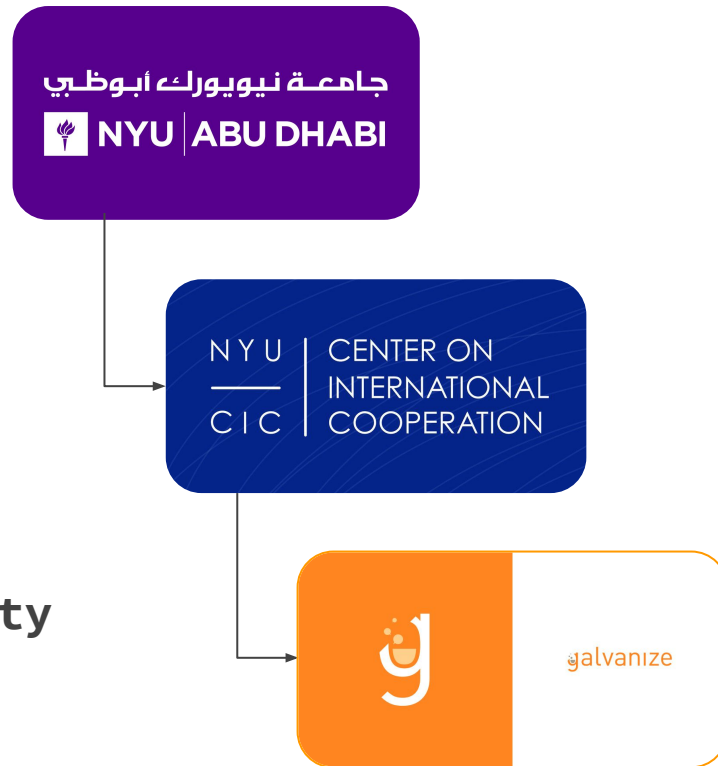# Predicting Film Success

Ryan Rappa

Me / my project

___

- **Poli sci + econ background**

- **Interest in film & entertainment**

- **Project: predict film profitability**

Project workflow

Profits vs. cast, writer, director
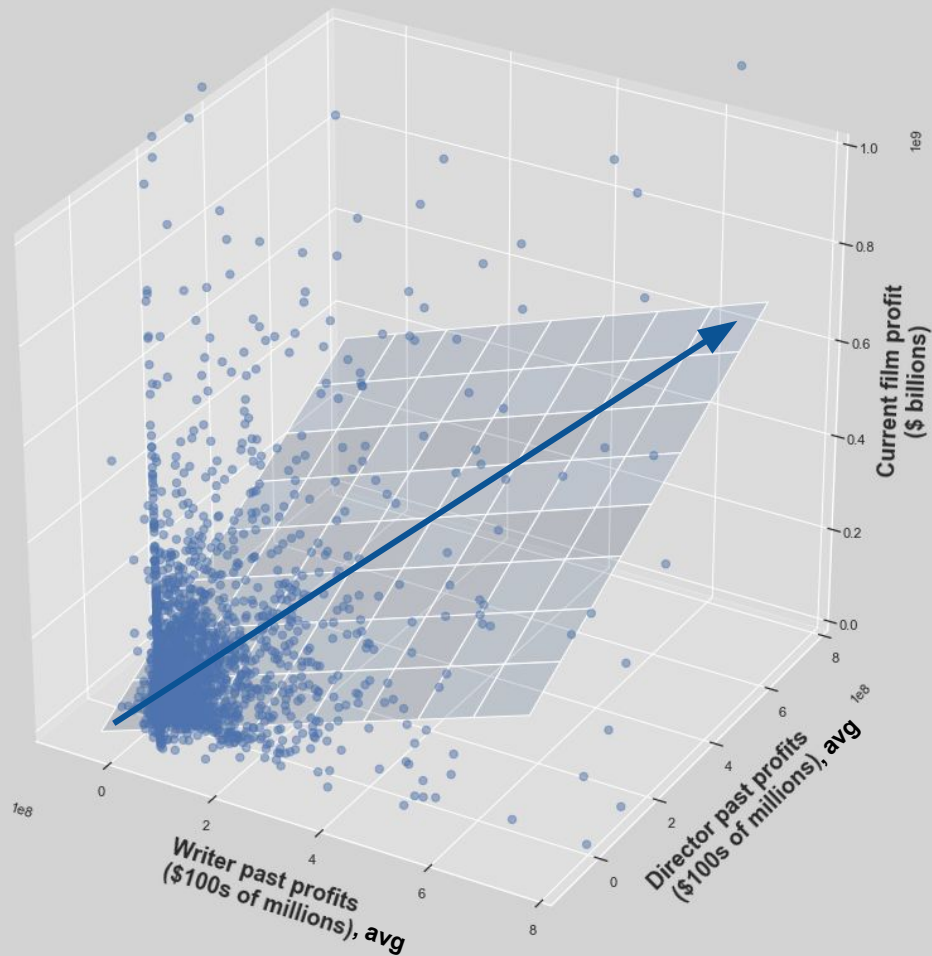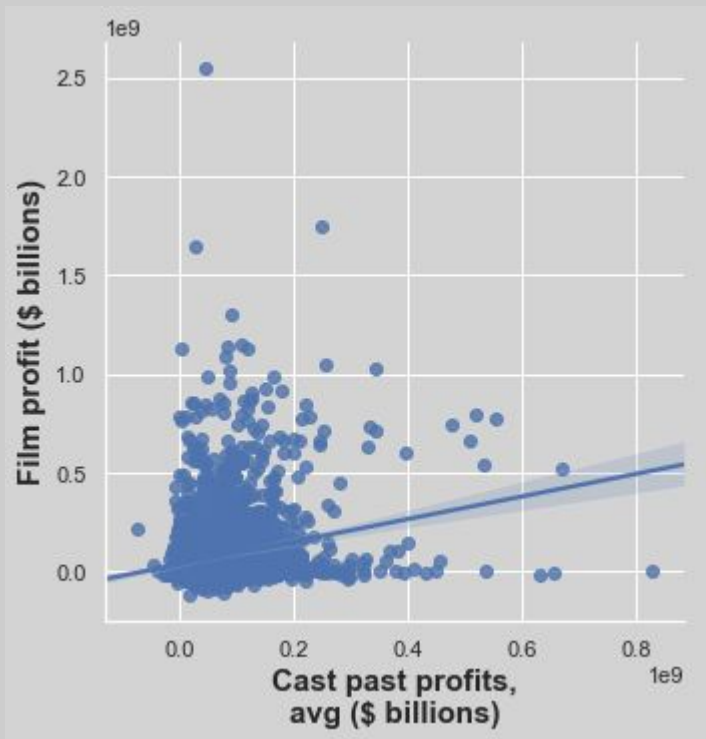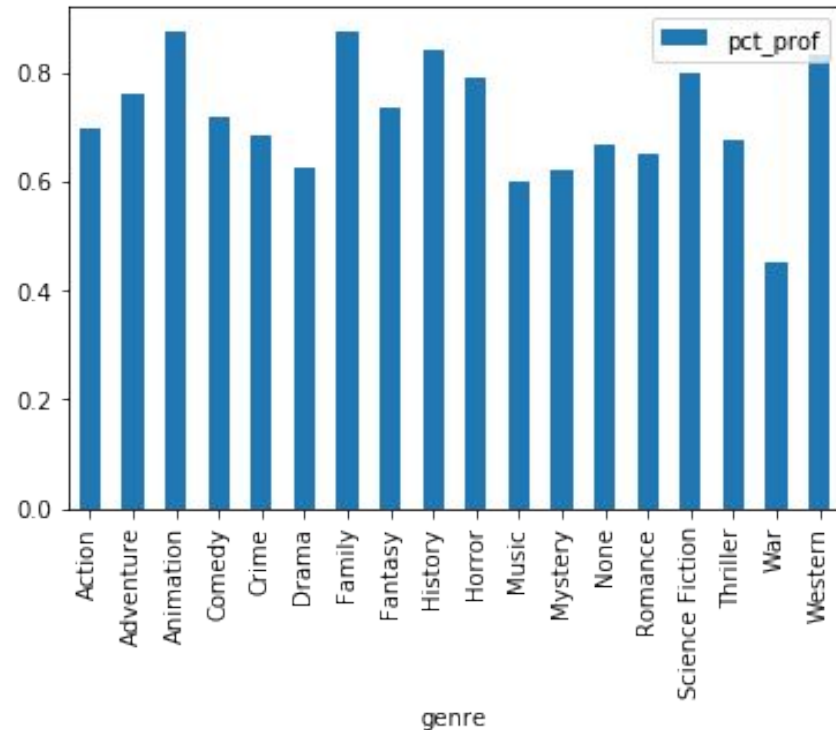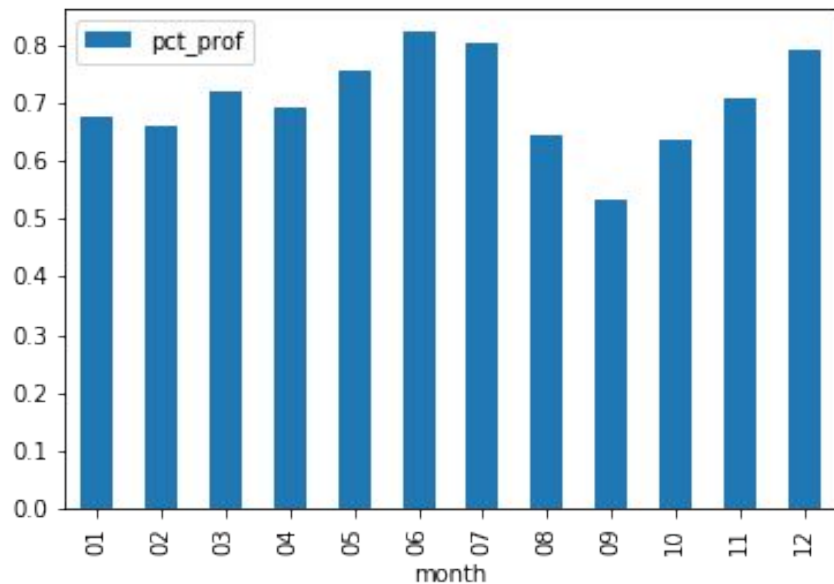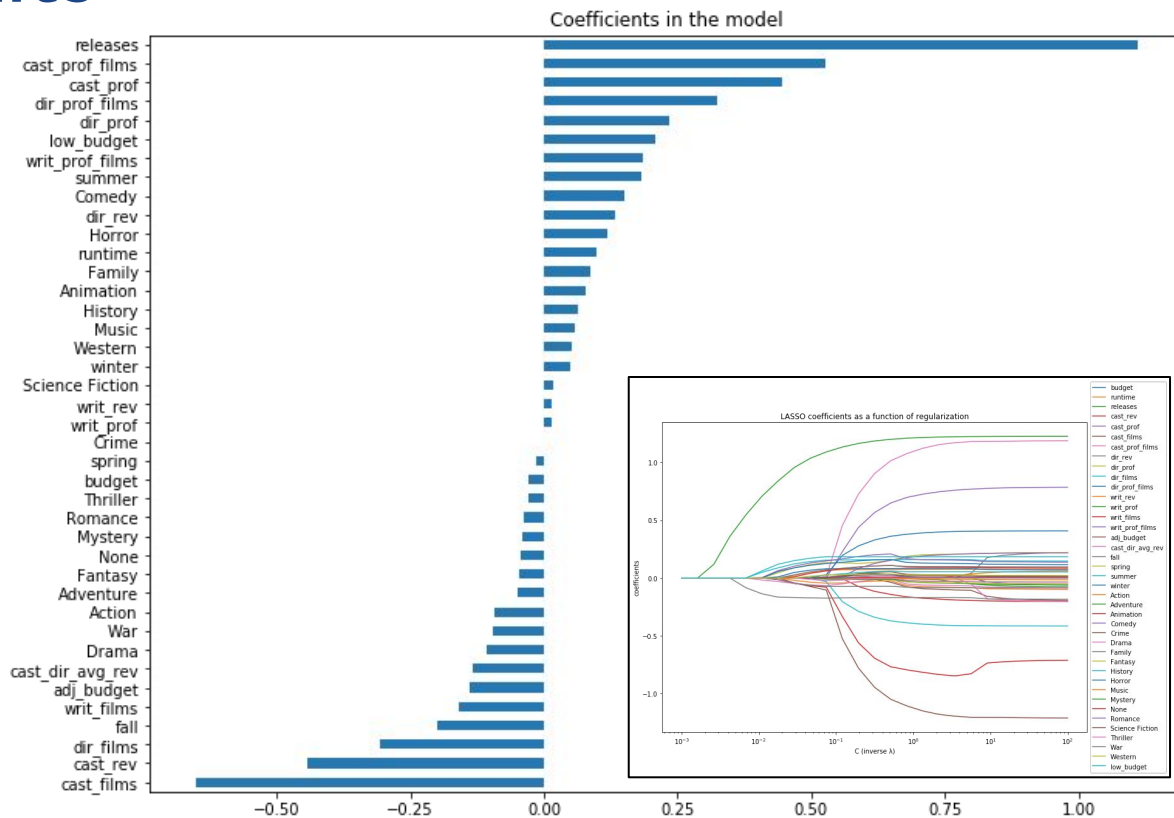
Profitability by month, genre

Robust coefficients

Predictors of success:

- ○ Int'l releases
- ○ Cast past success
- ○ Director past success
- ○ "Low Budget" (<= $3m)
- ○ Season
  - ■ May-Jul best
  - ■ Aug-Oct worst
- ○ Genre
  - ■ Animation/Family
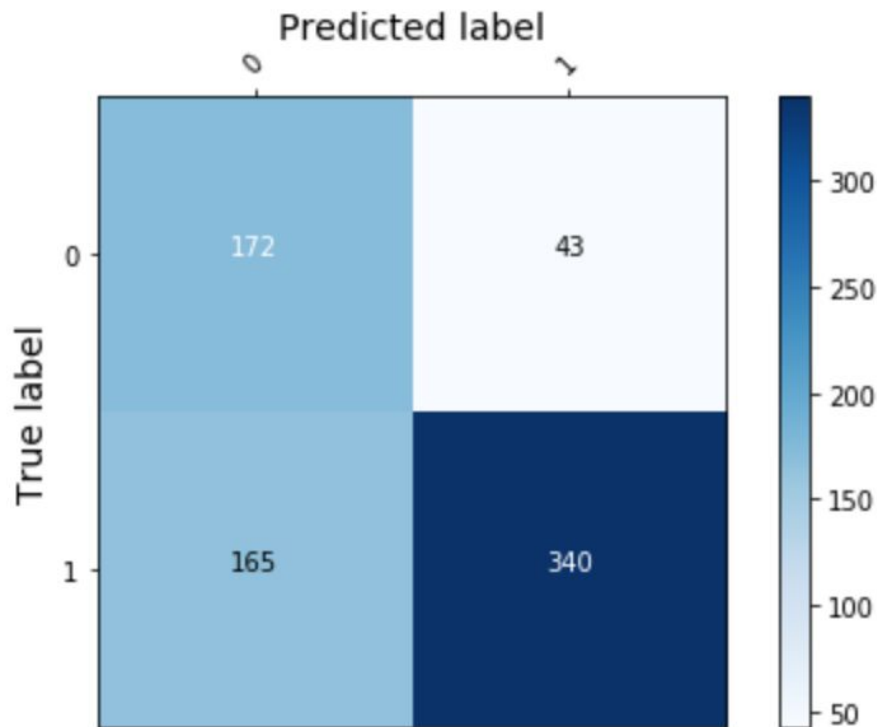 - ■ Horror

Coefficients in the model

Robust predictions

XGBoost w/ 75% decision boundary:

- **Accuracy: 71%**

- **Precision: 89%**

- **Recall (TPR): 67%**

- **Fallout (FPR): 20%**

```
---- Classification Report ----
             precision    recall  f1-score   support

          0       0.51      0.80      0.62       215
          1       0.89      0.67      0.77       505
```
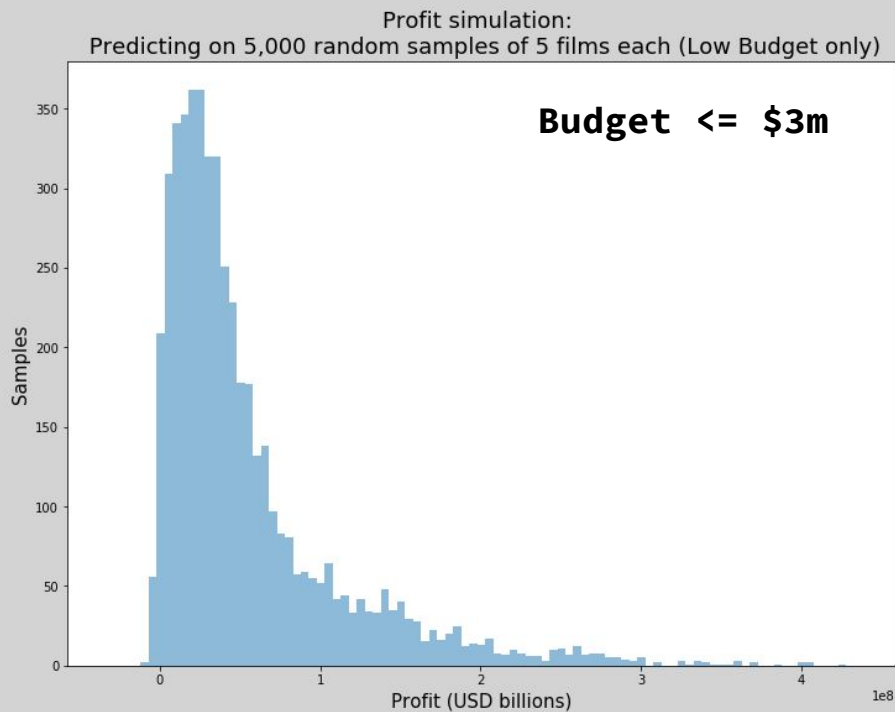
Profit simulation:
Predicting on 5,000 random samples of 5 films each

Percent of samples with positive profit: 0.9546
Median (USD millions): +249.6

Profit simulation:
Predicting on 5,000 random samples of 5 films each (Low Budget only)

Budget <= $3m

Profit simulation:
Predicting on 5,000 random samples of 5 films each (higher budget only)

Budget > $3m
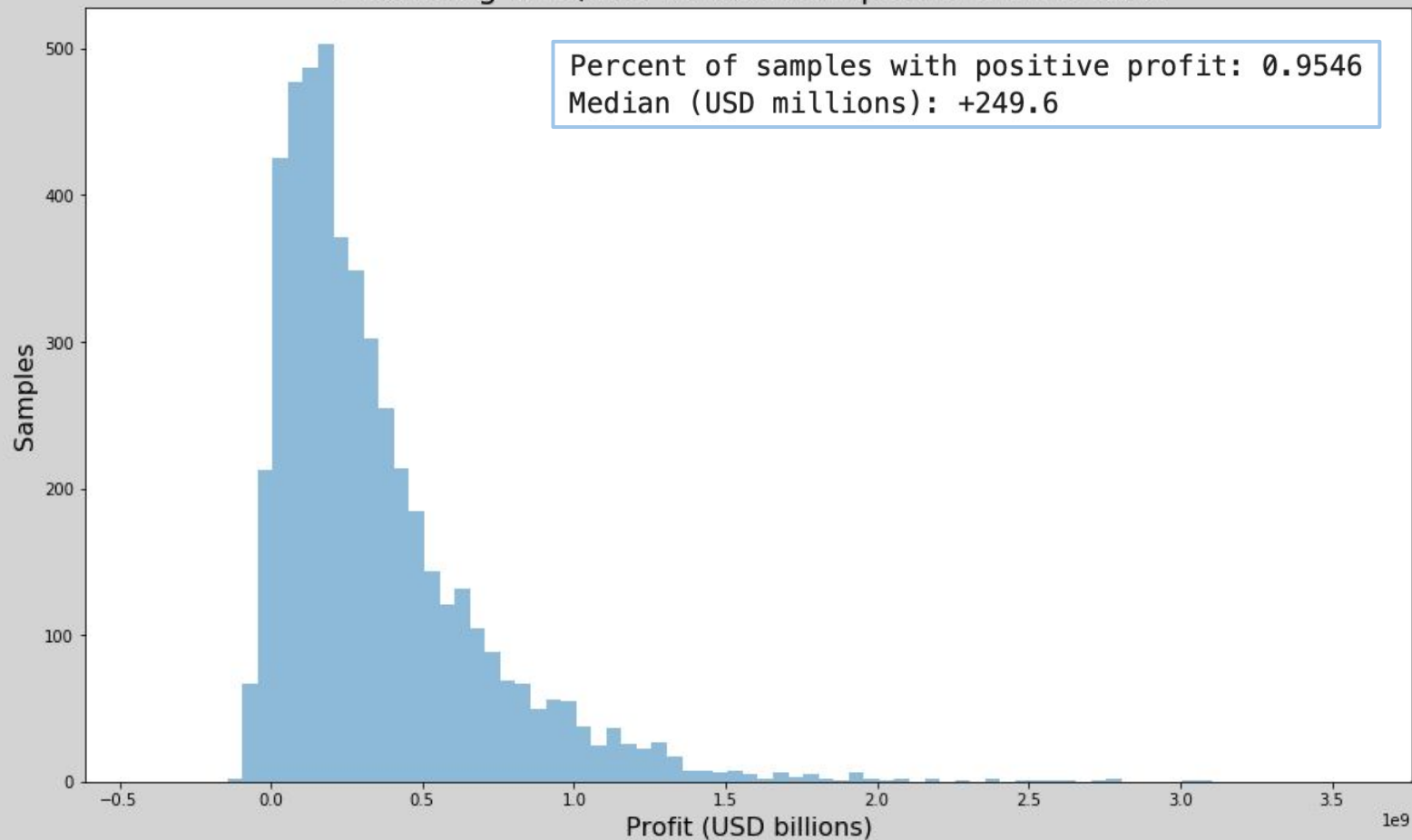
Percent of samples with positive profit: 0.9744
Median (USD millions): +36.1
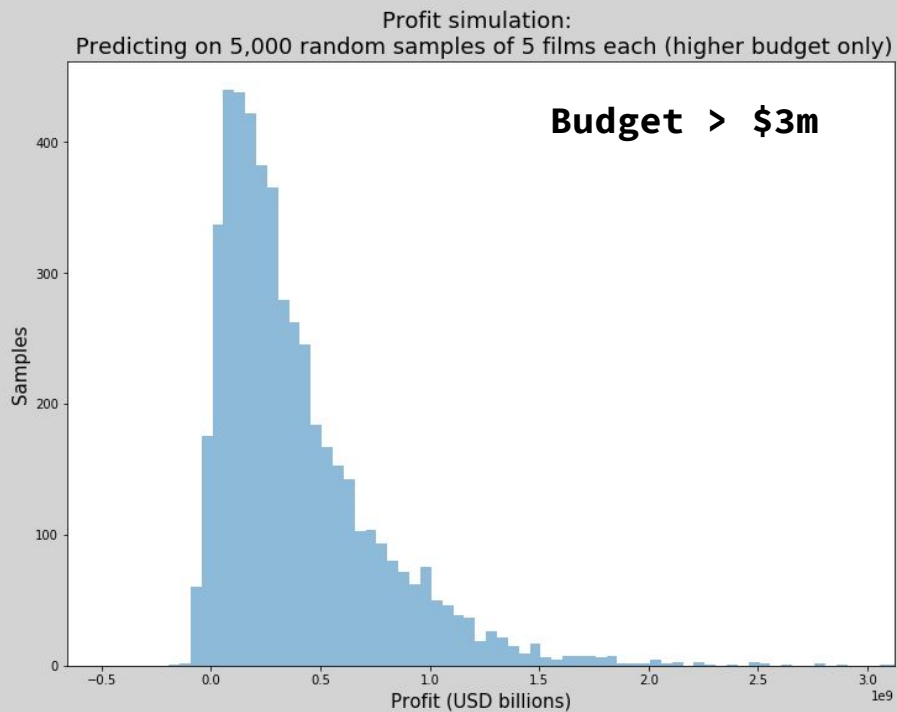
Percent of samples with positive profit: 0.958
Median (USD millions): +290.8

Next steps / variables to consider

— — —

- **Collect more data**
- Story/script
- Posters
- Trailers
- Marketing strategy
- Domestic vs. foreign box office
- Ancillary revenues
- Piracy

Thank you
— — —