**MA 678 Midterm Proposal**

Jinyu Li

* **Personal Statement**: My interest is to work in tech/internet company, or consulting company as a data scientist/analyst or a consultant dealing with user portrait, behavior prediction, forecasting, something related with machine learning and so on. The industry I have interest in may be various now. I prefer to refine the fields and the industry that I love by taking some practices. Then I would like to choose something related with user portrait which I lay my interest on but haven’t set food on yet. Plus, I am also a potential game player. Consequently, I choose the data of steam store games to somehow have a basic sense of user behaviors and test my interest in this field.
* **Question:** Are there some differences of ratings of games among different types of games?

How variables like amounts of owner, play time, price of games and so on correlated with the ratings of games?

* **Data source:** Kaggle--<https://www.kaggle.com/nikdavis/steam-store-games>
* **Proposed Timeline of work:**

EDA: 11.5 – 11.9

Data Cleaning: 11.10 – 11.14

Modeling and Validation: 11.15 – 11.19

Write up: 11.20-11.25