MA678 Midterm Project Proposal

Runxin Yu

* + The reason why you are interested in this project

One reason is that I am very interested in the film industry. I have noticed that as time passed by, the film industry has become more and more commercialize with less and less artistry. Another thing that I realize is that in order to attract more viewers, more and more directors tend to hire celebrities with huge and solid fun bases without regarding the quality of the movie itself. So, does it really worth to devote a huge amount of money to a movie? Do celebrities with huge and solid fun bases truly have any contributions to the movie?

* + The Dataset you are going to use for the project
    - [Movie Industry](https://www.kaggle.com/danielgrijalvas/movies)
    - Mainly focusing on variables budget, company, genre, name, score, star, and director.
  + The question you are going to answer with this project
    - Do budgets have any association with the score of the movie? Is it true that the more money spends on the film the higher the score will be?
    - Do directors and main casts influence the score of the movie?
    - Are these associations differ between genres as well as production companies?
  + Potential problem that you foresee with this project
    - Other variables would affect the movie score such as the number of people voting, the gender of each voter, race of each voter, etc.
  + The timeline that you will work on this project
    - Start with basic data cleaning and initial EDA as soon as get the feedbacks
    - Finishing up the EDA by Nov.15th in order to see whether there are some other models that could be fitted to the data.
    - Then fit proper models, including a multilevel model, to the data and have it done by Nov.20th.
    - Go to office hours on Nov.21st in order to be sure that I am on the right pass and make some changes if needed.
    - Have the final version done by Nov.26th and submit it by Nov.30th.
  + Future direction that you might take this project to
    - Data analysis in the entertainment industries
    - (Tech companies?)