Kendall Brown 8564403
Pstat 134 Final Project Reviews

1. Benjamin Fox, Meet Gala, Alex Gordee
    a. Depression and Suicide in the United States.
        i. Goal of identifying the most attributing factors to depression and suicide. Uses logistic regression to identify depression in people and by correlation how likely they are to commit suicide. Tiredness shown to be the most impactful predictor
        ii. Very interesting topic. Beginning of the presentation seems to have a lot of fluff, giving comparative statistics that are interesting but do not really provide anything outside of what was already stated in the first couple of slides.
        iii. Very thorough predictors chosen.
        iv. Acknowledgement of class imbalance.
        v. Why not exam standard of living instead of income as a measurement of depression?
        vi. Why not account for world events such as wars and economic depressions?
        vii. What percentage of overall variance can be described by tiredness.
        viii. If I was working on this project I would like to tackle other health problems such as obesity, heart disease, and cancer.
        ix. Overall, really solid presentation. Going forward I would just like to see a more in-depth analysis into health problems and why they spike after year 2000.

2. Michael Chiang, Harvey Lao, Christian Taruc, Patrick Vidican
    a. Movie Genre Classification
        i. If movie posters can be used as a predictor variable for movie genre classification. Data gathered from Kaggle. Identified Drama and a number of other genres as being redundant and limited the total number of genres. Used a computational neural-network to classify movie posters. Achieved an accuracy of around 60%. Identified that artistic meaning is hard to train into machine learning processes.
        ii. I really liked how it was shown that RGB and HUE values were practically useless. Additionally, the use of neural network was interesting as I have never seen one used before.

iii. Would it not have been easier to build a classification model if the on poster text data was used to mine more predictor variables?
iv. I would have liked to see if it was possible to predict box office revenue could be predicted from the poster. For instance if it can be discovered that a major studio is pushing a movie with multiple A-list actors then can we infer a high earnings?
v. I would like to seperate the movie posters by some sort of 'era' peramator. I can see how a movie poster from the 40s, 80s, and 2010s might have different artistic styles could lead to pad prediction if measured against the styles of different eras.

3. Sam ONeill, Aaron Plesset, Xiaoxiong Xu, Xiaoyue Zhu
    a. Beer Advocate
        i. Pulling Data from a custom BeerAdvocate.com to find the highest rated beer according to user reviews. Used user-to-item and item-to-item correlations to find recommendations for specific users. Trained and tuned prediction models to achieve the most accurate models.
        ii. I liked the use of correlations between items and users as it added a level of depth to finding what specific users would like.
        iii. I also liked the result of the project and how it was able to accurately recommend beers to users.
        iv. If a beer had too few reviews it was excluded, It would have been neat to build groups around users that did rate those reviews and find other correlated users to base recommendations. For instance, if user A liked an unpopular beer, and user B is correlated to user A, then it might be appropriate to recommend user B the unpopular beer.
        v. As far as changes I would make to the project there aren't any I could think of as of now, the methodology used is rather flexible and can be extrapolated to other recommendation services.