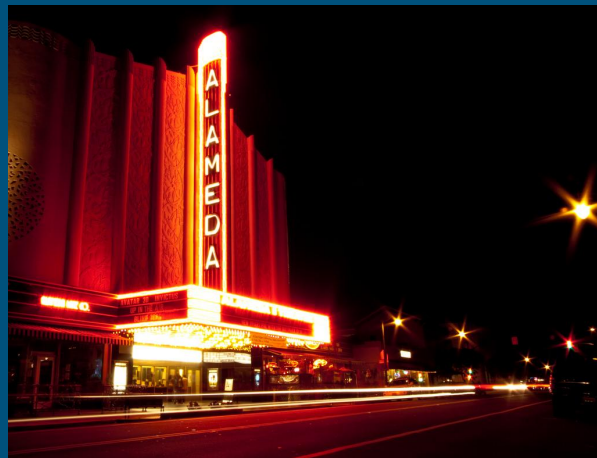# Ratings Effects on Revenue

Samuel Wong

Background

- Theaters choose which movies to show

- Limited number of theater space

- Goal is to maximize profitability

- Exists a point to decide when to stop showing a movie and bring in another

- Being able to predict the potential revenue of a movie can help with the decision

Initial Model Attempt

- Gathered data from Box Office Mojo, IMDB, and Rotten Tomatoes

- Ratings and gross domestic revenue (Box Office) for top 4000 movies

- Examined features including the critic ratings, user ratings, number of reviews, and Tomatometer linearly

- Not very predictive

Feature Selection

- Removed features that did not have much predictive power

- Noticed the number of user reviews had both small and very large numbers

- Decided to add extra features besides "rating-like" ones

  - MPAA Rating (dummy)

  - Big 8 Movie Studios (dummy)

- Much more of the variation of revenue explained

Deciding on a Model

- Tried the following models:
  - Lasso
  - Ridge
  - ElasticNet
  - Gradient Boosting
  - Random Forest Model

- Obtained the highest R squared value of using the Random Forest Model

- Ultimately, R squared value of **0.47** vs. original model value of 0.19

Predictive Power

Observed vs Predicted Domestic Total Gross

Residual Plot of Rating Feature
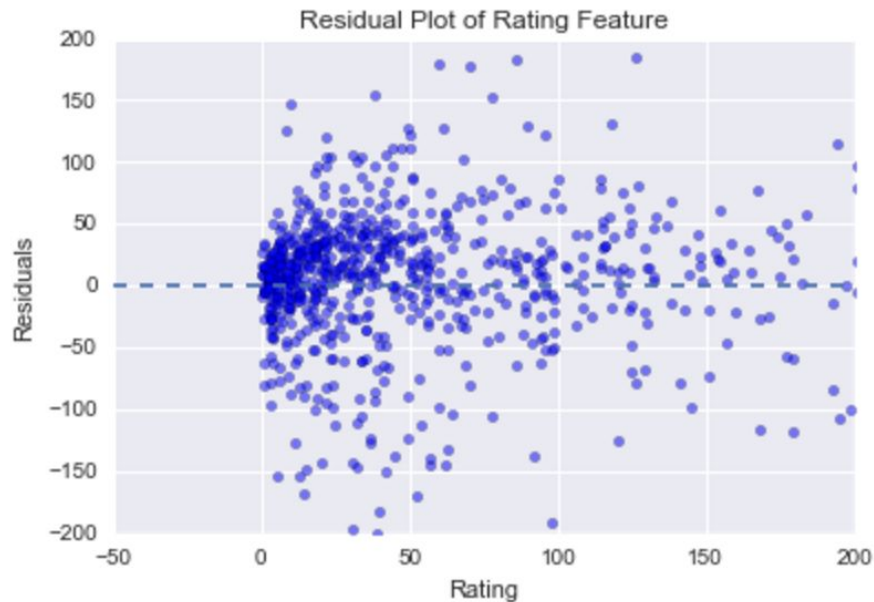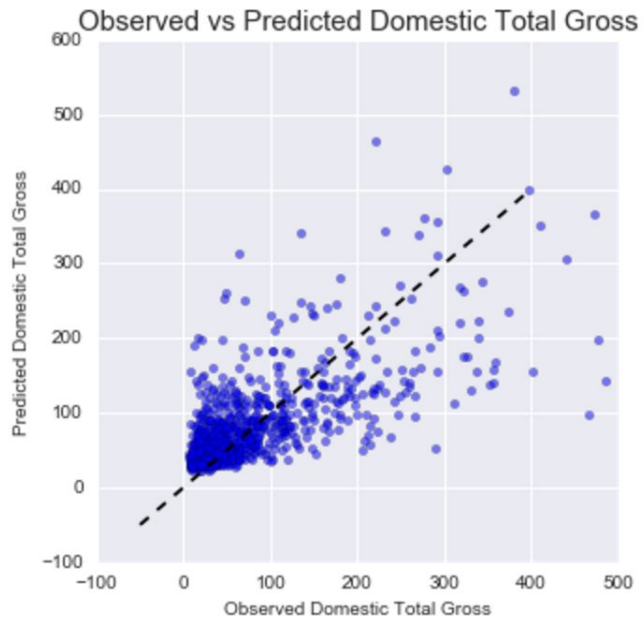
Future Improvements

- Examine the point in time where a movie no longer is as profitable
  - Value the final stream of revenue
  - Account for other factors such as reputation and contract provisions, series

- Modeling Improvements
  - More features
  - Interaction terms
  - Higher polynomials