ST 516: Midterm Project

Due by 8:30, Tuesday, October 8

Problem Statement:

> You have taken a new job working for the transportation department of the city of Washington, DC, and your boss wants to better understand the behavior of customers for the bike sharing system in the city. Data for the daily number or rides for both registered and non-registered users over a two year period has been collected, as well as other information about timing of the rides and weather conditions that may impact the customers' decisions to ride or not.
>
> Your job is to fit two (possibly sets of) models, one for the registered customers and another for the casual, non-registered riders to predict the number of rides on a given day. Your boss wants to prioritize the ability to accurately predict the number of rides for a given day, but thinks it would be nice to be able to explicitly state what factors influence ridership and by how much. Finally, she is interested to know if ridership grew from the first year to the second as well as any differences you notice in the behavior of the registered versus casual riders.

General Instructions:

- You will work in teams of four (a couple of teams could have +/- 1 team members since there are 82 students in the class). You should sign up for your teams on the Google Sheets link provided on Moodle.
- Your team will submit one written report no more than 5 pages in length in the following format:
    - Executive summary of your findings
    - Introduction (explain what you are trying to learn and why)
    - Data (description of the data and sources – you can use the UCI link provided to get this information)
    - Methods (describe the modeling and error estimation methods you used and why you chose them; this is also the section to discuss results of diagnostics and steps needed to correct problems, if any)
    - Results (discuss how well your model model(s) fit the data and enables you to interpret the relationship between individual predictors and response)
- One team member should submit a final report and accompanying R code using the following format
    - Report: pdf named in the format 'lastname_firstname_Proj1.pdf'
    - Code: R script named 'lastname_firstname_Proj1.r'.