Movie Analysis for Microsoft

By Jinsol Cantrall

Summary

"What is a successful film?"

1. The most **profitable** - money
2.   The most **popular**- ratings

Outline

- Starting a movie business
- The Numbers, iMDB (titles), and iMDB(ratings)
- ELT(Extract Transform Load) and visualization
- Results
- Conclusions

Business Problem

Microsoft stakeholders have noticed their competitors creating original video content. In response, Microsoft has decided to create a new movie studio, but they are unfamiliar with the film sector. I have been given several datasets, and are asked to put together a report on the movie industry that will allow Microsoft producers to make informed decisions.

Data

1. **The Numbers**
   - Dataset contains budgets and gross

2. **iMDB (titles)**
   - Dataset contains titles, years, running time, and genres

3. **iMDB (ratings)**
   - Dataset contains average rating and number of votes

Methods

1. Used ELT(Extract Transform Load)
   Import data ➜ apply functions ➜ group data ➜ check for duplicates and remove
   ➜ re-check for duplicates ➜ create pivot tables ➜ Deal with missing data.
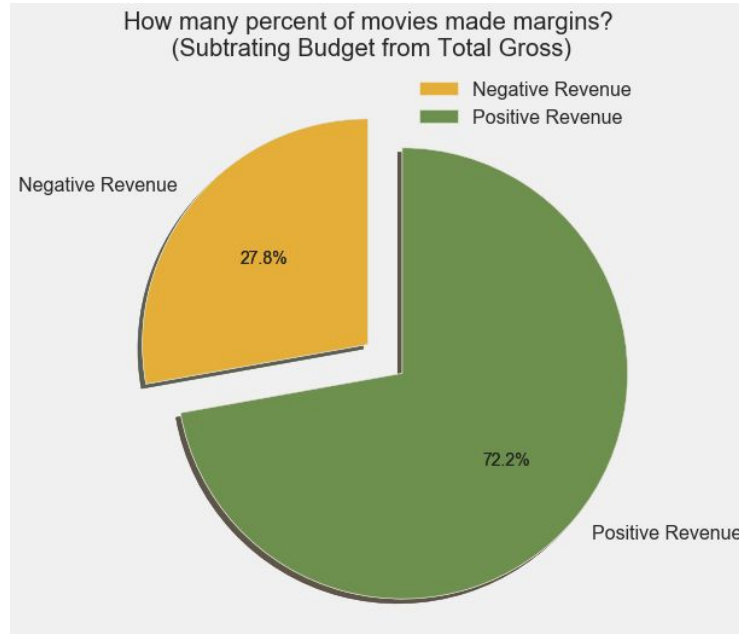
2. Visualizing Data
   Use functions to make meaningful graphs. (ex. Matplotlib and seaborn)
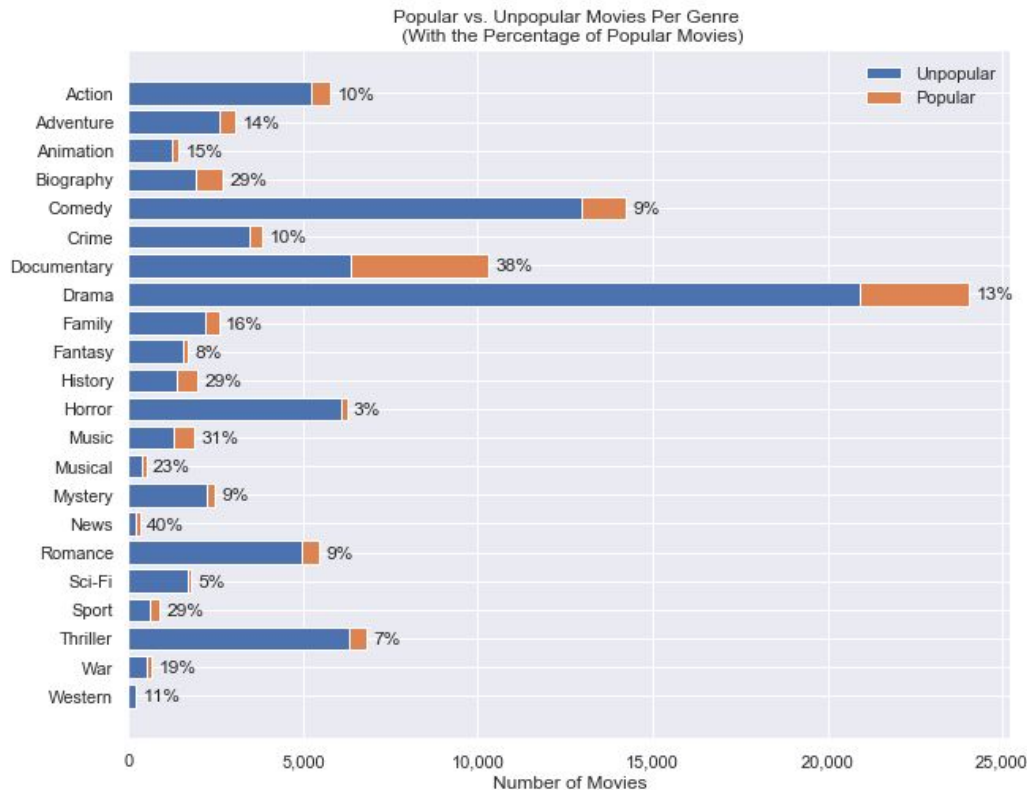
3. Infer and interpret results

Results

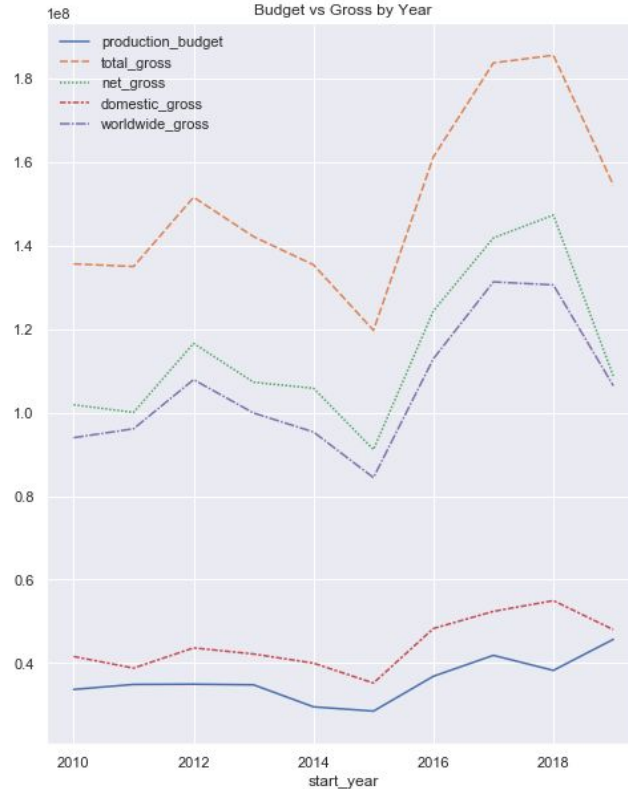Answering questions with visualizations
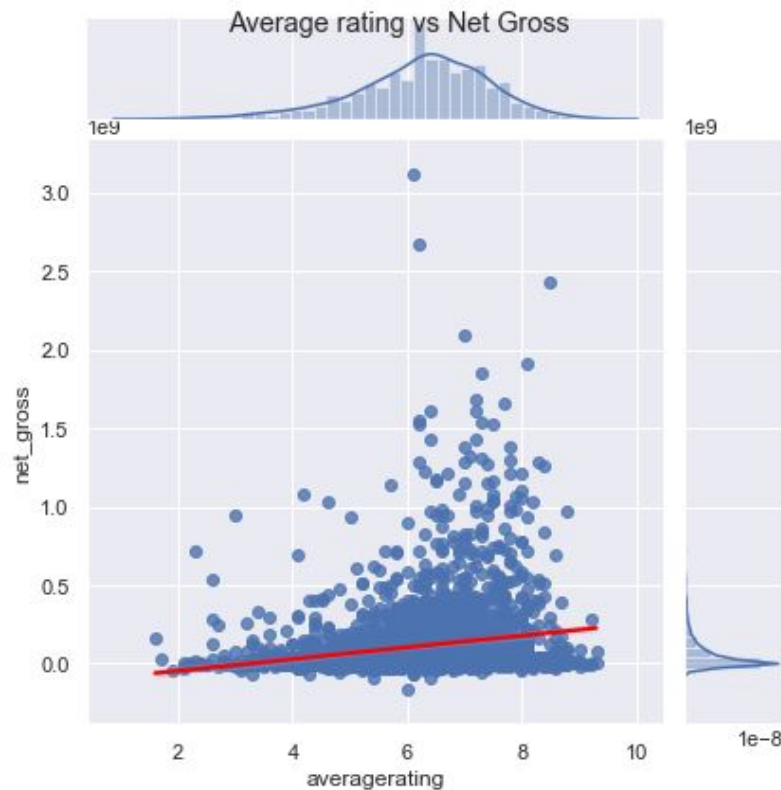
Question 1: Is it profitable to create film?

How many percent of movies made margins?
(Subtrating Budget from Total Gross)

Negative Revenue
Positive Revenue

Negative Revenue
27.8%

72.2%

Positive Revenue

Question 2: Which genres of movie is popular?

Popular vs. Unpopular Movies Per Genre
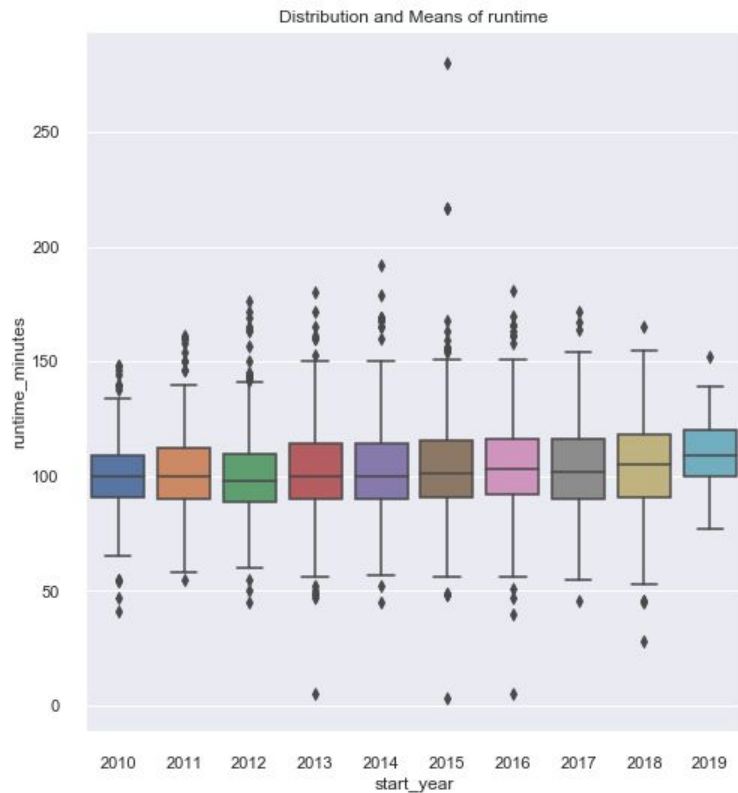(With the Percentage of Popular Movies)

Question 3: Does the large budget movie brings more money?

Question 4: Running time could be a huge factor of popularity of the movie?

Average rating vs Net Gross

Note: Average running time has been increased over decades

Distribution and Means of runtime

Conclusions

Starting movie business seems not bad. As you can see, most movies(72.2%) made more than the budget. Highest net gross movies are not always the most highest average rate nor most reviewed. As far as genres, Comedy Drama seems like safest bet since both genres are popular genre. Comedies, Documentaries and Dramas have a lot of popular movies, in terms of pure numbers But News, Documentaries and Music have a lot of popular movies, in terms of percentages.

Thank You!

Email: jinsol.cantrall@gmail.com
GitHub: @jinsolnim