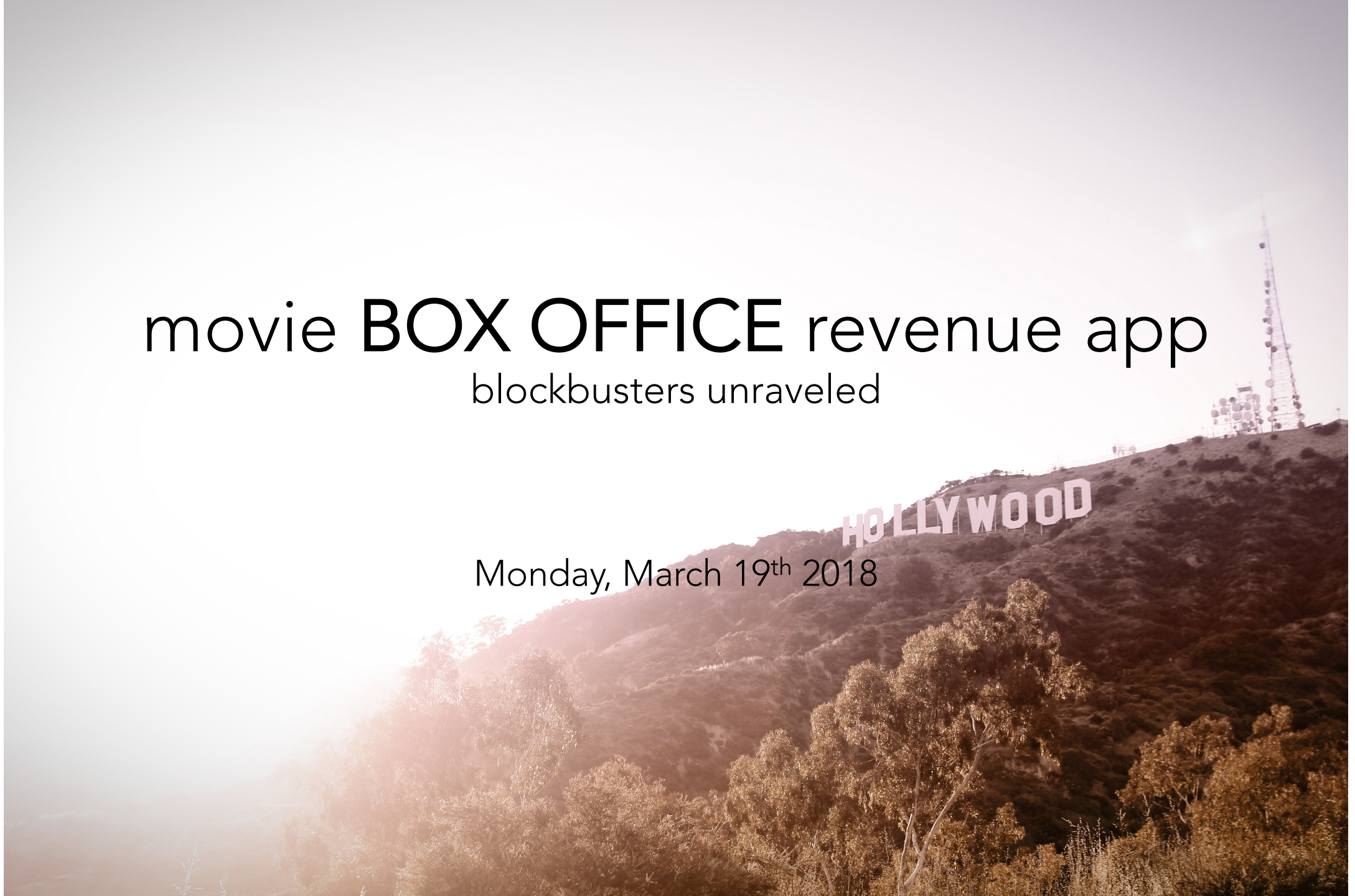
# movie BOX OFFICE revenue app

blockbusters unraveled

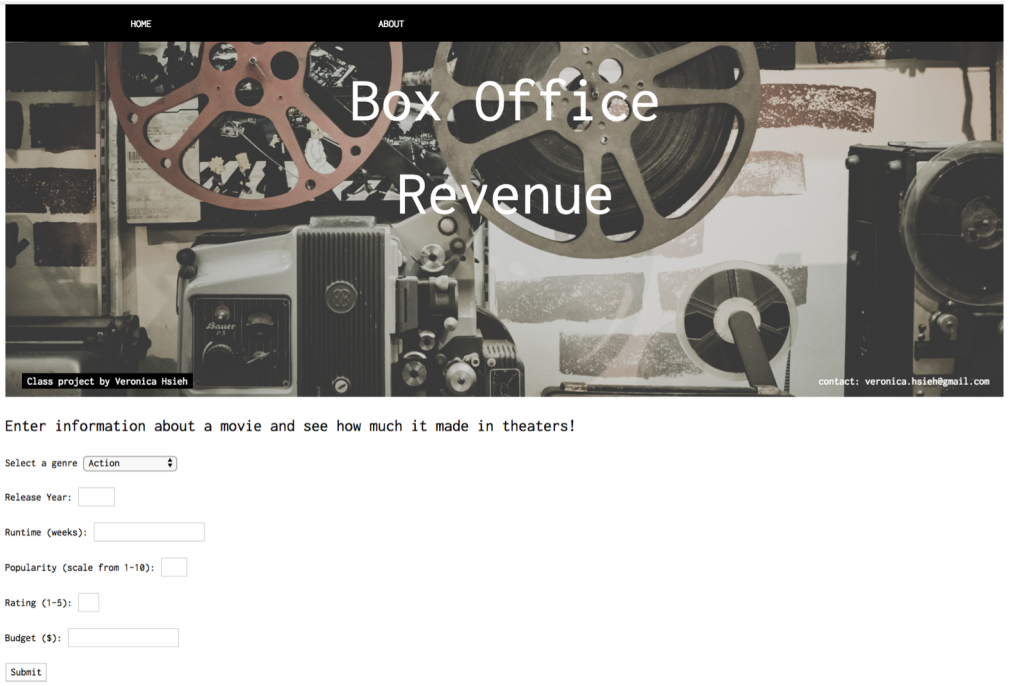



Monday, March 19th 2018

# MOTIVATION FOR APP

GOAL

> Build a model to accurately predict movie revenue




REASON

> Understand what factors lead a movie to be a hit

# DATA



MovieLens dataset from Kaggle

Original dataset had 45,467 observations & 24 columns

After EDA & data cleaning, 5369 observations were used for model building & development

Transformations:

- Log scaling
- Categorical variables to dummies

# PREDICTION MODEL



DecisionTreeRegressor from Sklearn

Movie metadata used:

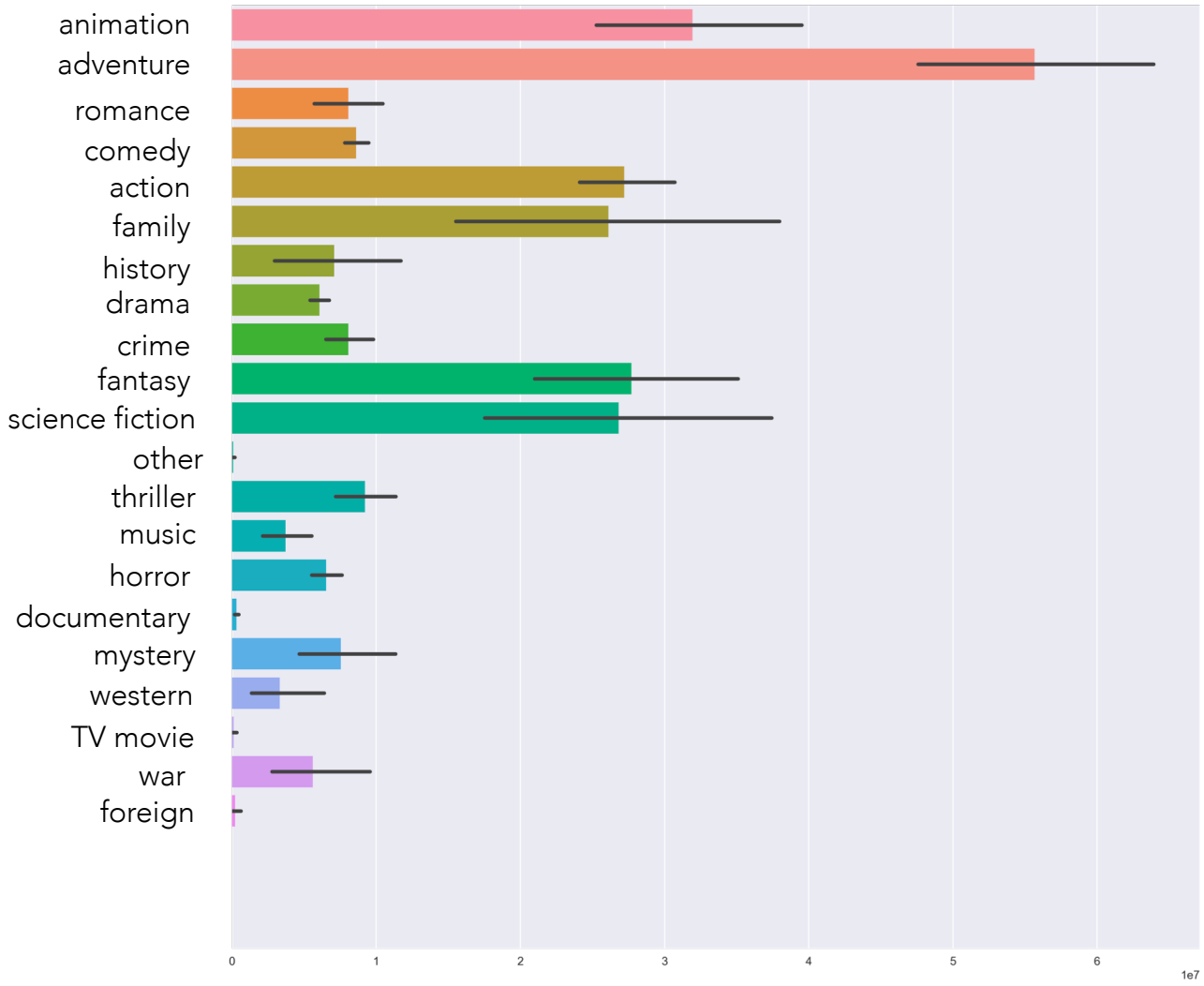
- Genre
- Runtime
- Budget
- Release Year
- Popularity
- Rating

Train-Test split: 80/20

Evaluation Metric: R-squared

# INSIGHTS

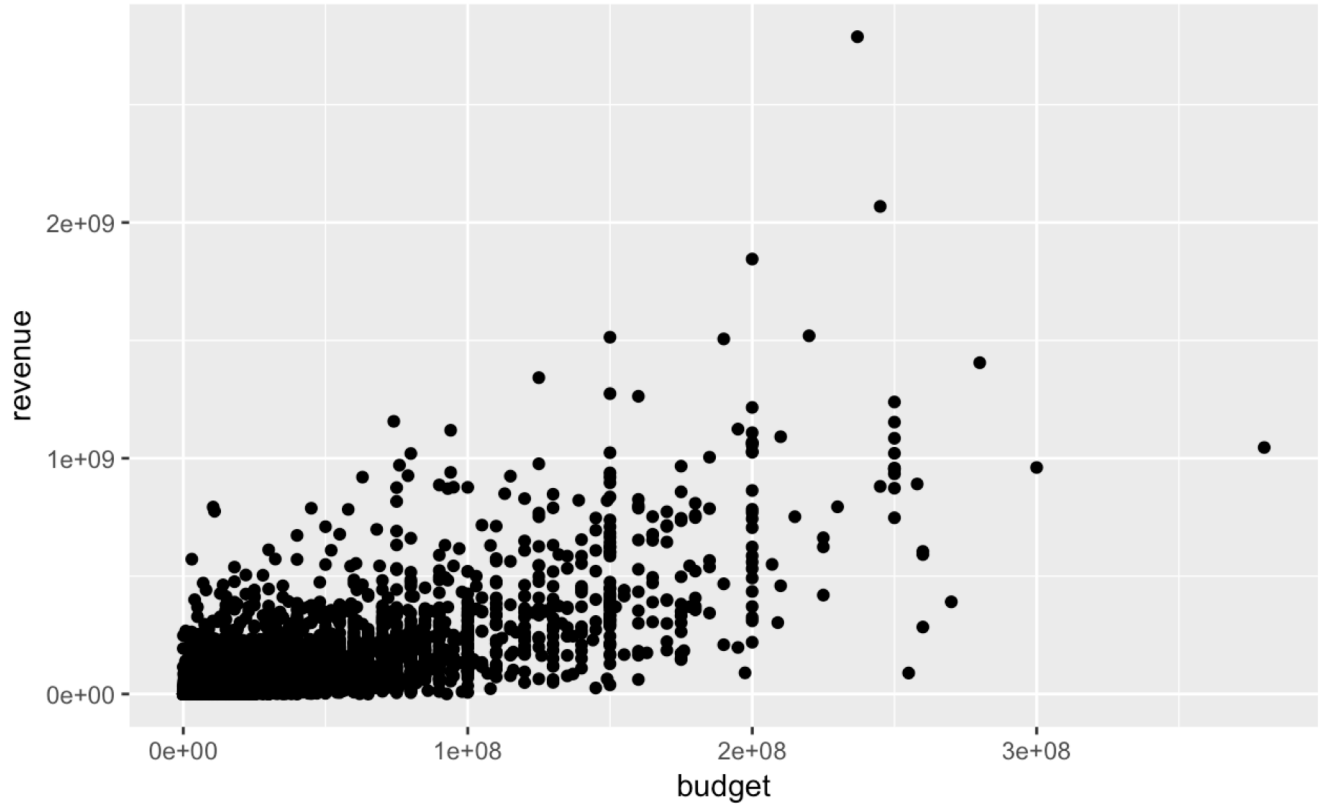
## Movie goers enjoy watching genres that take them outside of reality




*Revenue by genre*

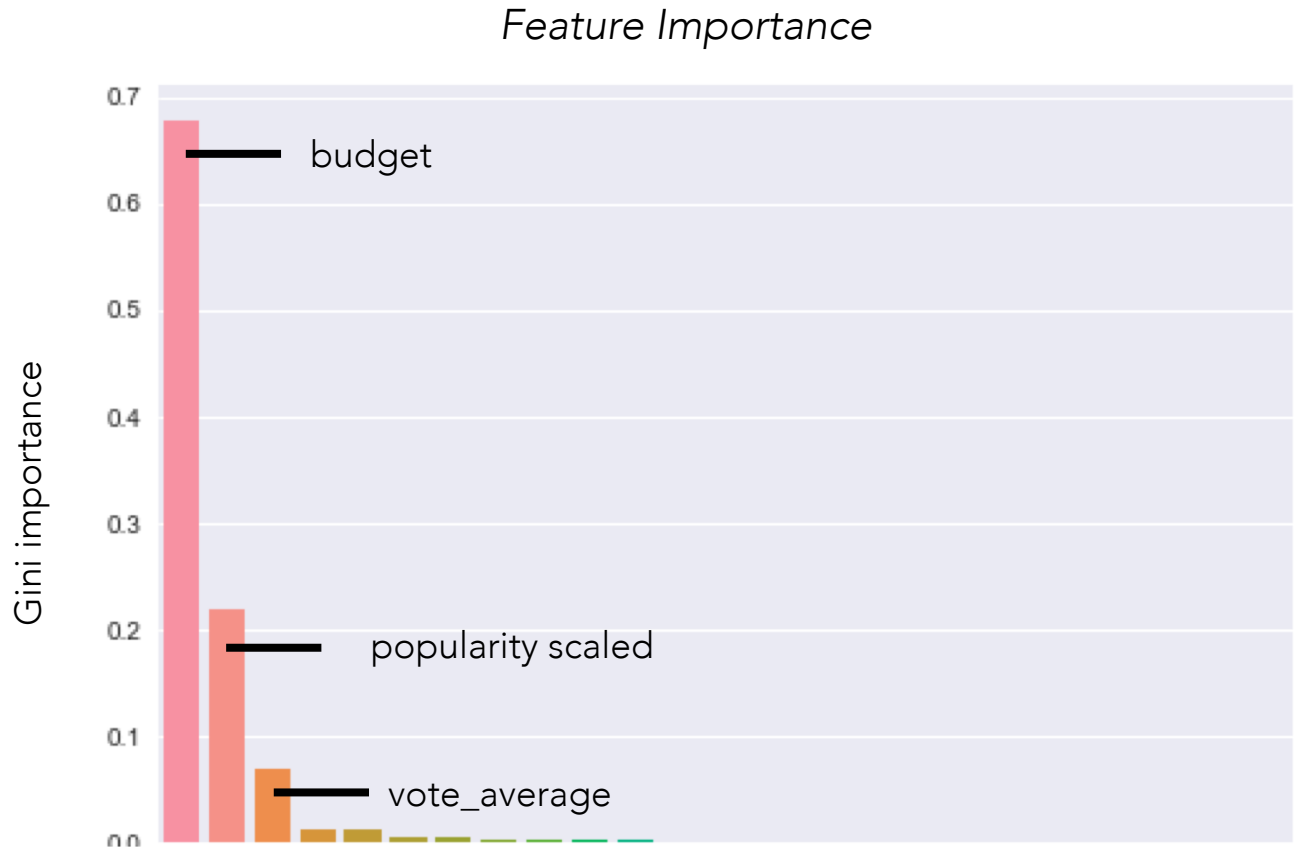
# INSIGHTS

Movie revenue and budget have a strong correlation of .73



# INSIGHTS

Movie budget and public perception of the movie are most influential in determining a movie's financial success



# THANK YOU


GitHub @ vhsieh920

email veronicahsieh2018@u.northwestern.edu