



MICROSOFT MOVIE STUDIO PROJECT

MARTIN NGUGI

BACKGROUND

Microsoft is venturing into the film industry by establishing a new movie studio. With limited knowledge about creating movies, Microsoft aims to gain insights into the types of films that are currently succeeding at the box office. By analyzing the prevailing trends and success factors in the film industry, Microsoft aims to make informed decisions that will guide their film creation process. This exploration is crucial for Microsoft to thrive in the highly competitive market by aligning their strategies with the evolving landscape of the film industry.



DATA EXPLORATION

- Six datasets were used to arrive to conclusions of this exploratory data analysis;
1. The 'bom.movie_gross.csv' dataset is a box office dataset showing the domestic and foreign gross revenues of each movie.
 2. The 'rt.movie_info.tsv.gz' dataset shows the Runtime of the various movies that has been produced before.
 3. The 'tn.movie_budgets.csv' dataset shows production budgets and their corresponding worldwide gross revenue.
 4. The 'tmdb.movies.csv' dataset shows the original languages of different movies and their percentage popularity.
 5. The 'imdb.title.basics.csv.gz' and 'imdb.title.ratings.csv.gz' datasets were joined together so as to help in in finding the highest and lowest rated genres.

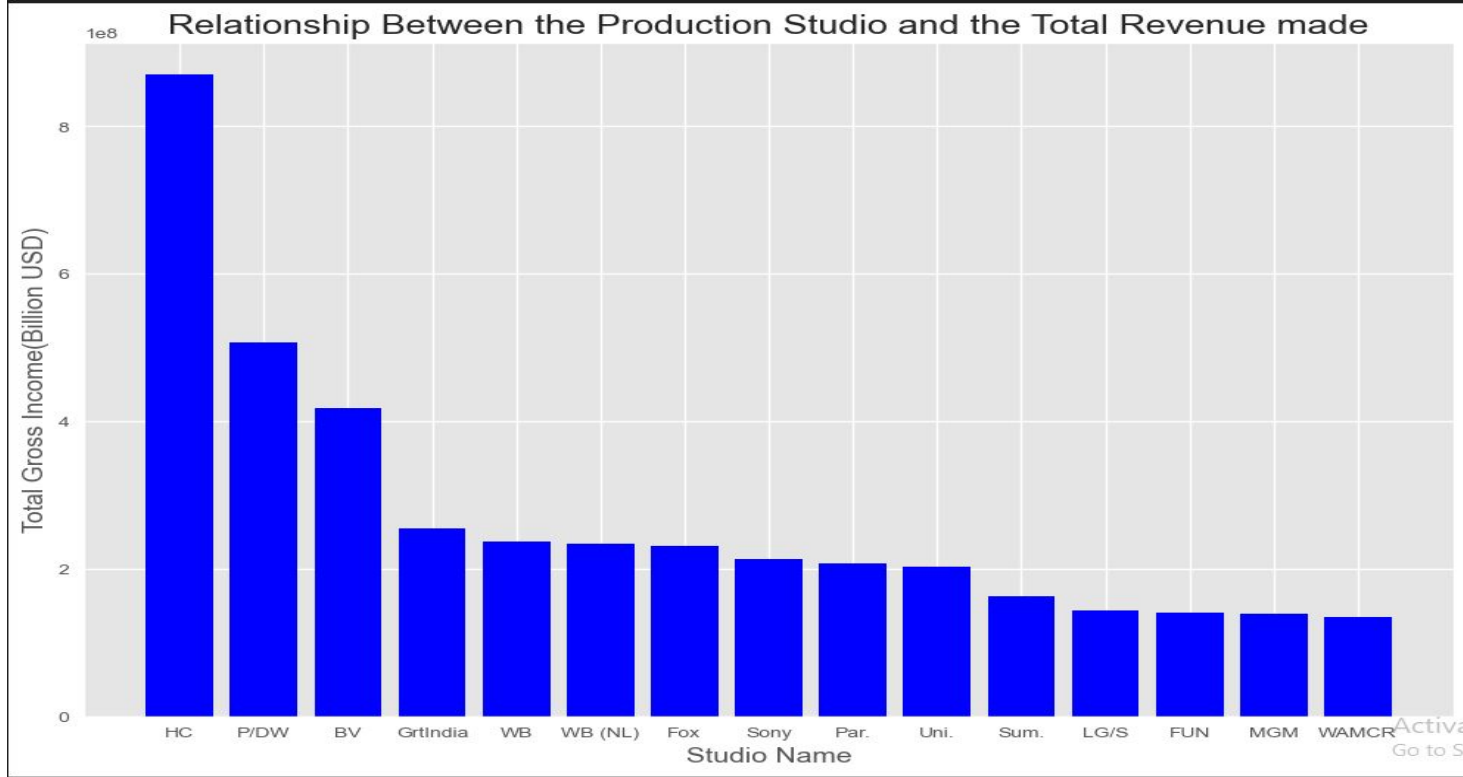


OBJECTIVES

1. Which studio produces the highest grossing movies.
2. Find the average runtime of movies and what would be the recommended runtime for movies to be produced in the studio.
3. What is the relationship between the movie budget and the gross income generated?
4. which is the most preferred original language of a movie.
5. Find the genres with the highest and lowest ratings.



RELATIONSHIP BETWEEN THE PRODUCTION STUDIOS AND THEIR TOTAL REVENUE MADE.



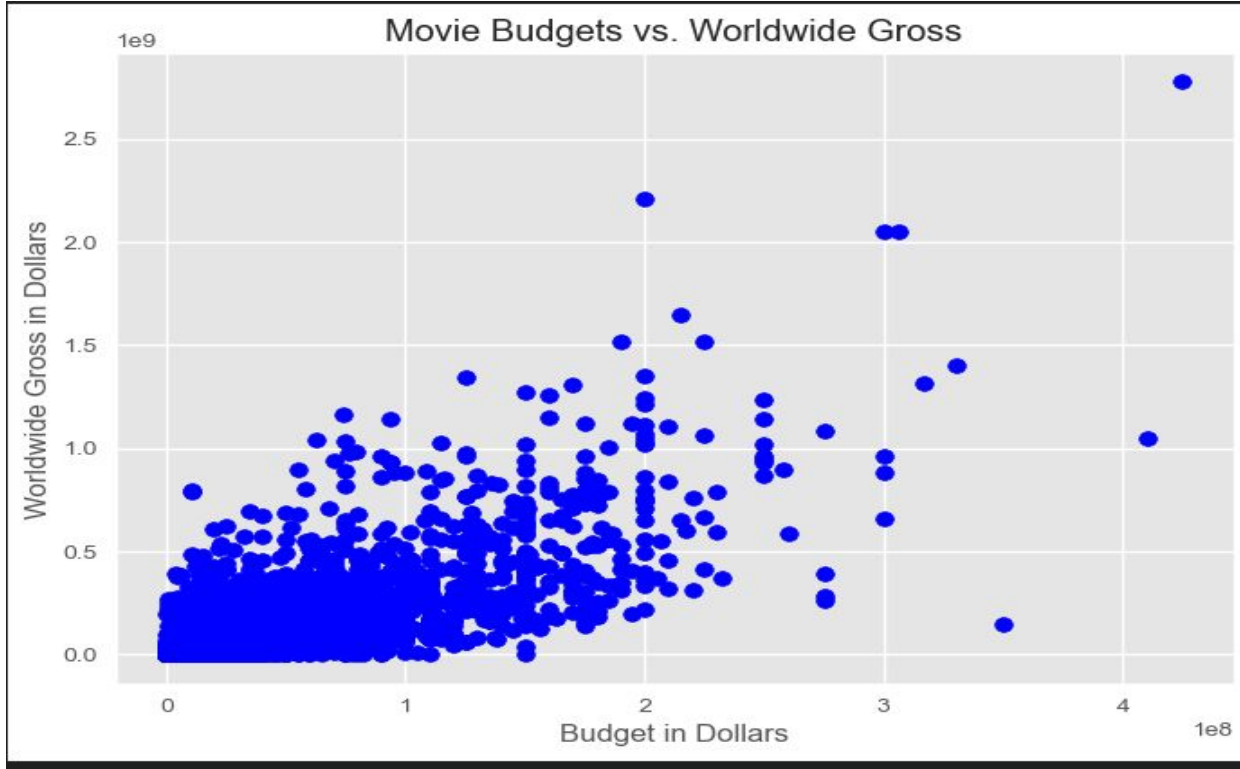
- The Bar plot shows the top 15 earning studios
- The top 5 earning studios are; HC, P/DW, BV, GrtIndia and WB.

AVERAGE MOVIE RUNTIME

```
count    1560.000000
mean      101.968590
std       28.276622
min        0.000000
25%       90.000000
50%      100.000000
75%      114.000000
max      358.000000
Name: runtime, dtype: float64
```

- From the above image it is recommended that the average runtime of a movie should be between 100 minutes(mean) and 120 minutes(upper percentile).

COMPARISON BETWEEN MOVIE PRODUCTION BUDGET AND ITS NET GROSS INCOME

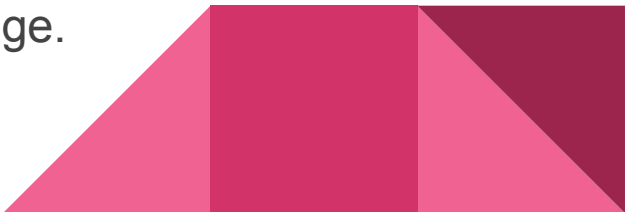


The Scatter plot reveals a moderately positive correlation between the gross income and the budget of a movie, suggesting that an increase in budget tends to result in a higher gross income. This positive relationship implies that by utilizing the insights from our analysis and assembling the appropriate resources, we can potentially gain a competitive edge in our investments and achieve profitability.

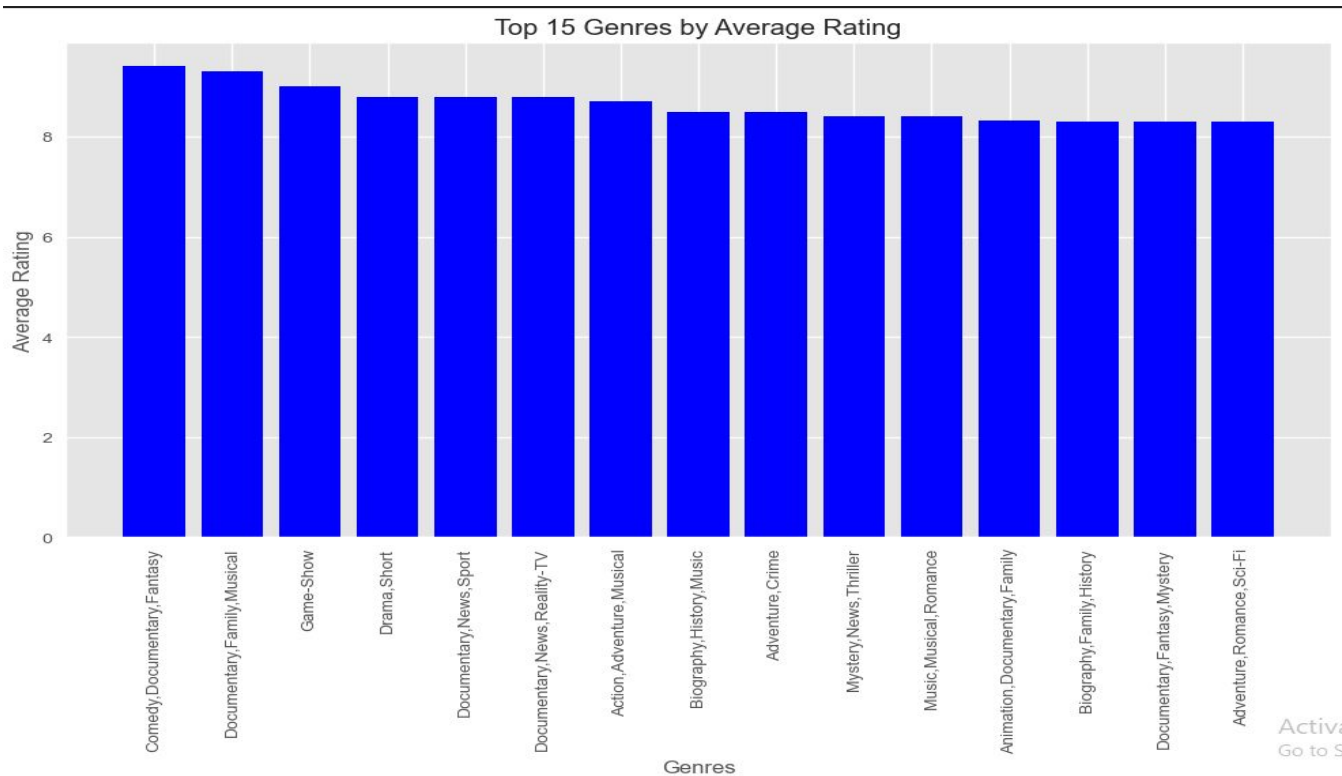
MOVIE LANGUAGE PREFERENCE

```
en    23291
fr      507
es      455
ru      298
ja      265
...
bo         1
si         1
sl         1
hz         1
dz         1
Name: original_language, Length: 76, dtype: int64
```

It can be seen that most movies released, their original language was english. This is also so since English is the most preferred universal language.

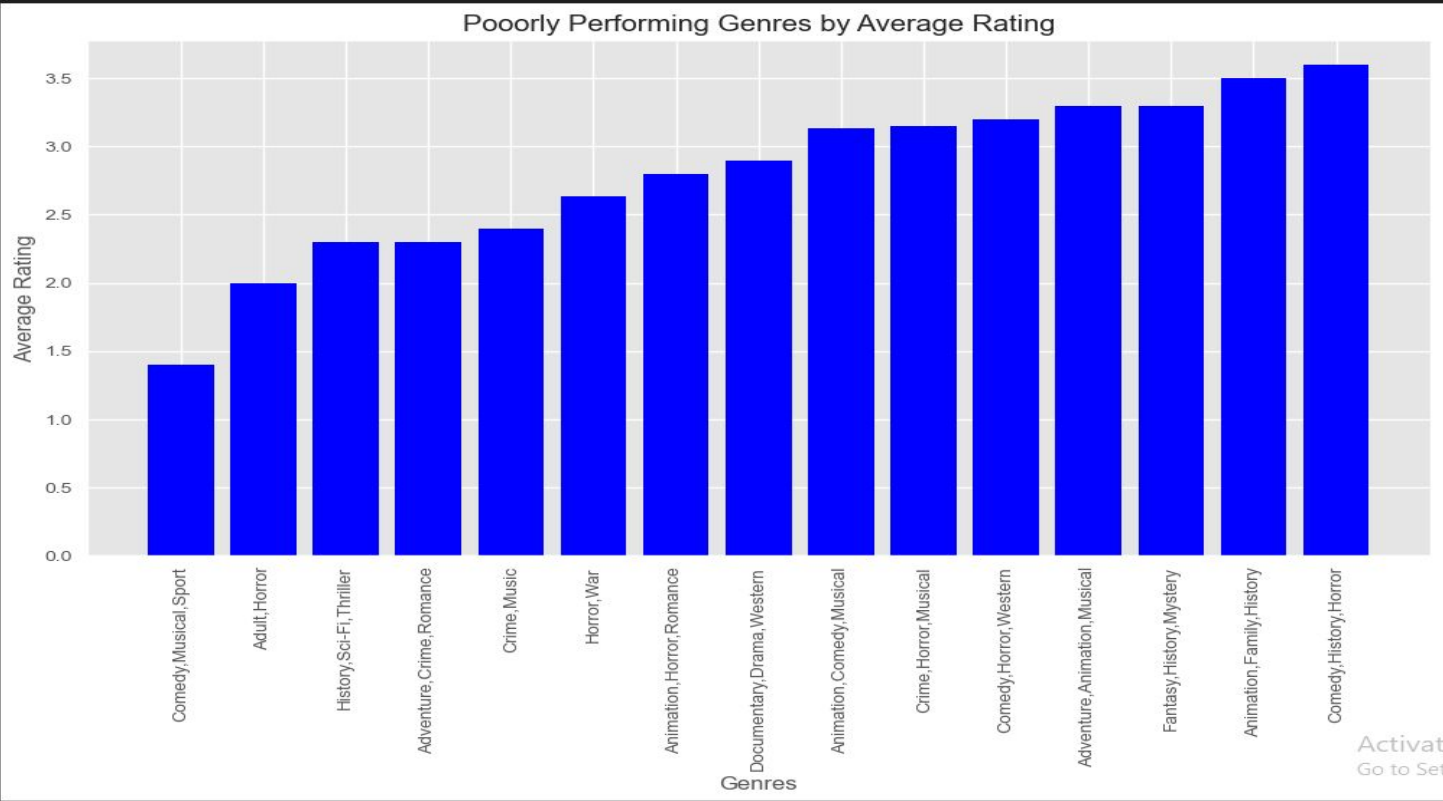


GENRES WITH THE HIGHEST RATINGS



The above bar plot represents the different types or genres of films and their corresponding ratings. We see that the most highly rated genres include Science Fiction & Fantasy, Action & Adventure, Kids & family, Comedy, Documentaries among other highly rated genres.


GENRES WITH THE LOWEST RATINGS




The bar plot represents the poorly performing genres by rating. We see that the poorly rated genres include Horror, History, Romance, War among other poorly rated genres.

CONCLUSION

After analyzing the data obtained from various sources such as the TMDb dataset, BOM movies, reviews, and movie budgets, several insights were identified to guide Microsoft's decision-making process for creating successful films. The analysis showed that the most successful genres by ratings were action, adventure, comedy, science fiction and Documentaries. The analysis also highlighted the importance of having a high production budget for successful films. These findings provide actionable insights to guide Microsoft's new movie studio in creating successful films and competing in the highly competitive movie industry.



RECOMMENDATIONS

1. They should strive and work closely or partner with the highest grossing companies in the industry as they could borrow on some of their planning and execution of projects for better return on business.
 2. The average run time of a movie being produced should be averagely 100 minutes to 120 minutes.
 3. The new studio should invest heavily on their production budget of a movie and with the right resources, this will help in coming up with better movies hence having high prospects on its returns to the business.
 4. The new studio should focus more in producing movies with their original language being English.
 5. The new studio should produce genres that attract good ratings such as Science Fiction & Fantasy, Action & Adventure, Kids & family, Comedy, Documentaries as it will lead to more viewership which translates to more returns on investment.
- 

THANK YOU

~MARTIN NGUGI

DATA SCIENTIST

EMAIL: nwainaina76@gmail.com

GITTHUB: <https://github.com/martin-ngugi>

