

# REGRESSION PROJECT

---

# EXPORTING AMERICAN MOVIE ADAPTATIONS



## INTRODUCTION

- ▶ **Motivation:** Movies Worldwide, Inc., wants a decision tool to guide their choice on which American movie adaptations to export.
- ▶ **Research Question:** Can a model predict international gross revenue based on readily available movie data?
- ▶ **Recommendation:** Explore country-specific box office sales to predict understand what movie factors contribute to high international gross revenue.



[Photo by Photo by Alamy.com](#)

# METHODOLOGY

- ▶ Web-scraped movie dataset from [boxofficemojo.com](https://www.boxofficemojo.com)

- ▶ Adaptations, include:

- ▶ Books, television shows \*

- ▶ Plays

- ▶ Events, video games

- ▶ Movies from 1978 to 2022

\*Books include young adult novels, contemporary novels, children's books, and comic books. Television shows include children shows and cartoons.

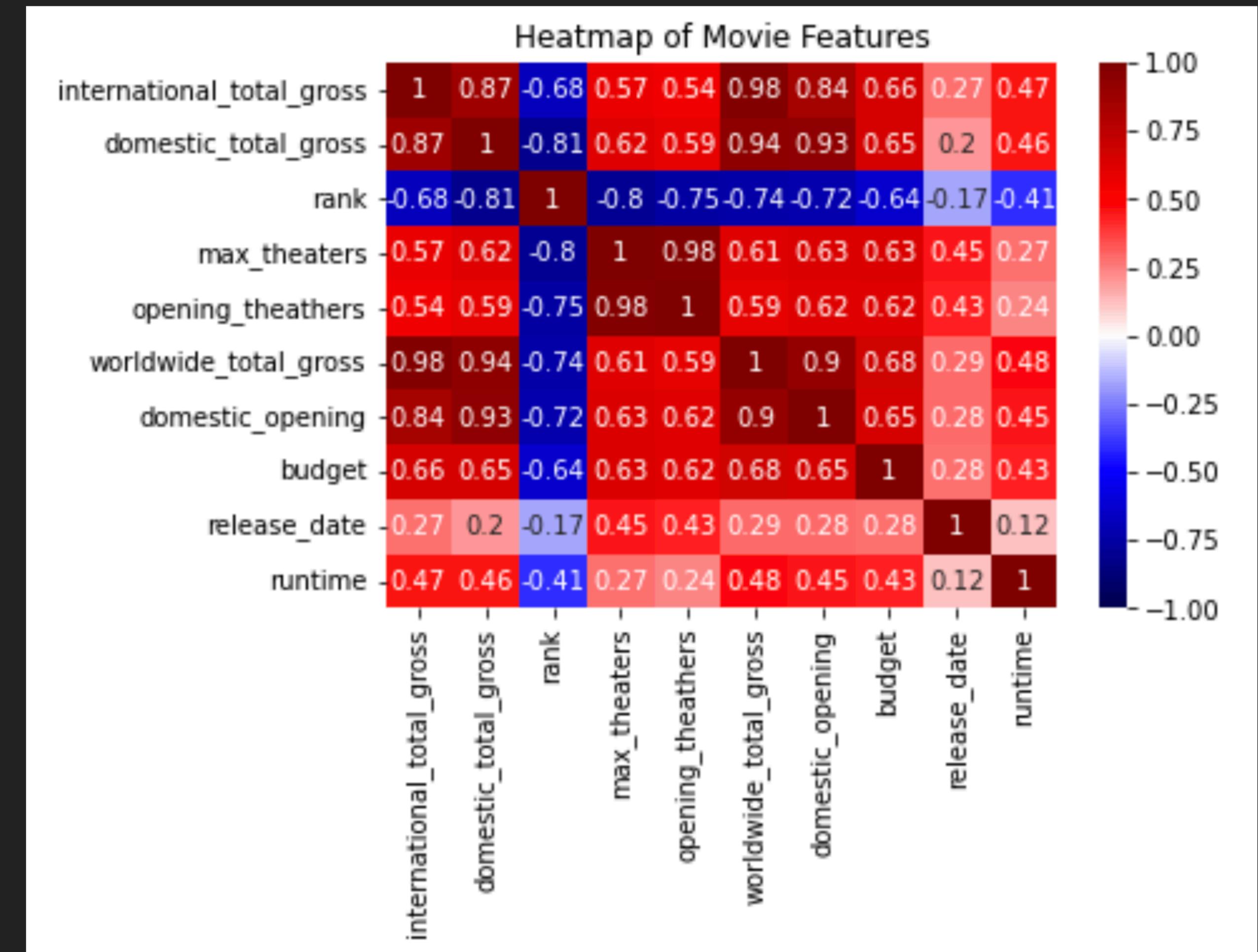
The screenshot shows the Box Office Mojo website interface. At the top, there is a navigation bar with the site's logo, a search bar labeled 'Search for Titles', and social media links for Facebook and Twitter. Below the header, the text 'Genre Keyword: Adaptation' is displayed. A table follows, listing three movies along with their respective box office metrics: Rank, Title, Lifetime Gross, Max Theaters, Opening, Open Th, Release Date, and Distributor.

Rank	Title	Lifetime Gross	Max Theaters	Opening	Open Th	Release Date	Distributor
1	Avengers: Endgame	\$858,373,000	4,662	\$357,115,007	4,662	Apr 26, 2019	Walt Disney Studios Motion Pictures
2	Spider-Man: No Way Home	\$794,343,553	4,336	\$260,138,569	4,336	Dec 17, 2021	Sony Pictures Entertainment (SPE)
3	Black Panther	\$700,059,566	4,084	\$202,003,951	4,020	Feb 16, 2018	Walt Disney Studios Motion Pictures

# RESULTS

- ▶ Known **high** multicollinearity
- ▶ Log transformation:
  - ▶ **domestic total gross**
- ▶ Feature engineering:
  - ▶ **Profit = domestic gross - budget**
  - ▶ **Opening profit = domestic opening - budget**

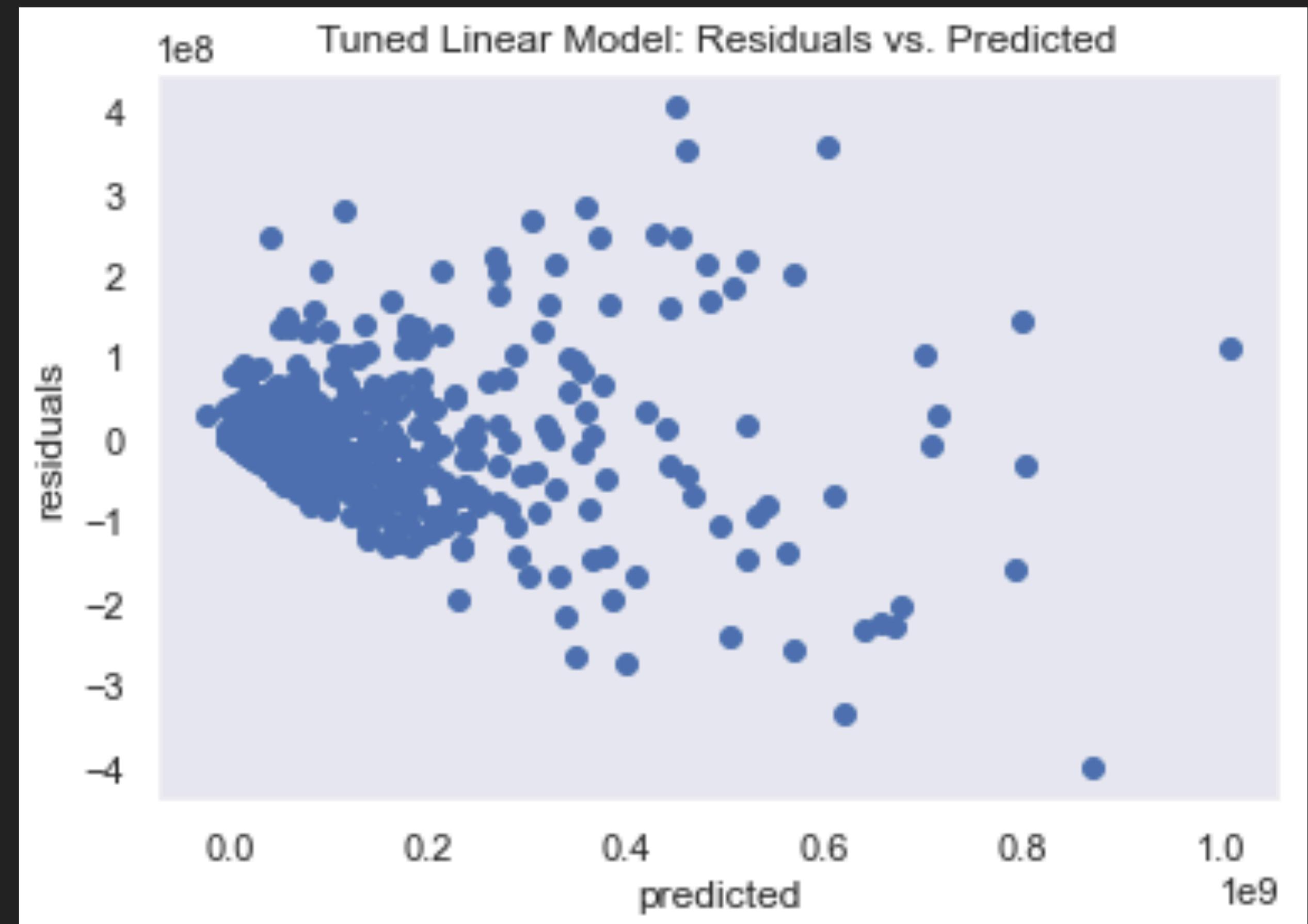
Figure 1.



# RESULTS

- ▶ Evaluating with  $R^2$ :
  - ▶ Ridge regression model
  - ▶ Train ~ test
  - ▶ **good fit** 
- ▶ Double checking with MAE:
  - ▶ Linear model better ....  
but off by \$45.5 million
  - ▶ Tuning lowered error...  
by to \$2.8 million 

Figure 2.



## CONCLUSIONS

- ▶ This model will not predict international movie gross.
- ▶ Recommend building model based on movie gross data from intended markets.
- ▶ Recommend exploring subset of movies with high international gross and looking for trends or characteristics that contribute to box office sales.



Photo by Kristi Luhaers on Unsplash

## FUTURE WORK

- ▶ Further EDA to look for patterns among high international grossing movies.
- ▶ Scrape country specific gross data to build model.
- ▶ Explore relationship between bestselling books and their movie counterpart international gross.



Photo by Kentaro Toma on Unsplash

## APPENDIX

- Blog post, code, and slides are available at [github.com/slp22/regression-project](https://github.com/slp22/regression-project)

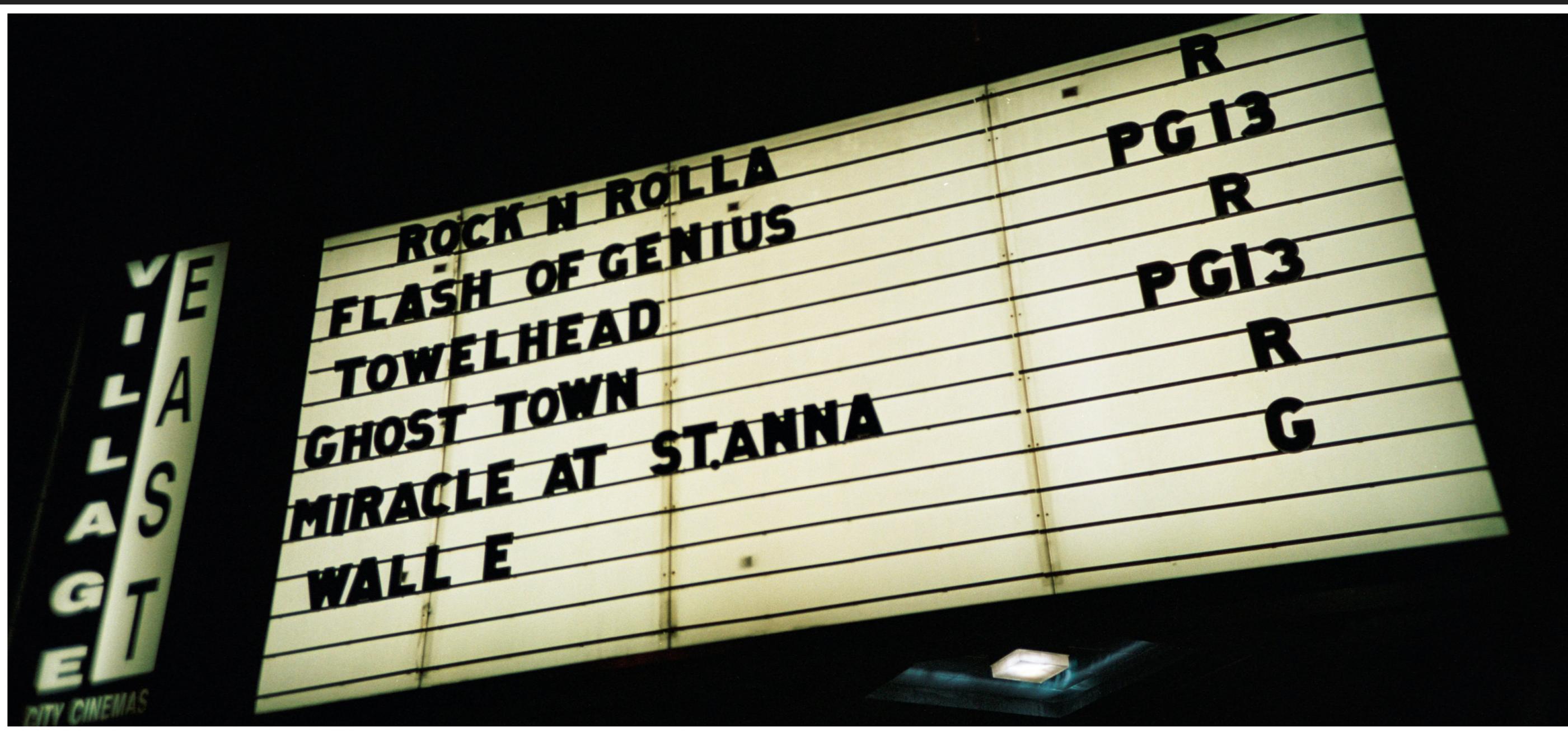


Photo by Gemma Evans on Unsplash