Assignment #4: Data Analysis and Regression (50 points)

Data: The data for this assignment is associated with the Two Month's Salary case study.

Assignment Instructions:

In this assignment we build regression models for the Two Month's Salary case, building on what we have learned from previous assignments in the course.

(1) Define the Sample Population

- Define the appropriate sample population for your statistical problem.

(2) Exploratory Data Analysis and Simple Linear Regression

- Utilize statistical graphics to explore the data from the case.

- Define a simple linear regression model between a continuous explanatory variable and the response. Report a table with the coefficient estimates, t-values, p-values, etc.

(3) Multiple Linear Regression Model Specification

- Consider alternative model specifications for building a multiple linear regression for this case. What are the predictor/explanatory variables, and how should these variables be represented in a multiple linear regression model?

(4) Multiple Linear Regression Model Fitting

- Fit two multiple linear regression models, describe and evaluate the results. Explain the analysis of variance table for the analysis. Which predictors are statistically significant? Which predictors would you recommend to be included in models to be presented to the decision maker?

(5) Model Comparisons and Recommendation

- This model comparison should be its own section in your assignment report. Which model would you recommend to management, and how can management use this model in making decisions?

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 document should be submitted in pdf format. Name your file **Assignment4_YourLastName.pdf**

The assignment pdf file and accompanying plain text files for Python programs should be included in a zip archive using standard zip compression with the name **Assignment4_YourLastName.zip**

Here is a reasonable section outline for this assignment report.

Section 1: Define the Sample Population

Section 2: Exploratory Data Analysis and Simple Linear Regression

Section 3: Multiple Linear Regression Model Specification

- Section 3.1: Model #1 (Name and Describe Predictor Variables)
- Section 3.2: Model #2 (Name and Describe Predictor Variables)

Section 4: Multiple Linear Regression Model Fitting

- Section 4.1: Model #1 (Describe and Interpret Fitted Model)
- Section 4.2: Model #2 (Describe and Interpret Fitted Model)

Section 5: Model Comparisons and Recommendation