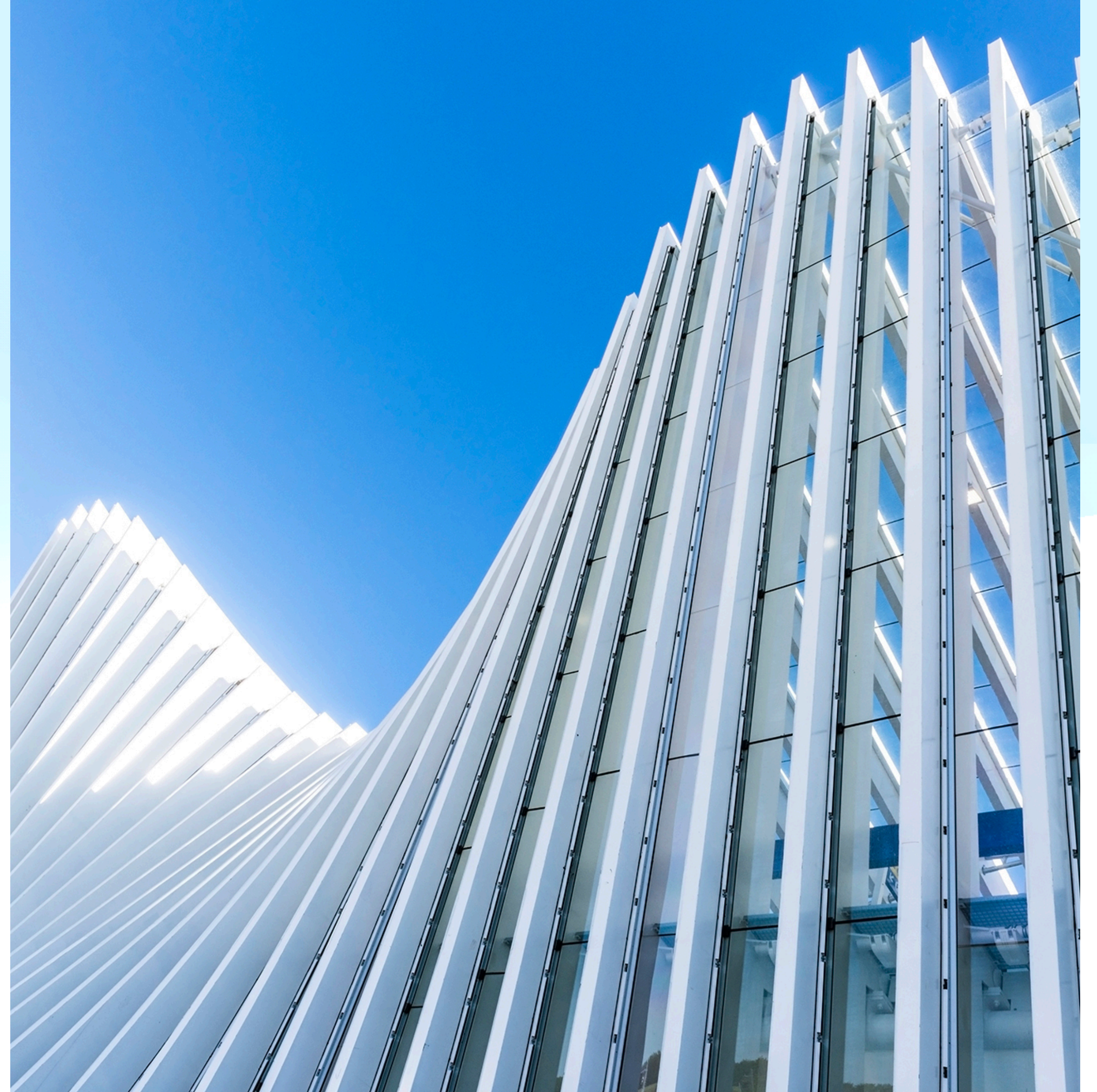


PHASE_1 PROJECT

**TITLE:
EXPLORATORY DATA ANALYSIS
FOR MICROSOFT'S UPCOMING
FILM STUDIO.**



INTRODUCTION

Project overview

The aim for carrying out this project was to analyze data from companies such as Box Office Mojo, IMDB, Rotten Tomatoes and MovieDB, since they are excelling in the film industry.

Main objective of the study:

To identify the genres of film or films that are currently doing well in the market.

The Study was significant since it fulfilled the objectives and saw drama as a genre was doing well in the market together with The monkey 3 Movie was doing well worldwide.

I would therefore recommend we start by producing Drama movies.

1.BUSINESS UNDERSTANDING

Why the research?

Microsoft recognizes the need to educate itself on current landscape of box office performance and this is crucial for any company that wants to generate income from its operations.

2.0 DATA MINING

2.1 Source of data

The project used data obtained from other seasoned firms,
That is:

Box office Mojo

Rotten Tomatoes

IMDB SQL Database

2.2 Data understanding

- 1 IMDB SQL DATABASE

- It is a database made up of 8 tables, Each with different set of Columns and Rows.

- The table names are:

- movie_basics.

Persons

- Directors.

Principals

- Knownfor

Writers

- movie_akas

- Movie_ratings

- Persons

2.2 DATA UNDERSTANDING

- **2 .Box Office Mojo Data**
- Made up of rows and columns
- Had five columns as shown below;

Title	Studio	Domestic_gross	Foreign_gross	Year

2.2 DATA UNDERSTANDING

- 3.Rotten tomatoes Data

- Used the movie_info table and the reviews table

- *Movie_info columns;*

Id	Synopsis	Rating	Genre	Director	Writer	theater_dat	dvd_date	Currenc	box_offic	runtime	Studio

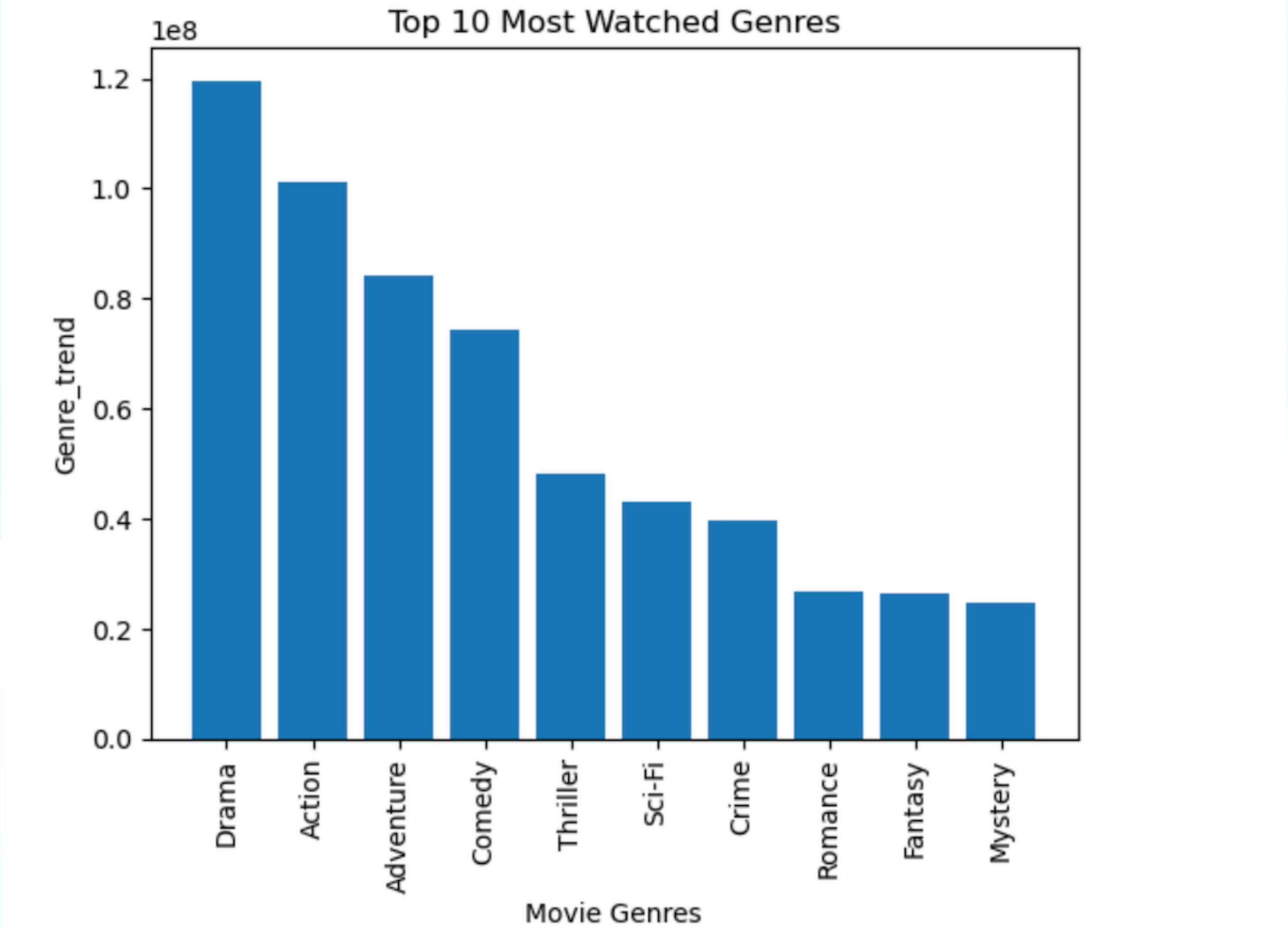
- *Reviews columns;*

Id	Review	Rating	Fresh	Critic	top_critic	Publisher	Date

- Notice, id is common in both tables, I merged the tables and did the analysis

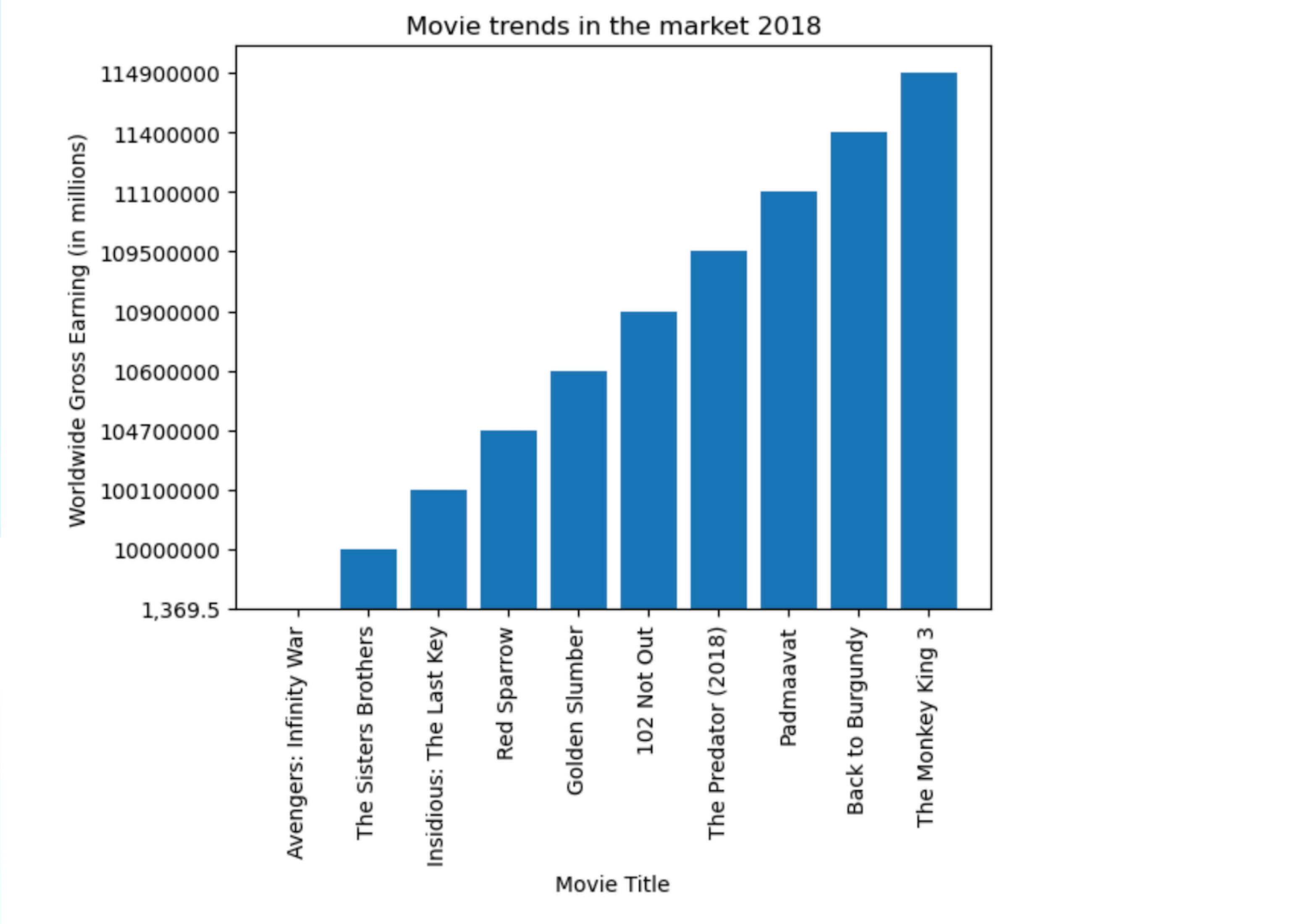
3.0 DATA ANALYSIS

fig.1 Bar plot showing Current Boxmovie trends in the market by genres



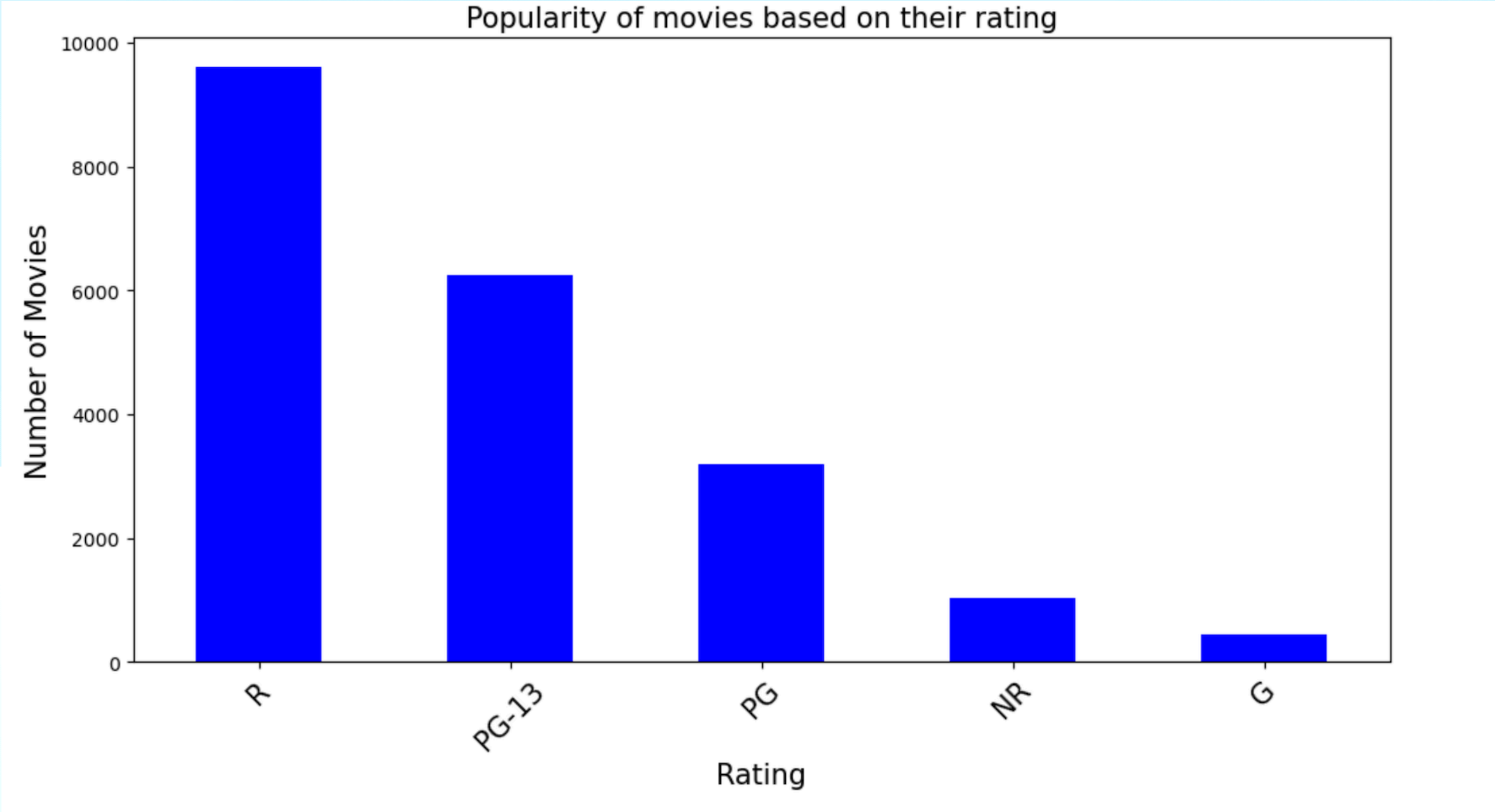
The most loved genre among the people is Drama, it has the highest number in genre_trend, Adventure comes in third. From the sample, Mystery Genre comes last and it is struggling on that spot. People showed less or no interest in the genre together with fantasy and romance genre.

Fig 2.Bar plot showing movie trends in the year 2018



The monkey King 3 movie is leading on world wide gross Earnings.It is doing great worldwide and has the highest foreign gross income of over 11.4 Millions compared to other movies.Back to Burgundy is following closely at number 2 from the sample and Avengers:Infinity has no World gross earnings.

Fig. 3 popularity of movies based on their rating



Movies rated R are the most popular followed by the ones rated PG-13 and the movies rated G are struggling in the movie market and they came last this shows that they are less popular.

4. Conclusion and recommendations

Conclusion

- The project yielded the desired results. The main objective and specific objectives were all satisfied. The results showed that the most loved genre among the people was Drama, the monkey King 3 movie was leading on world wide gross Earnings, and also Movies rated R were the most popular.

Recommendations

- I would recommend Microsoft studio to produce more of these listed genres; Drama, Action movies and Adventure since they were the most loved genres and this will attract more audience and in return more revenue to Microsoft Studios.
- I would recommend microsoft to produce movies like The monkey King 3 movie [2018] since it was doing great worldwide and had the highest foreign gross income of over 11.4 Millions, which translates to more money to the company as compared to producing the other movies
- I would recommend Microsoft studio to produce movies rated R or rated PG-13 since they were the most consumed and this in turn bring money to the company

The End !

Thank you

Name:Gitahi Muchori

Linkedin profile:<https://www.linkedin.com/in/kevin-muchori-2739432b4/>