

STA321 Week #4 Assignment - MLR

Due: 11:30 PM, Sunday, 02/19/2023

Introduction

This assignment focuses on the multiple regression model using various techniques you learned from your previous courses. You will use the data set you selected last week. That data set will also be used for the next assignment that is based on this week's report.

Please study the first three sections of the class note. Your data analysis and write-up should be similar to what did in the case study in section 4 of the note. In fact, Section 4 can be considered a standalone statistical report.

Your analysis and write-up should have all components in Section 4.

- Description of your data set
- What are the research questions
- Exploratory analysis of the data set and prepare the analytic data for the regression
 - create new variables based on existing ones?
 - drop some irrelevant variables based on your judgment.
- Initial full model with all relevant variables and conduct residual diagnostic
 - special patterns in residual plots?
 - violation of model assumptions?
- Need to transform the response variable with Box-Cox?
- Want to transform explanatory variables to improve goodness-of-fit measures?
 - Please feel free to use my code to extract the goodness-of-fit measures. If you forgot the meaning of the goodness-of-fit measures, please check your old textbook or the eBook that I suggested on the course web page.
 - Several packages have a Box-Cox transformation function. The one that I used in the case study is from the library `{MASS}`. You can check the help document if you are sure how to use it.
- Build several candidate models and then select the best one as your model.
- Summarize the output including residual diagnostic plots
- Interpret the regression coefficient as I did in the case study.
- Conclusions and discussion.