

PHASE 1 PROJECT

Done by Flavine Atieno
Data Science Class(dsf-pt04)

Part Time

23/07/2023

INTRODUCTION

Movies Production Data Analysis

- This project's purpose is to provide valuable insights for Microsoft's new studio venture.
- The task is to explore the types of movies that are performing well at the box office and transform and translate these findings into actionable recommendations for Microsoft.

Business Problems

Microsoft needs to:

1. Establish and understand the relationship between ratings and revenue of different genres.
2. Assess the trend in movie release years.
3. Identify profitable movies genres and studios that made them.
4. Explore International market opportunities of movies.

Data Sources

Data sources used for the analysis

1.Box Office Mojo

➤ `bom.movie_gross(csv file)`

2.IMDb

➤ `imdb.title.basics(csv file)`

➤ `imdb.title.ratings(csv file)`

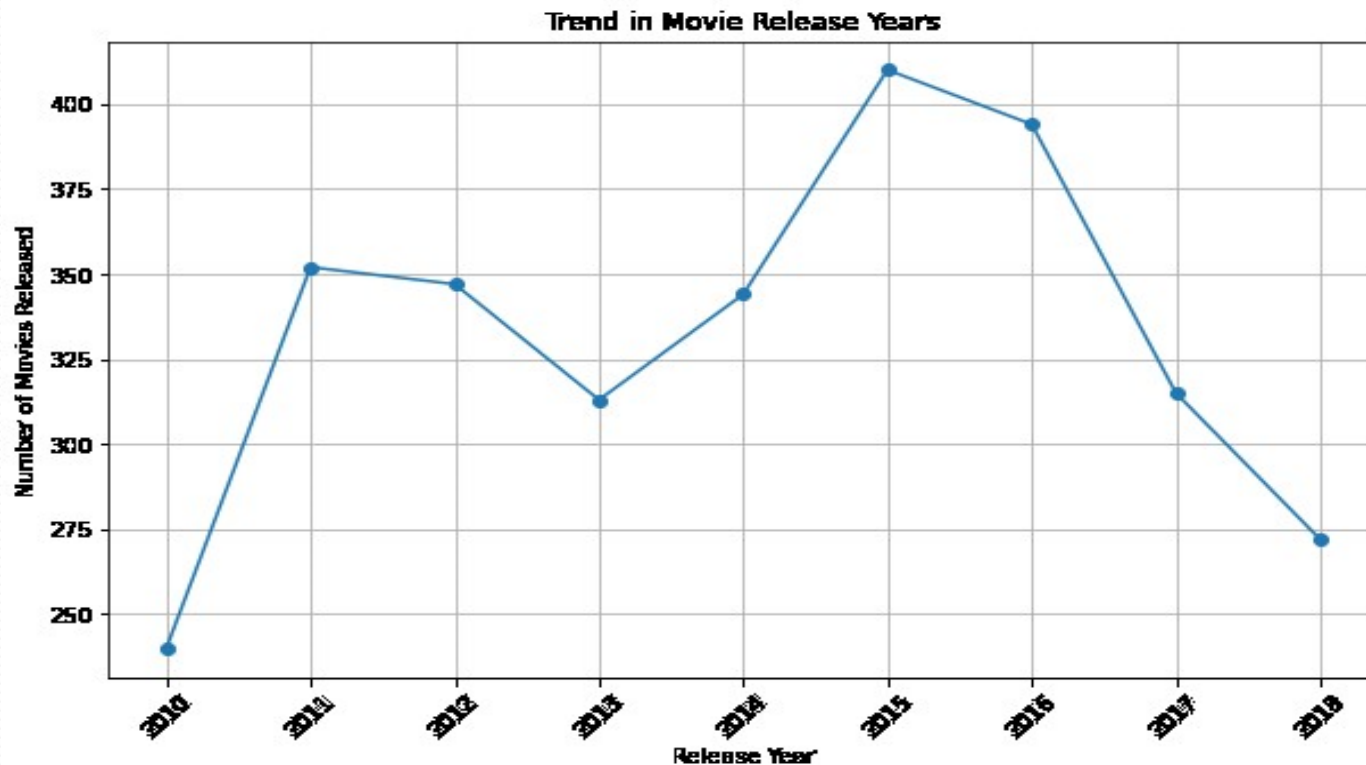
Data Cleaning and Exploration

- The data cleaning process, including:
 - Handling missing values
 - Removing duplicates
 - Removing white spaces and special characters
 - Data type conversions
- Something I noticed from the datasets is that they are of different lengths, when creating relational databases it returns a relatively small data frame to work with.

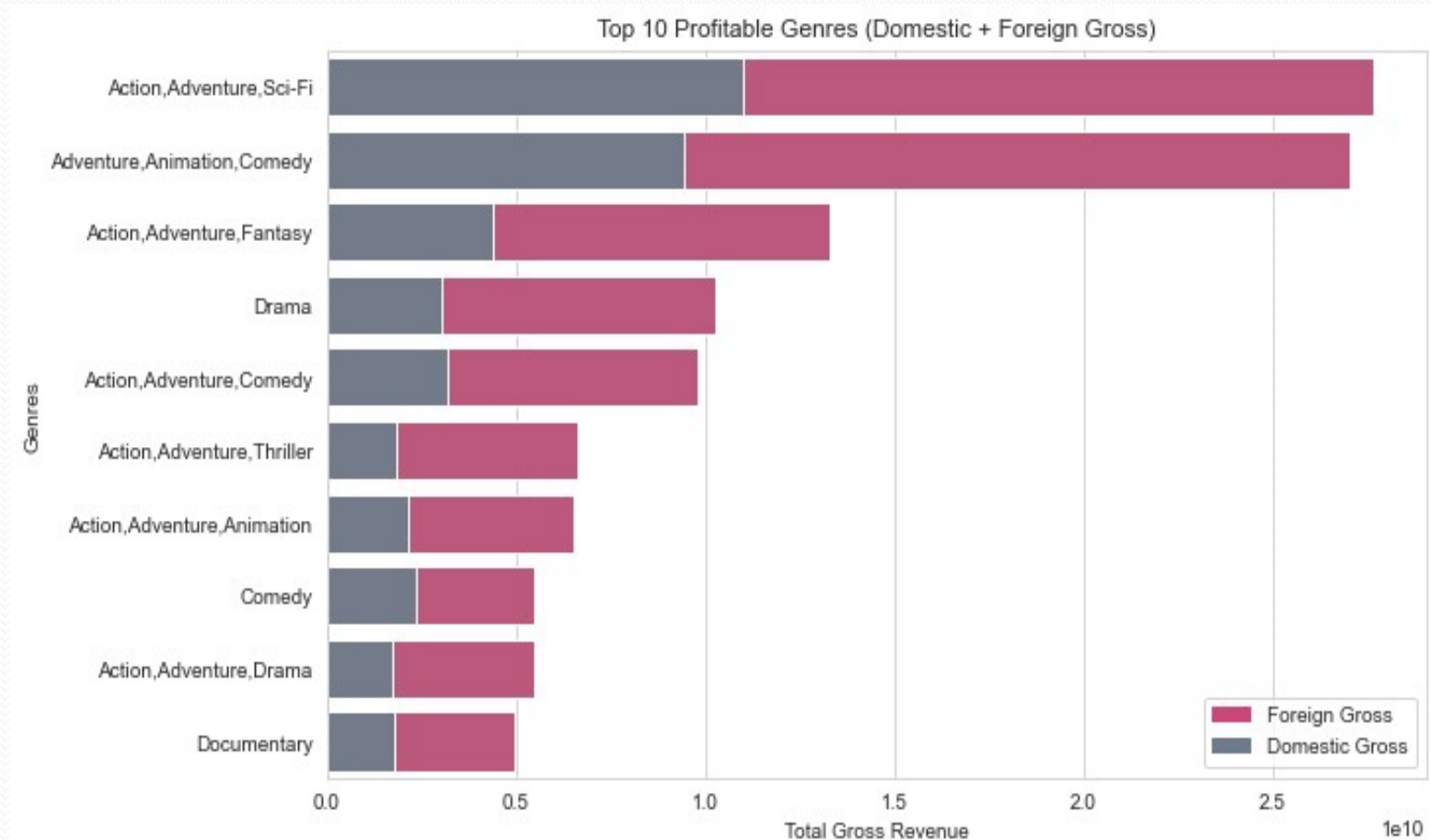
Relationship between ratings and revenue of different genres.

- The correlation-coefficient of the between ratings and domestic gross is 0.12 which indicates that the two have no relationship, ratings does not influence the revenue generated from the movie genres.

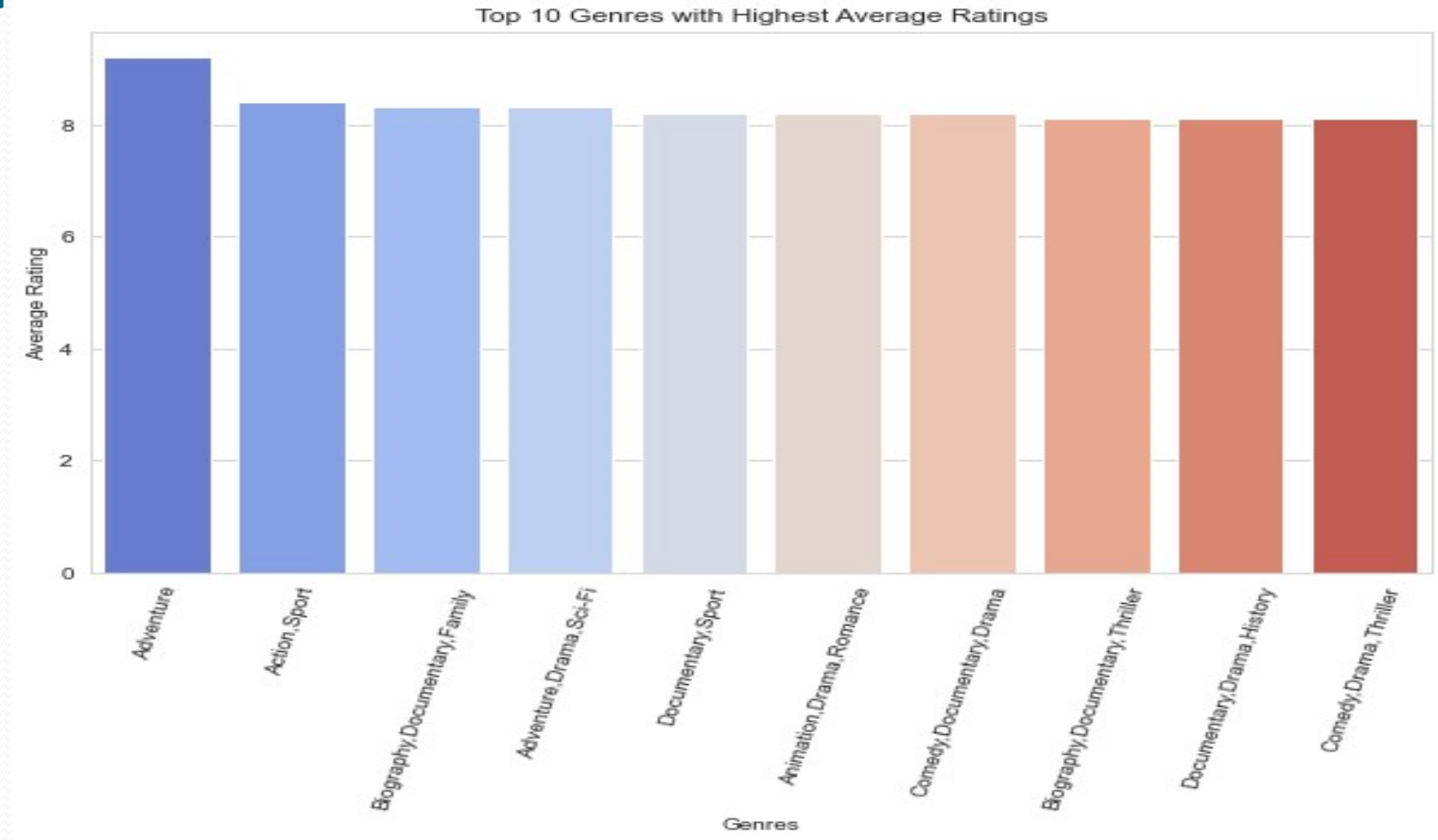
Assessing the trend in movie release years



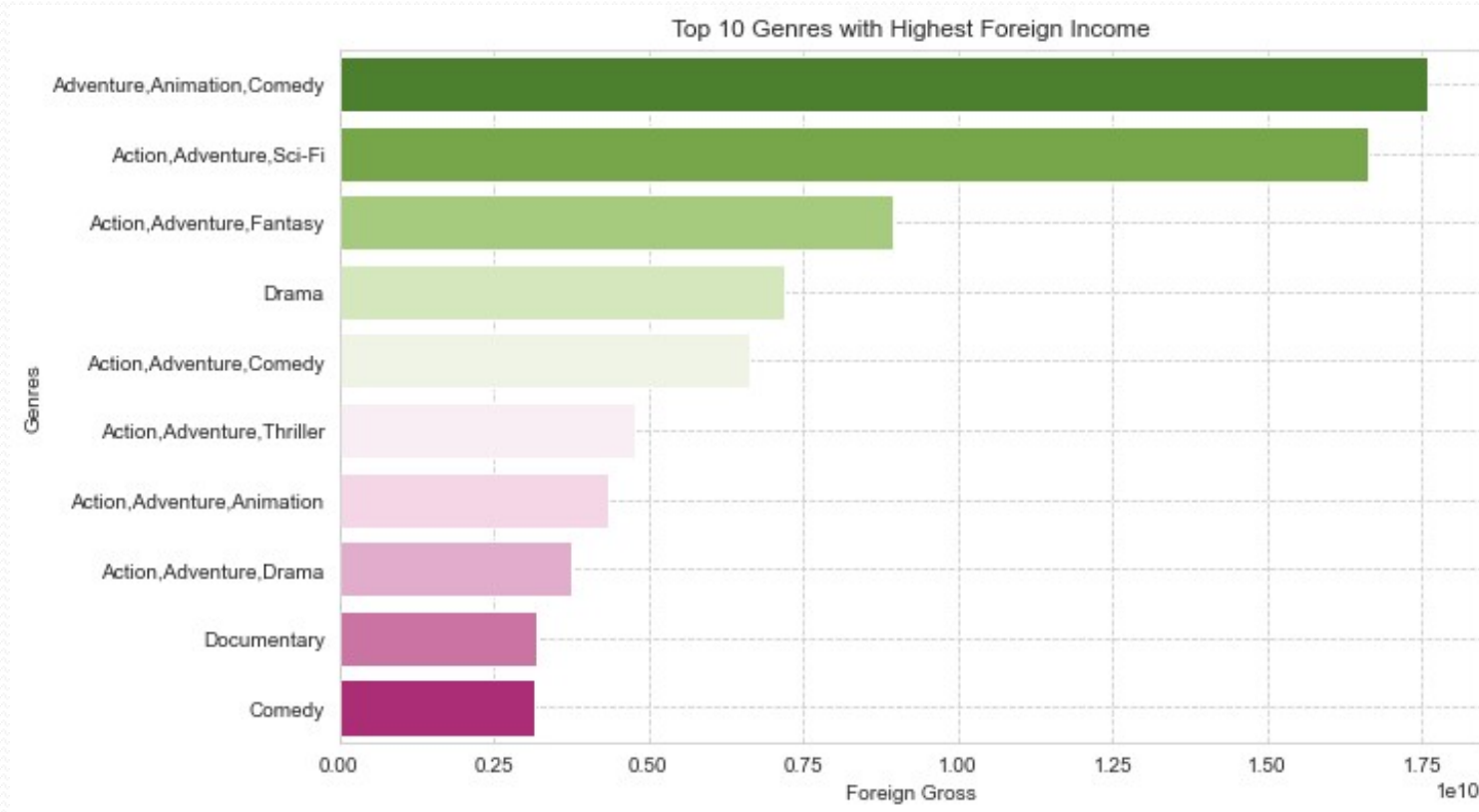
Profitable Movie Genres



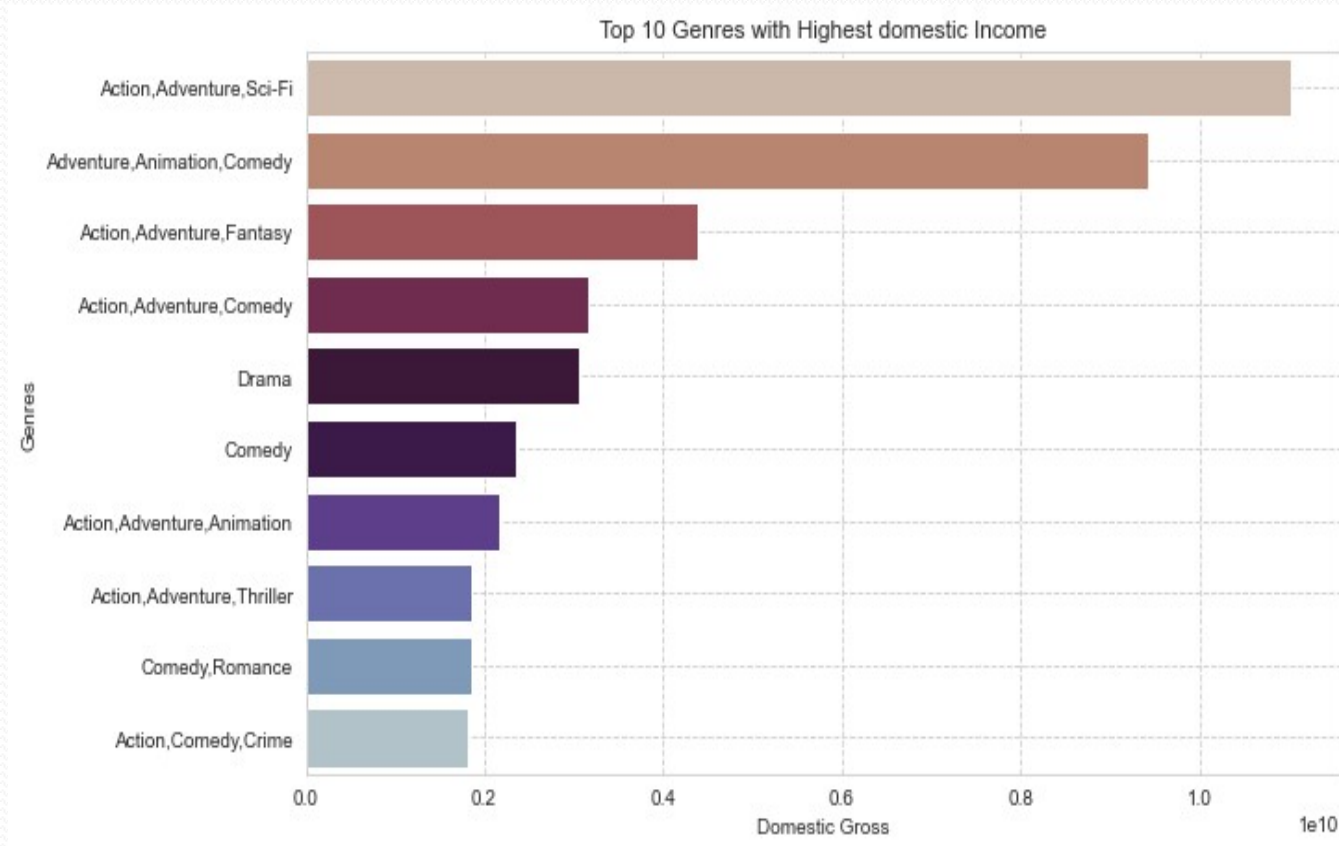
Top 10 Genres with Highest Average Ratings



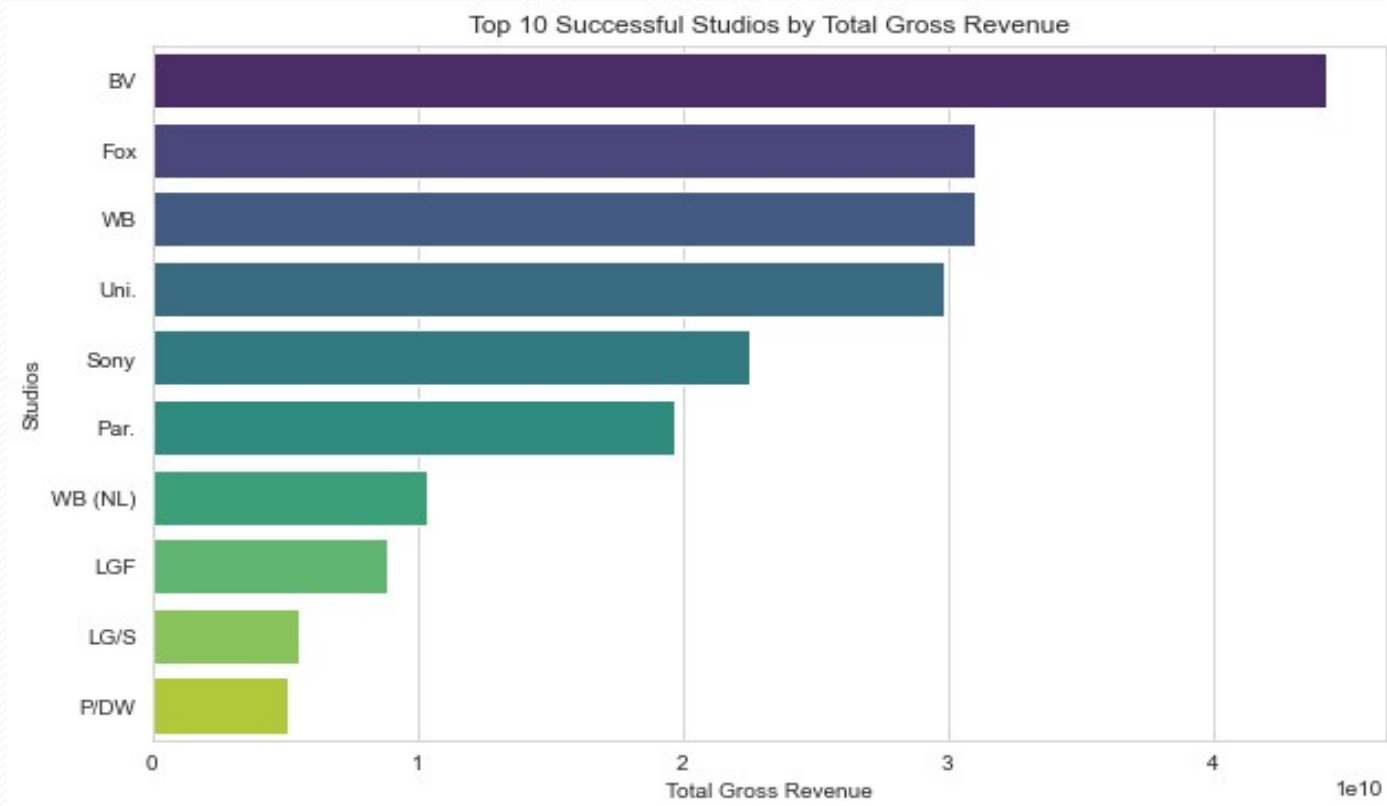
Top 10 Genres with Highest Foreign Income



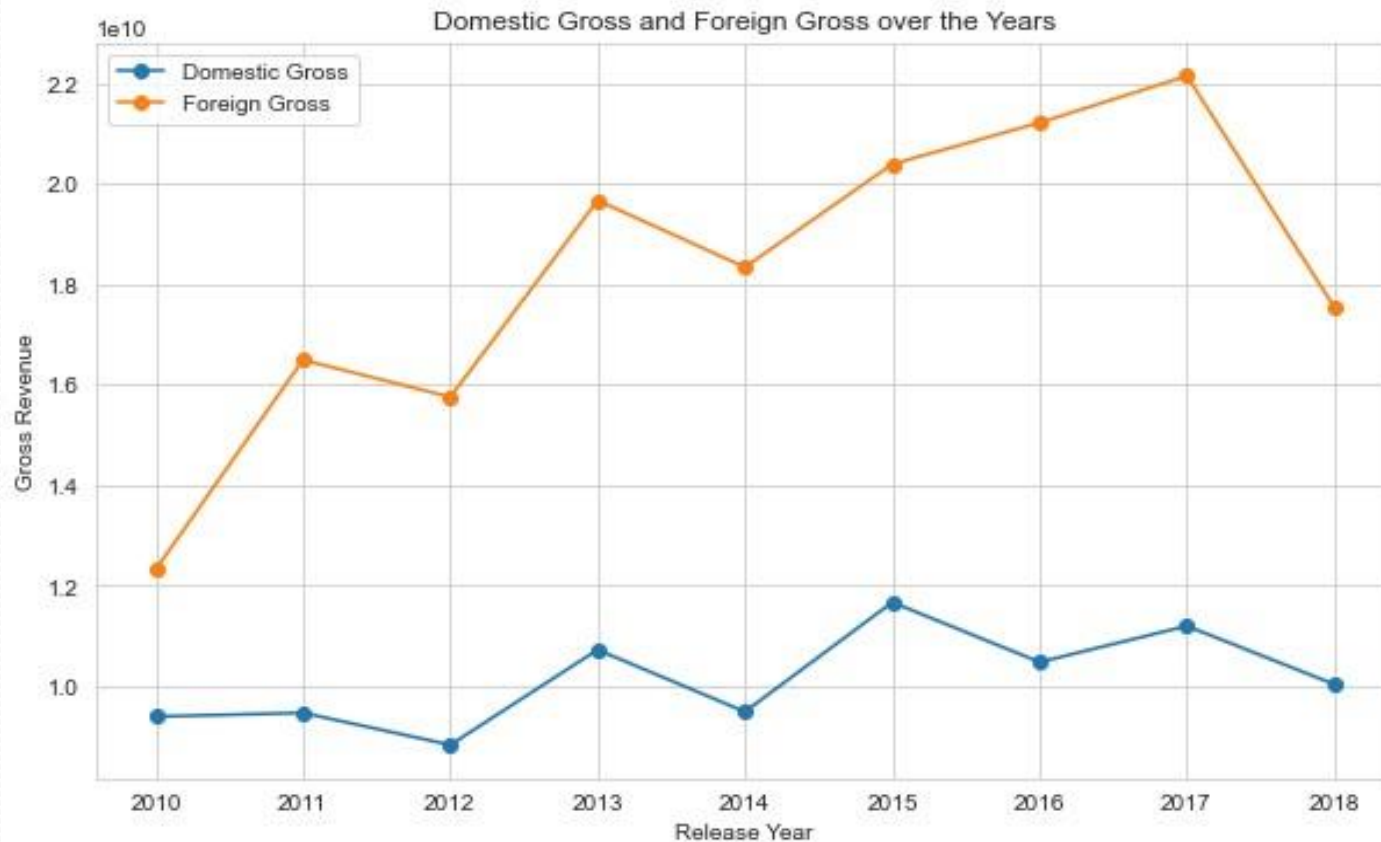
Top 10 Genres with Highest domestic Income



Successful Studios by Total Gross Revenue



Exploring International market opportunities of movies.



Recommendations

- Microsoft should;
 1. Focus on popular genres which are highly rated and highly voted for to increase chances of success.
 2. Explore partnership with international markets since foreign revenue is higher than the domestic revenue.
 3. Focus on high revenue generating studios to learn from them such as Buena Vista, Fox and Warner Bros.
 4. Study the trend in movies release-years.

Conclusion

- In conclusion all I can say is that the analysis of the three datasets combined gives us an insight of the key features that play a role in the success of movie production, its not just one aspect but many including rating, release year, popularity, the target audience whether domestic or international and the revenues they make.



FIN

THE END

THANK YOU