

PREDICTING WORLDWIDE MOVIE REVENUE

Ignasi Sols
October 1st, 2021

Why I am interested in cinema

IMDb Menu All Search IMDb

Luis Lucia (1914–1984) SEE RANK

Writer | Director | Production Manager

+ Add or change photo on IMDbPro »

Luis Lucia was born on May 24, 1914 in Valencia, Valencia, Comunidad Valenciana, Spain as Luis Lucia Mingarro. He was a writer and director, known for *Zampo y yo* (1966), *El príncipe encadenado* (1960) and *Jeromín* (1953). He died on March 15, 1984 in Madrid, Spain. See full bio »

Born: May 24, 1914 in Valencia, Valencia, Comunidad Valenciana, Spain

Died: March 15, 1984 (age 69) in Madrid, Spain

4 wins. See more awards »

Known For

Zampo y yo
Writer (1966)

El príncipe encadenado
Director (1960)

Jeromín
Writer (1953)

Molokai, la isla maldita
Writer (1959)

Google luis lucia películas

All News Images Videos Maps More Tools SafeSearch on

Luis Lucia / Movies Most popular first

Tómbola 1962 Ray of Light 1960 An Angel Has Arrived 1961 Las cuatro bodas de M... 1967 Zampo and Me 1966 Canción de Juventud 1962 Morena Clara 1954 Sister San Sulpicio 1952 Solos los do 1968

Feedback

Business case:



Streaming



Theatres



Methodology

-BeautifulSoup,
Selenium

-At least 200
votes on IMDb

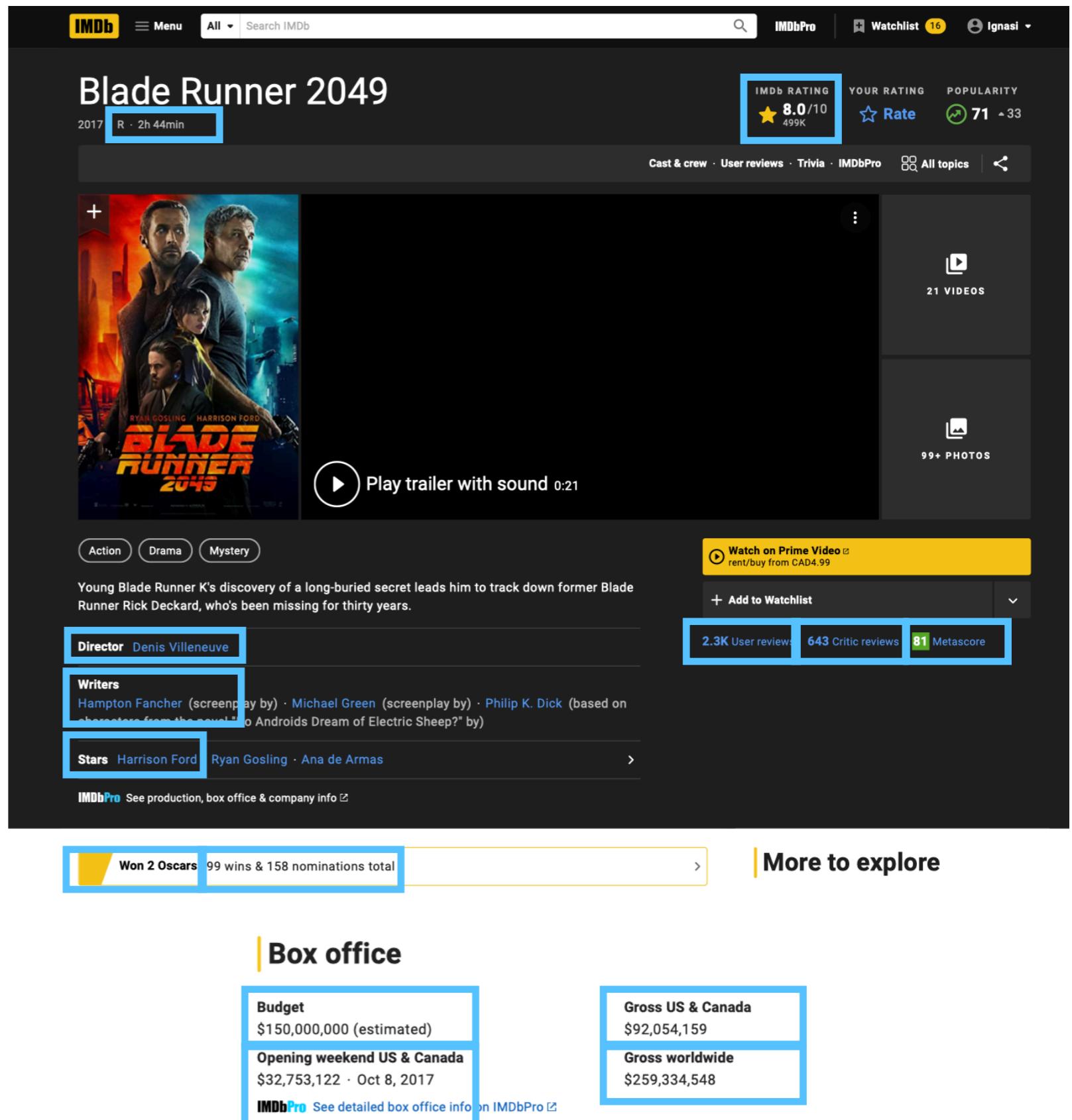
-US production

-English

-5719 websites
Web-scraped



-914 movies selected



The screenshot shows the IMDb movie page for "Blade Runner 2049". At the top, it displays the title "Blade Runner 2049", the year "2017", the rating "IMDb RATING 8.0/10 499K", and the popularity "71 33". Below the title is a large movie poster featuring Ryan Gosling, Harrison Ford, and Ana de Armas. A "Play trailer with sound" button is visible next to the poster. Underneath the poster, there are genre tags: Action, Drama, Mystery. A plot summary follows: "Young Blade Runner K's discovery of a long-buried secret leads him to track down former Blade Runner Rick Deckard, who's been missing for thirty years." Below the summary are sections for "Director Denis Villeneuve", "Writers Hampton Fancher (screenplay by), Michael Green (screenplay by), Philip K. Dick (based on characters from the novel 'Do Androids Dream of Electric Sheep?' by)", and "Stars Harrison Ford, Ryan Gosling, Ana de Armas". There is also a link to "IMDbPro See production, box office & company info". Further down, a box highlights "Won 2 Oscars" and "99 wins & 158 nominations total". To the right, there is a "More to explore" section with "Box office" and "Gross US & Canada" information.

Blade Runner 2049

2017 R · 2h 44min

IMDb RATING 8.0/10 499K

YOUR RATING Rate POPULARITY 71 33

Cast & crew · User reviews · Trivia · IMDbPro · All topics

21 VIDEOS

99+ PHOTOS

Action Drama Mystery

Young Blade Runner K's discovery of a long-buried secret leads him to track down former Blade Runner Rick Deckard, who's been missing for thirty years.

Director Denis Villeneuve

Writers Hampton Fancher (screenplay by) · Michael Green (screenplay by) · Philip K. Dick (based on characters from the novel "Do Androids Dream of Electric Sheep?" by)

Stars Harrison Ford · Ryan Gosling · Ana de Armas

IMDbPro See production, box office & company info

Won 2 Oscars 99 wins & 158 nominations total

Box office

Budget \$150,000,000 (estimated)

Opening weekend US & Canada \$32,753,122 · Oct 8, 2017

Gross US & Canada \$92,054,159

Gross worldwide \$259,334,548

Watch on Prime Video rent/buy from CAD4.99

Add to Watchlist

2.3K User reviews 643 Critic reviews 81 Metascore

More to explore

Methodology

Baseline model:



Added more features

-budget

-runtime

-Top actor?

-runtime

-Top actor?

-budget

-Action
-Adventure
-Comedy
-Crime
-Drama

-G
-PG-13
-R

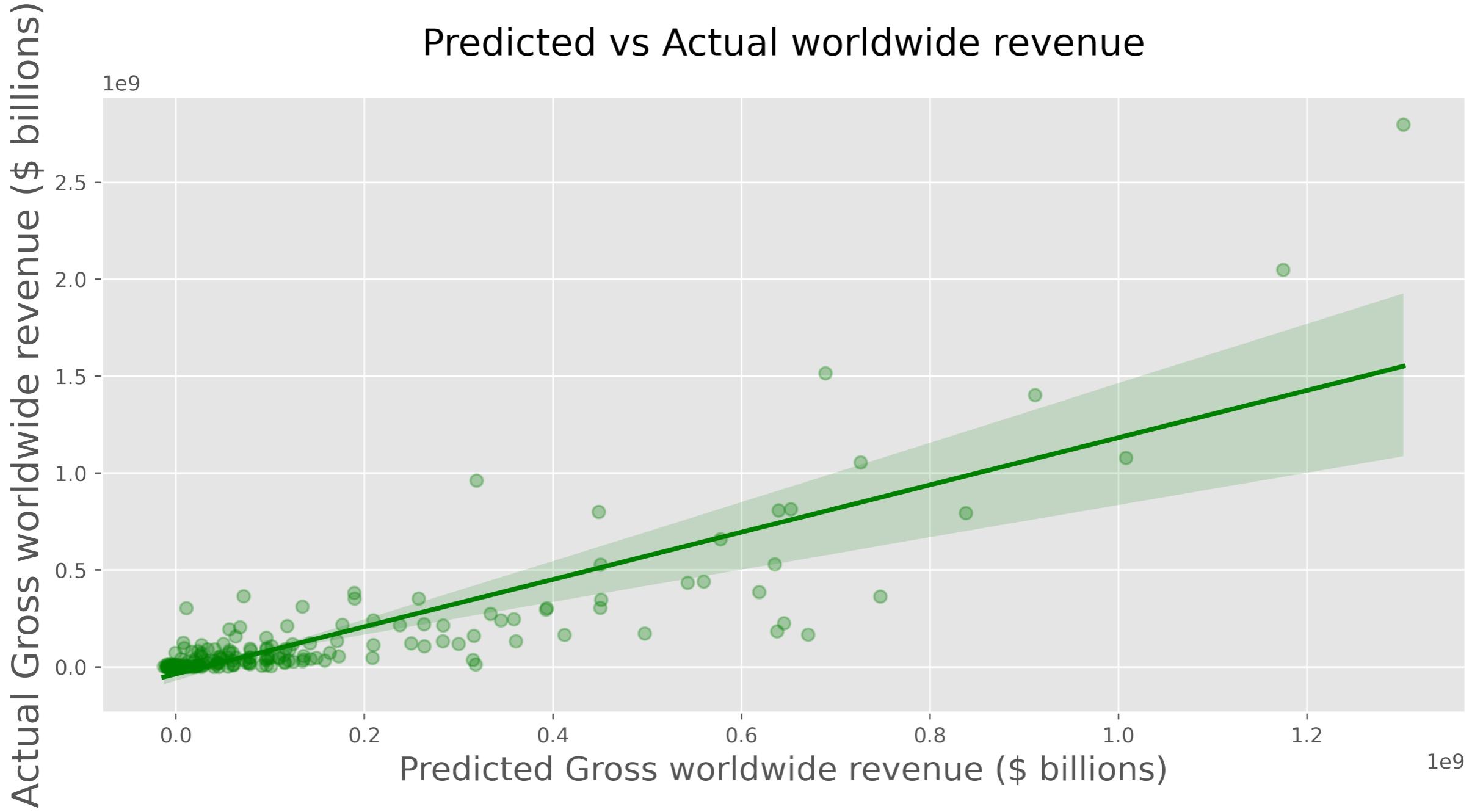
MPPA

Genre

Results

R^2 adjusted = 0.563

MAE= \$94,555M



Conclusion

-Our linear regression model poorly predicts gross worldwide revenue for movies produced in the US. We need better features to start with.



-Thank you!

Future directions

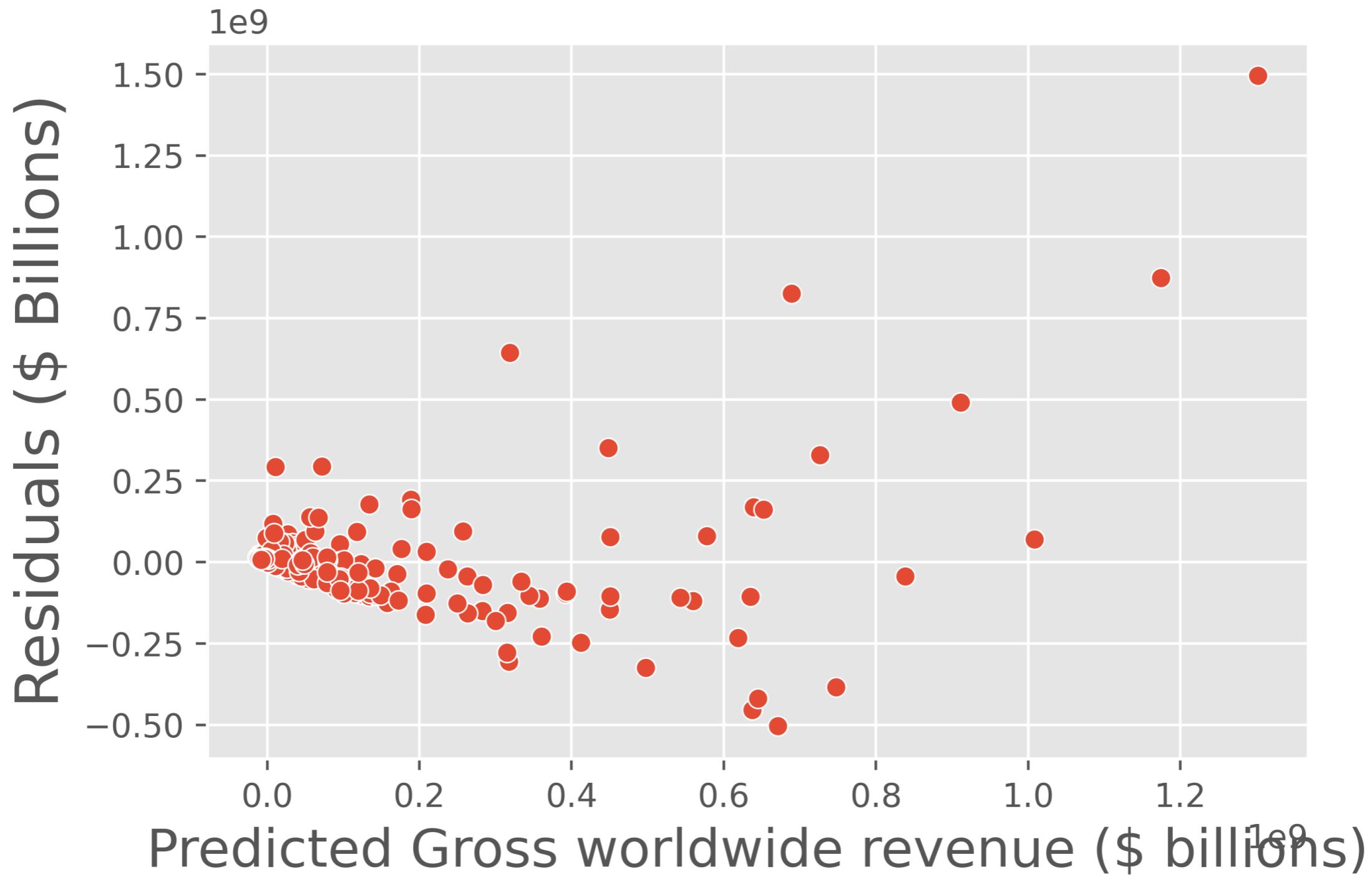
Improve features:

- NLP analyses of news articles regarding viewings prior to movie release. Did spectators stand up and clap, or not? (Sentiment analysis).
- Number of awards and nominations prior to movie release.
- Number of tweets, news after movie screening on awards.
- Investigate which movies do streaming platforms actually choose (some sort of reverse engineer their algorithms).
- WLS models to improve current model?

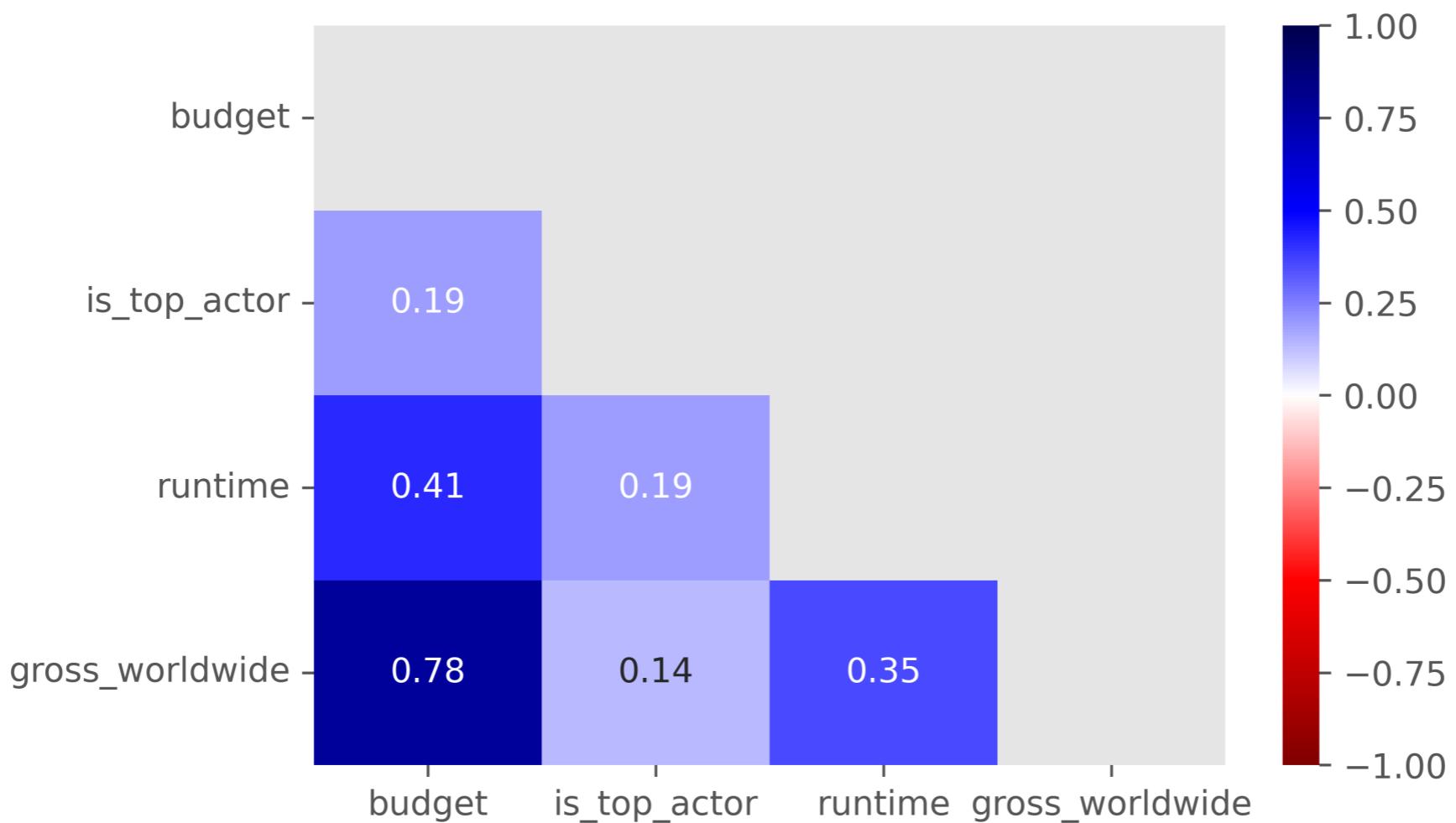
Appendix

Errors do not have constant variance

Residuals



Appendix:



Appendix:

