Phase One Final Project
Microsoft Film Project

Author: Nahum Odemba
Institution: Moringa School
Instructor: Antony Muiko
Date: 27-August-2022

BUSINESS UNDERSTANDING

Microsoft Company is planning to join the fun-filled movie industry.

Unfortunately, Microsoft has no knowledge about movie production, leave alone where to start.

My objective of this projects was to conduct data analysis using the provided data sets to recommend to Microsoft Company on how to approach the movie industry.

Understanding The Problem

The movie industry is booming. This multi-billion-dollar industry is getting even bigger every year.
Investing in the movie industry sounds fun and also a good business opportunity for a large company like Microsoft.
However, joining the industry without proper knowledge may lead to huge losses to the company.
Therefore, this project focuses on the best approaches.
To come up with the best recommendations, I based my analysis on the following factors:

- Movie genre vs ratings and number of votes: Which is the best movie genre to invest in depending on ratings and votes?
- Movie budget vs movie income (gross): Does high budget translate to high income?
- Movie income vs genre: Does the movie genre determine its income?

Technologies Used

- GitHub
- Jupyter Notebook
- Pandas Library
- Matplotlib
- Seaborn
- Sqlite3

Data Understanding

The tn.movie_budgets.csv: Provides data on production budget of a movie and the domestic and foreign gross of that movie.

The im.db is a relational database with tables containing information such as movie directors.

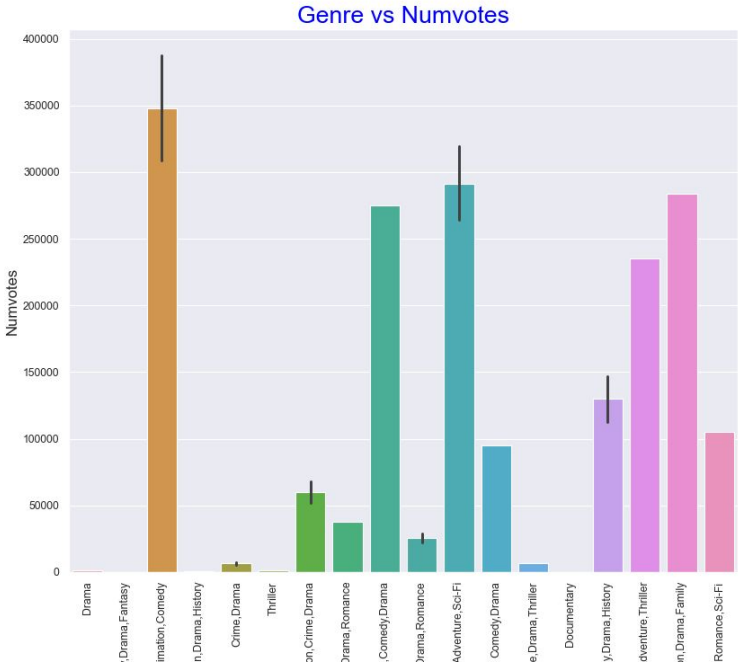
Using different technologies such as Pandas library, Jupyter Notebook, and Sqlite3 these files were read and prepared accordingly for analysis.
Data was then cleaned through dropping rows with NaN values and duplicates.

Data Analysis

Which movie genres have higher ratings with high number of votes:

I plotted a bar graph of genres vs. numvotes, for movies with ratings above 6 and numvotes above 100. The graph above shows that the following movie genres have high ratings and number of votes:

1. Adventure, Animation, Comedy
2. Action, Adventure, Sci-Fi
3. Action, Drama, Family



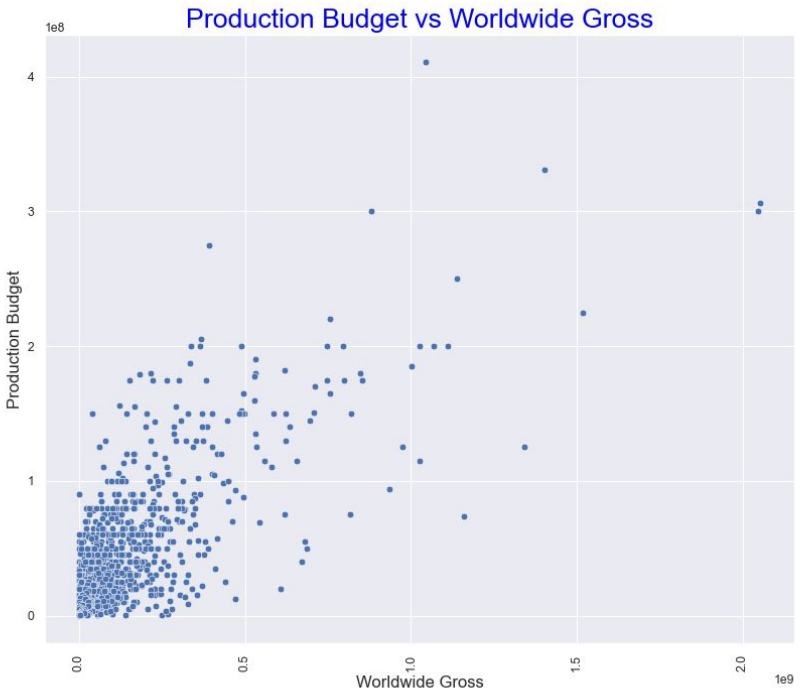
Data Analysis

Does high production cost translates to high movie income?

The graph clearly shows that correlation between production budget and worldwide gross.

There are movies with high production budget that do not translate to high worldwide gross.

Also, there are movies with low production budget but high worldwide gross.

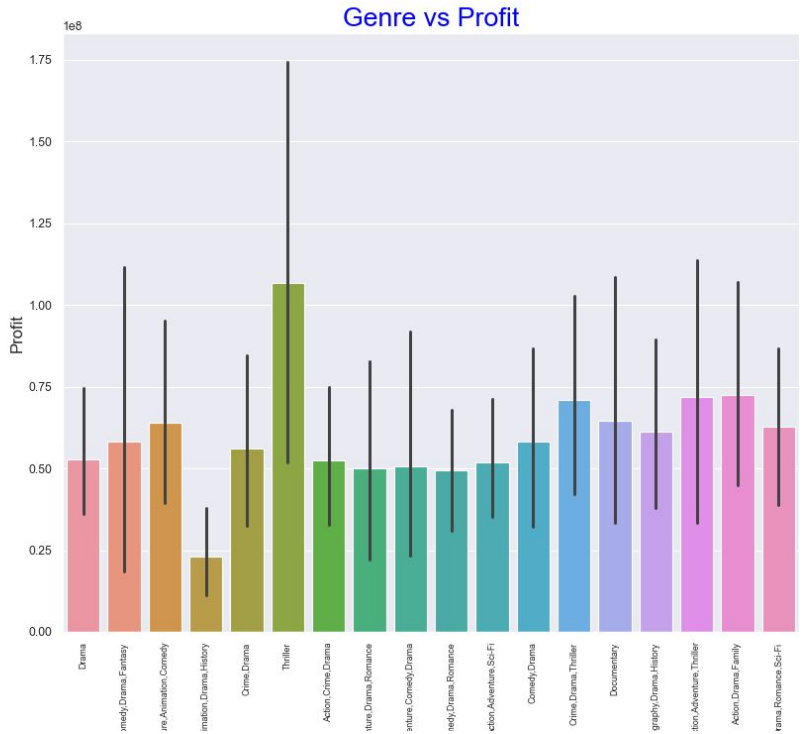


Data Analysis

Does movie income depends with its genre?

The chart shows that the following genre make the highest profits:

1. Thriller
2. Action, Drama, Family
3. Action, Adventure, Thriller
4. Crime, Drama, Thriller



Conclusion

- This data analysis proved wrong certain hypothesis about movie industry.
- For instance, it is hypothesised that high production budget translates to high movie income.
- From the analysis, it is proven that high production cost does not guarantee high income.
- Second, this analysis finds high movie rating does not translate into high income.
- For instance, Adventure, Animation, Comedy genre has a very high rating but low income compared to Thriller which has very low rating but high income. Action, Adventure, Sci-Fi is another example of genres with high ratings but average income.
- This conclusion, however, may be affected by the number of movies in said genres.

Recommendations

- ★ First, using the graphs that I have presented, I would recommend that Microsoft's new movie studio kicks off with Thriller movies.
- ★ Thrillers have the potential of making more money for the studio and the company.
- ★ This recommendation has been arrived because Thrillers have high profits from the analysis that I did with the available data.
- ★ Second, the production budget should be kept at the possible minimum because high production cost does not translate into high income. However, this should not be done at the expense of the movie quality.