Revenue predictive regression model

By Yordanos W.

Purpose

- Movie directors and investors
- Factors to consider
- Predict profitability

The Data

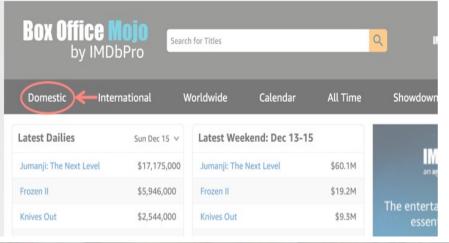
Data scraping

Beautiful soup

Websites Boxoffice mojo, IMDB

Movies 2000 - 2019 releases

Data points 2002 movies





Variables

Features Budget , Distributor, genre, Running time , Rating

Target Revenue

Project methodology

- Data scraping
- EDA
- Baseline model
- Feature engineering
- Model development and fitting
- Model evaluation
- Model improvements

Model selection

Lasso mode

 $R^2 = 0.47$

MAE = 0.66

What does it tell us?

Budget, running time - significant positive influence

Drama, R-rating, Other distributors - negative influence

Remarks

- Time sensitive
- Exceptions





Thank You!