Analysis By Example*.

Assignment Instructions:

For this assignment we will fit a simple linear regression model (a regression model with a single predictor variable) to the building_prices data set. First, select the regression model of your choice based on the EDA that you performed in Assignment #1. Second, find the best simple linear regression model using the *selection=rsquare* option in PROC REG with *start=1* and *stop=1* (see Chapter 9 in *SAS Statistics By Example*). Did you select the optimal regression model using EDA?

Using the second regression model obtained from the *selection=rsquare* option perform an assessment of the model adequacy by producing diagnostic and residual plots using the *plots = (fitplot diagnostics residualplot)* in PROC REG (see Section 9.11 in *The Little SAS Book (5th Edition)* and Chapter 9 in *SAS Statistics By Example*). Note that a check of model adequacy includes: (1) plot the fitted regression model over the scatterplot, (2) an assessment of the normality of the residuals using a Quantile-Quantile plot (QQ plot), and (3) an assessment of the specification of the predictor variable by plotting the predictor variable against the residuals, and (4) a check for potential outliers using Cook's Distance. Comment on each of these aspects of model adequacy.

Assignment Document:

All assignment reports should conform to the standards and style of the report template provided to you. Results should be presented and discussed in an organized manner with the discussion in close proximity of the results. The report should not contain unnecessary results or information. The report for this assignment should contain two fitted regression models and the summary table of the variable selection procedure. For the regression model selected by using the *selection=rsquare* option in PROC REG you should include a full discussion of the model adequacy and all of the necessary graphics needed to understand the assessment of the model adequacy. The document should be submitted in pdf format.