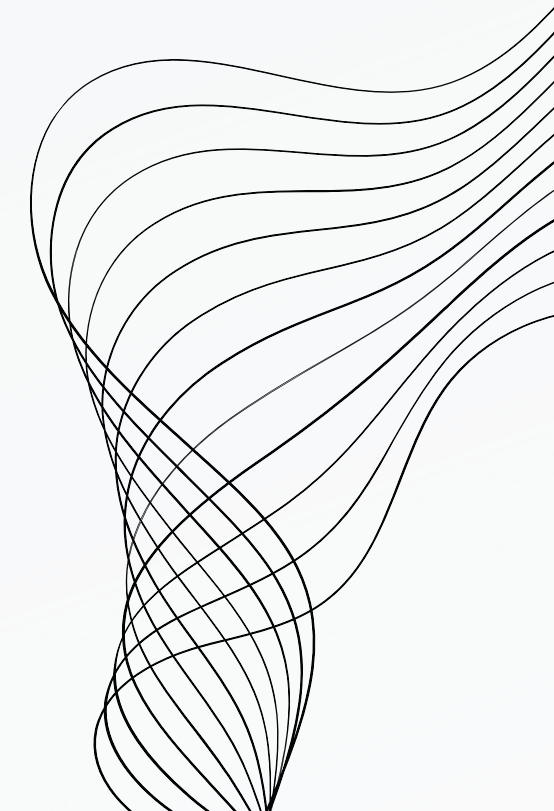


PHASE 1 PROJECT

INSIGHTS FROM MOVIE DATASET & RECOMMENDATIONS





CONTENT



01

INTRODUCTION

02

DATASET OVERVIEW

03

DESCRIPTIVE STATISTICS

04

GENRE AND RATINGS ANALYSIS

05

REVENUE AND BUDGET ANALYSIS

06

RECOMMENDATIONS

07

CONCLUSION

INTRODUCTION

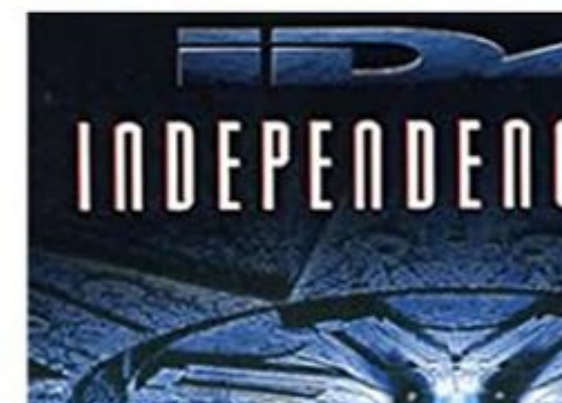


To determine the types of films that are performing well at the box office and provide actionable insights for Microsoft's new movie studio, an analysis of the current trends and successful genres would be essential



The project involves an analysis of movie datasets to try and establish correlations and relationships of elements in the dataset. The analysis is divided into two major aspects:

- Genre and ratings analysis
- Revenue and budget analysis

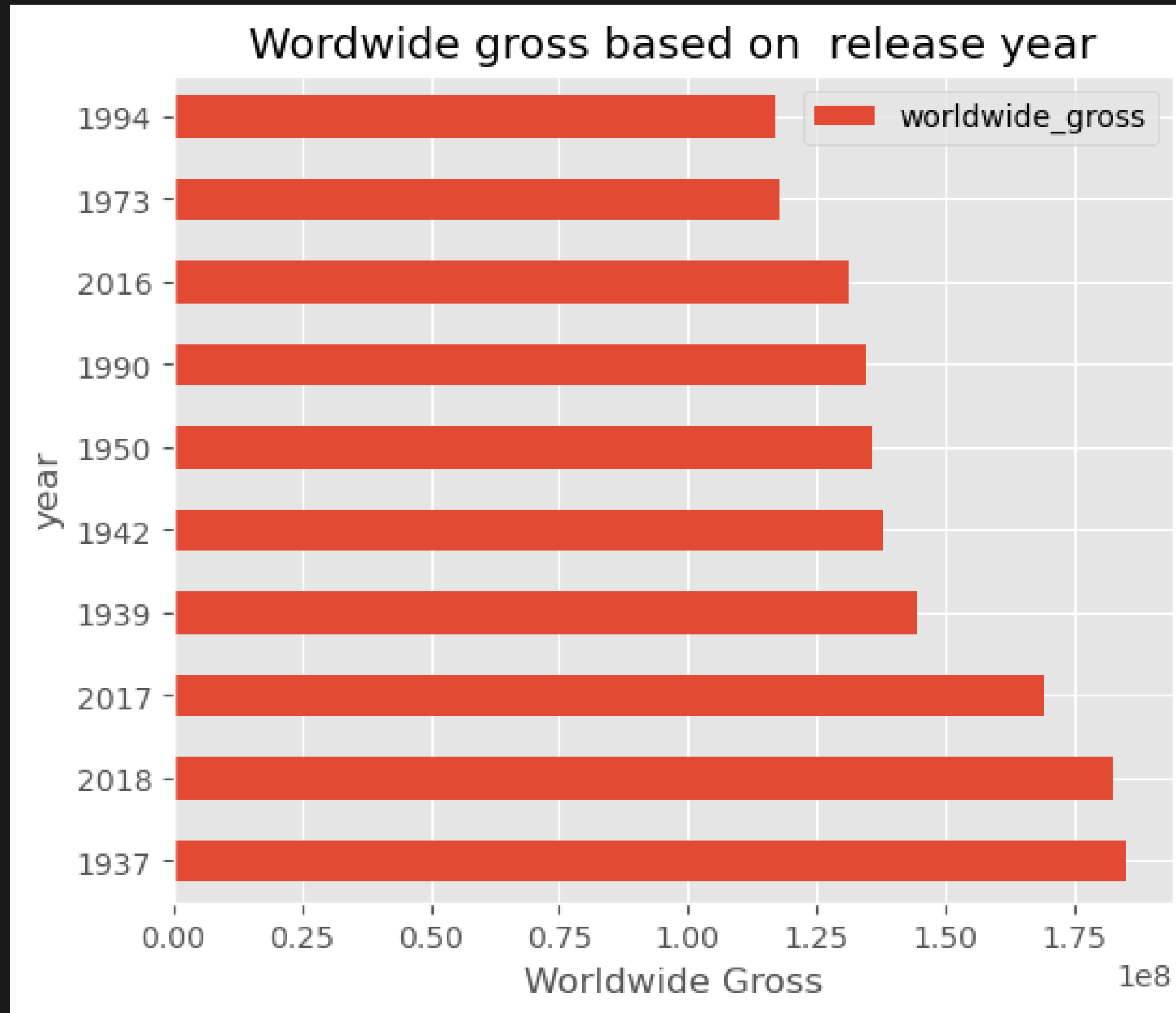


DATASET OVERVIEW

I used two datasets in my analysis:

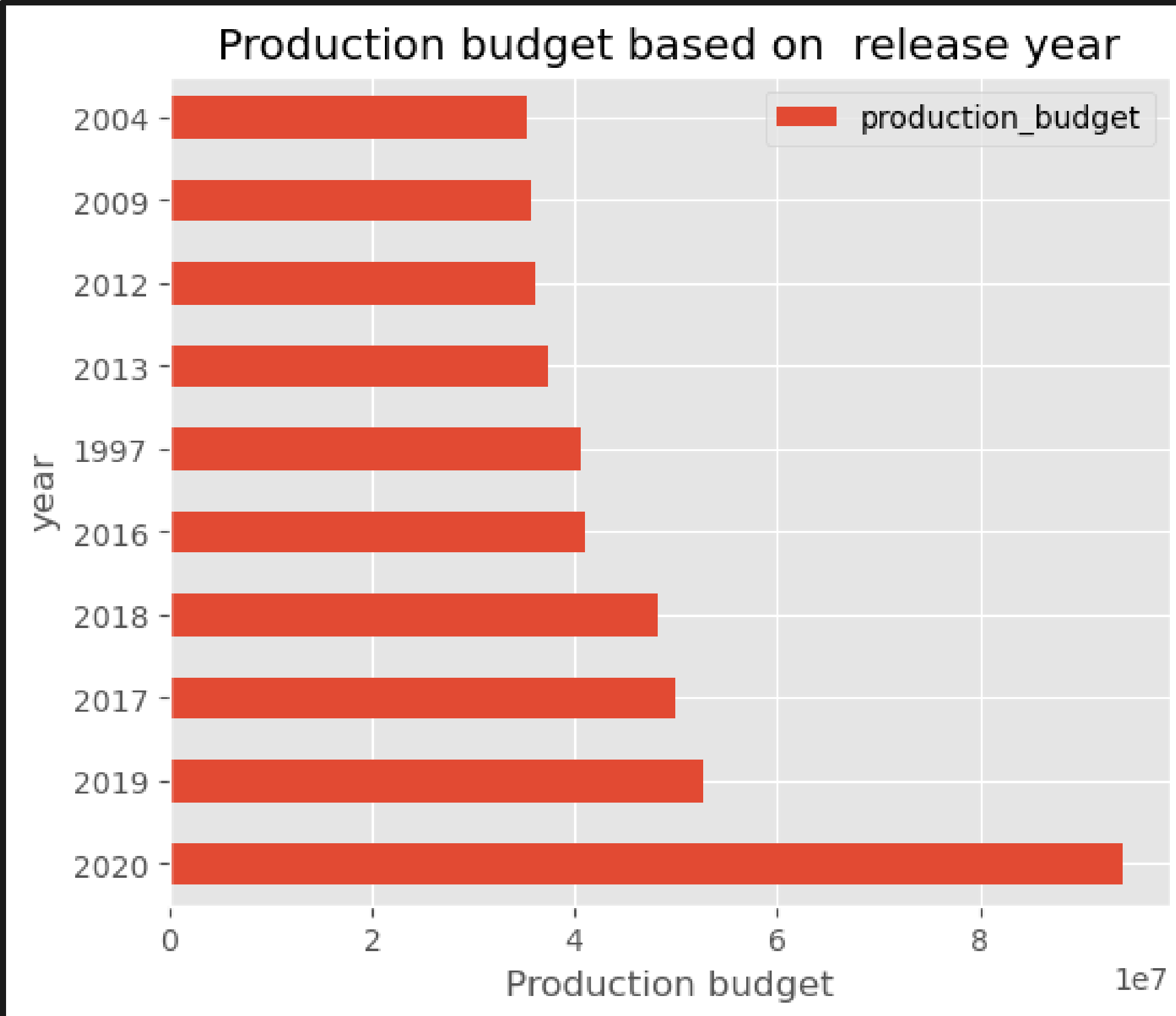
- **tn.movie_budget.csv**: The dataset contains list on movies, their release dates, production budget and domestic and worldwide gross.
- **tmdb.csv**: The dataset contains a list of movies, the popularity and vote average ratings.

DESCRIPTIVE STATISTICS



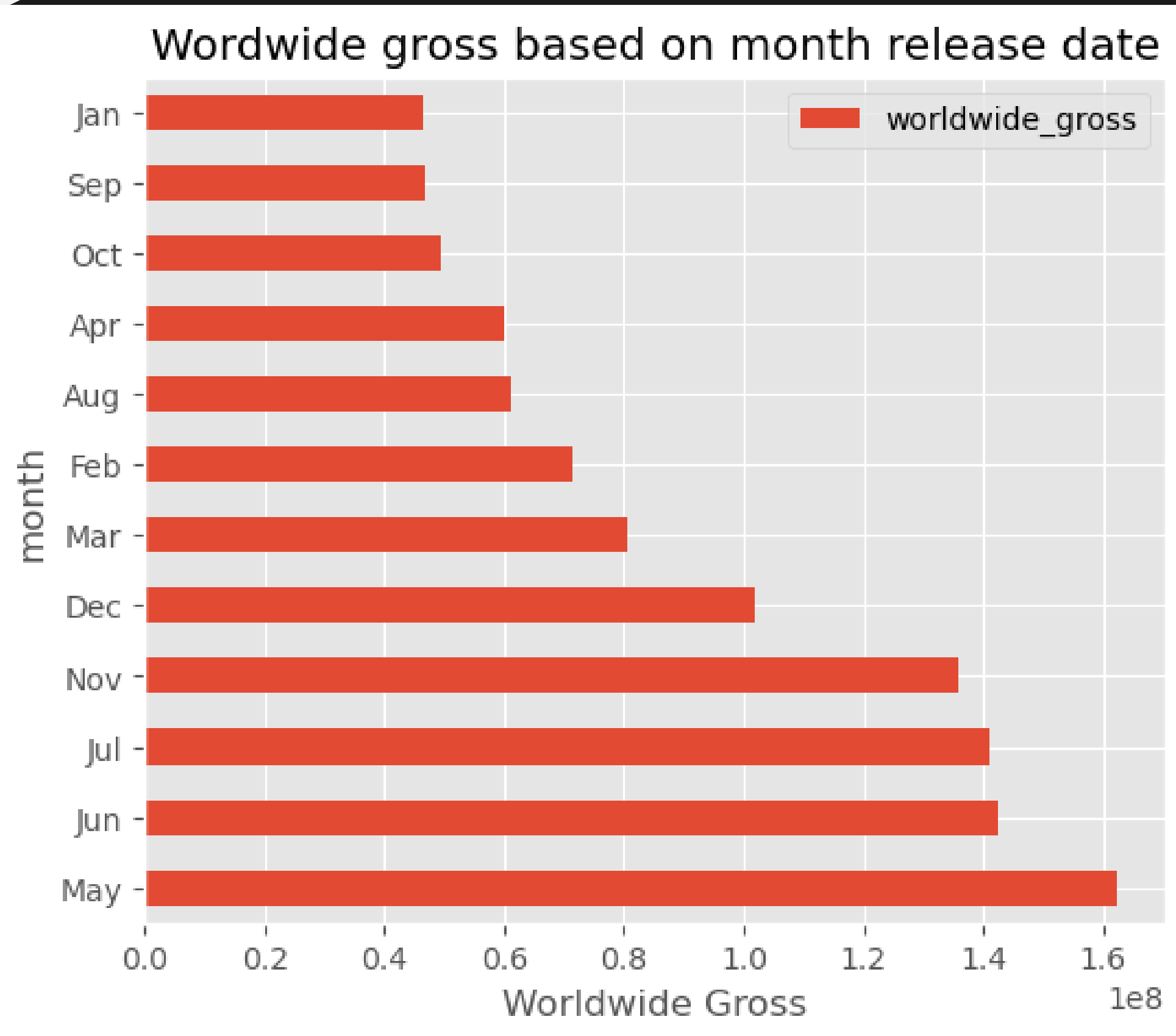
THE GRAPH SHOWS TOP 10 YEARS
IN TERMS OF HOW MUCH MOVIES
HAVE GROSSED.(AVERAGE
WORLDWIDE GROSS PER YEAR)

DESCRIPTIVE STATISTICS



THE GRAPH SHOWS TOP 10 YEARS IN TERMS OF HOW MUCH MONEY WAS USED IN THE PRODUCTION BUDGET. ITS CLEAR THAT OVER THE YEARS, PRODUCERS HAVE PUMPED IN MORE MONEY TO THE PRODUCTION PROCESS.

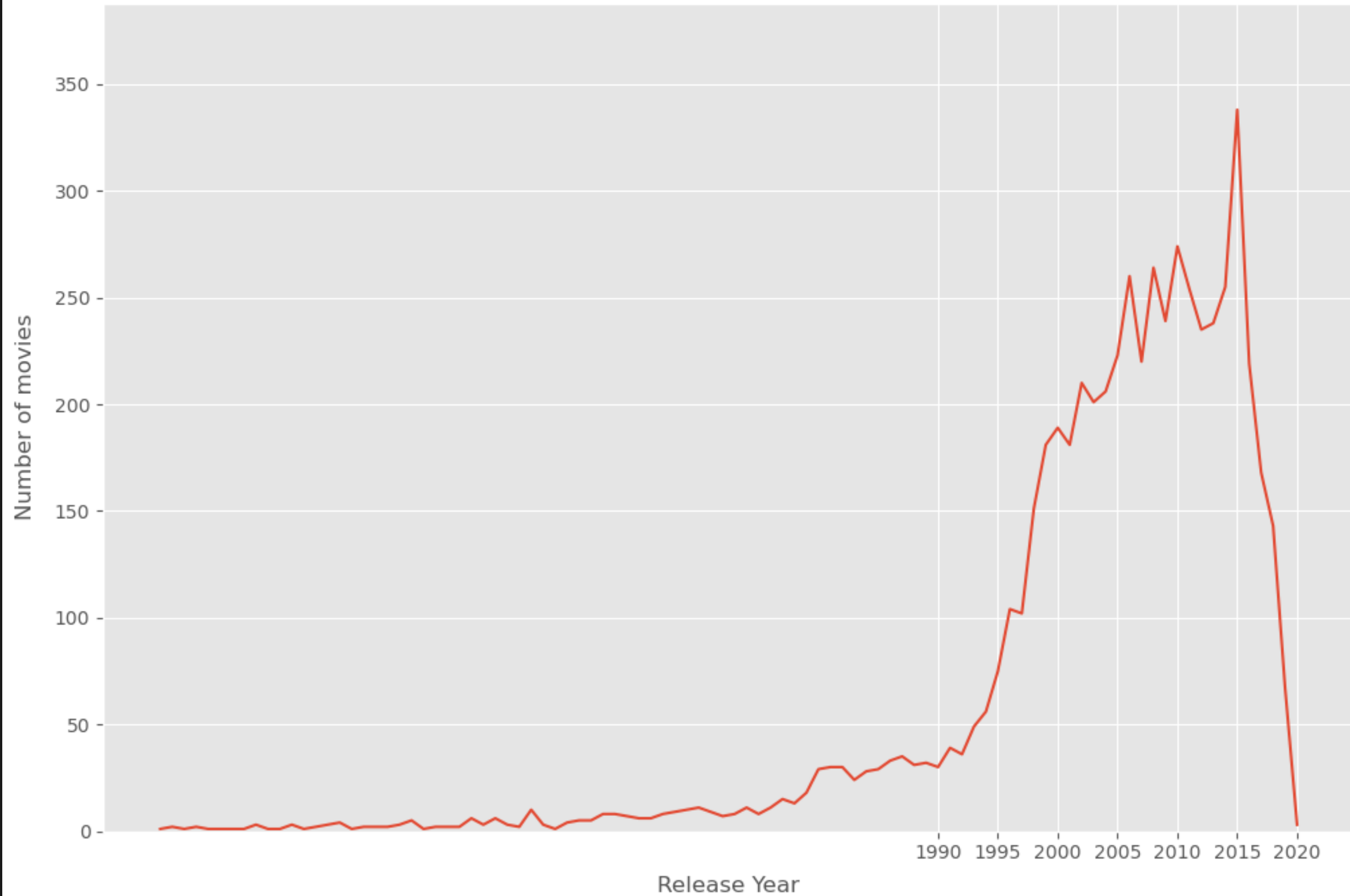
DESCRIPTIVE STATISTICS



THE GRAPH SHOWS AVERAGE MOVIES HAVE GROSSED BASED ON THE RELEASE MONTH. MOVIES RELEASED IN MAY HAVE GROSSED MORE AS COMPARE TO THE REST.

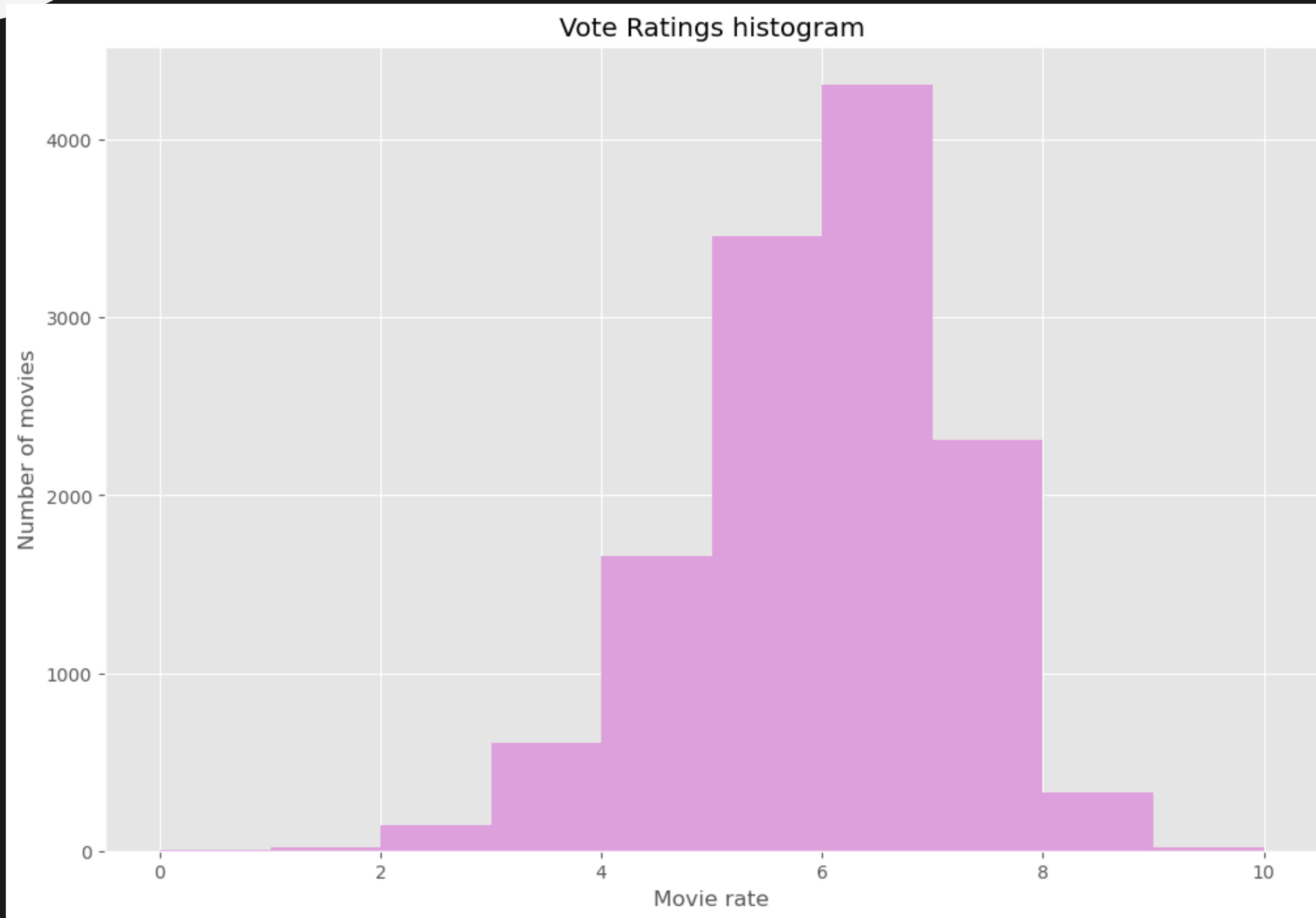
DESCRIPTIVE STATISTICS

Movies Grouped By Year Of Release



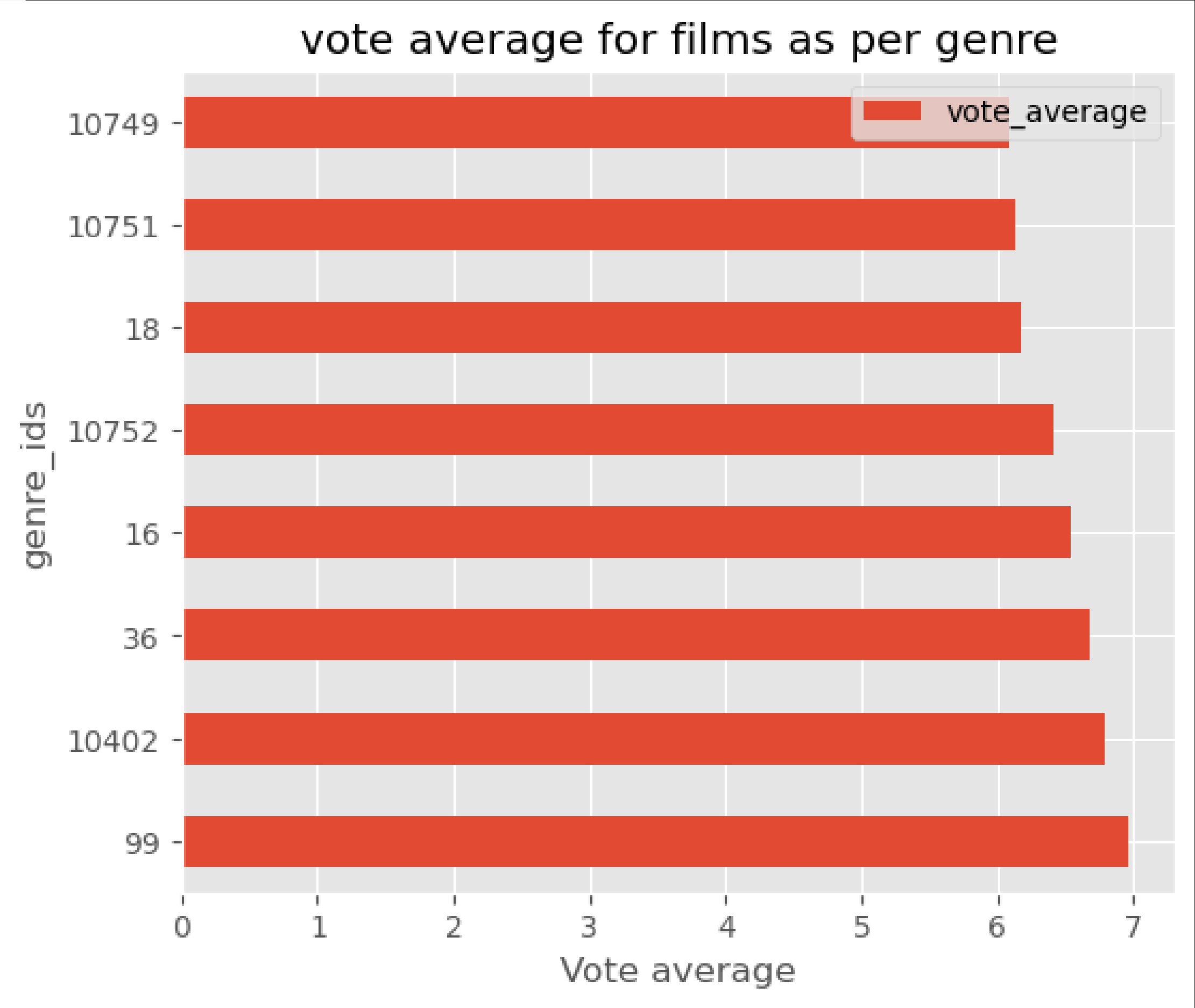
THE GRAPH SHOWS THE
DISTRIBUTION OF MOVIE
RELEASES OVER THE YEARS

DESCRIPTIVE STATISTICS



THE HISTOGRAM SHOWS THE
DISTRIBUTION OF THE RATINGS OF
MOVIES MOST FALL BETWEEN 6
AND 8

GENRE AND RATING ANALYSIS

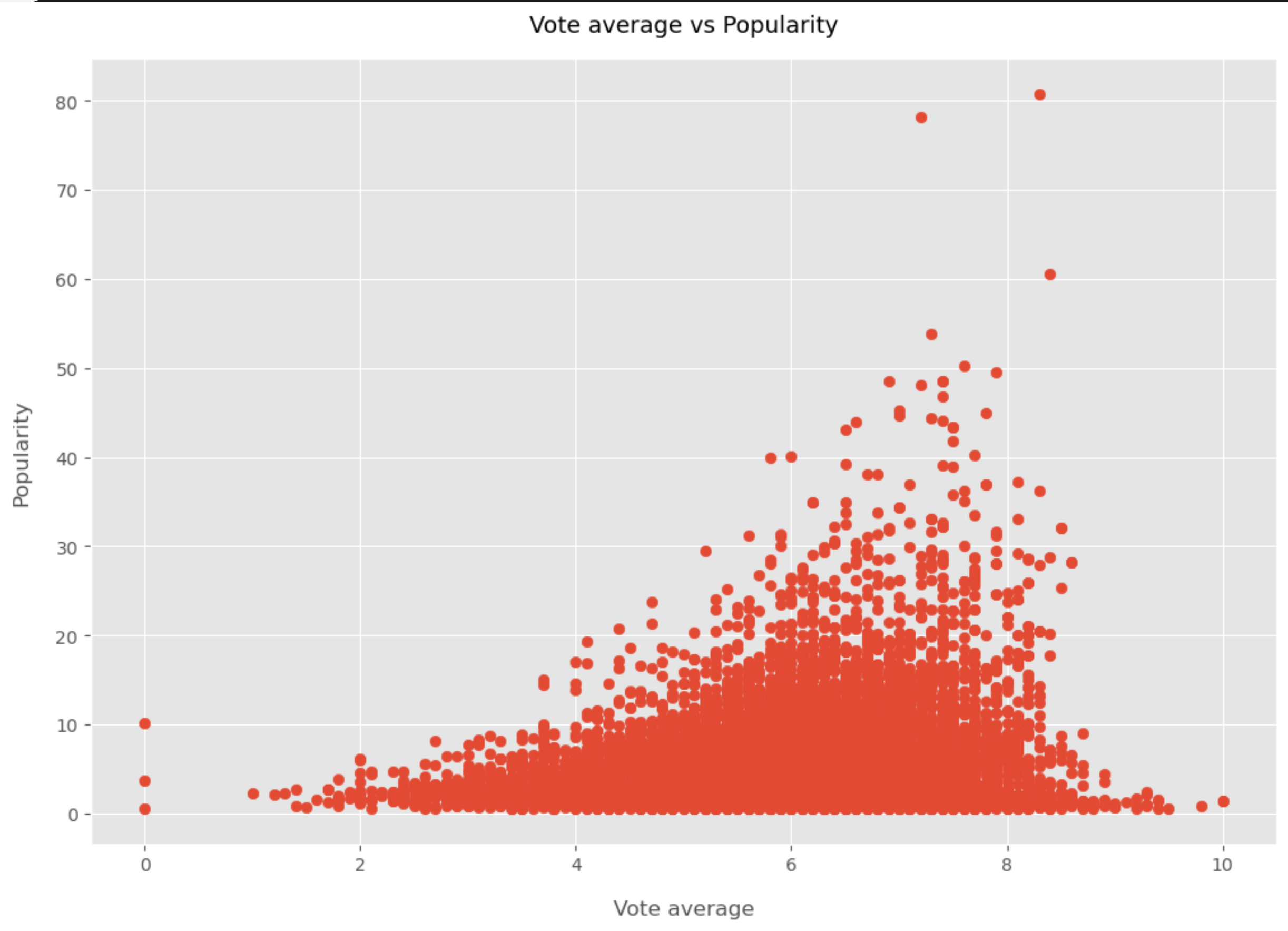


THE GRAPH SHOWS TOP 8 RATED GENRES:

KEY:

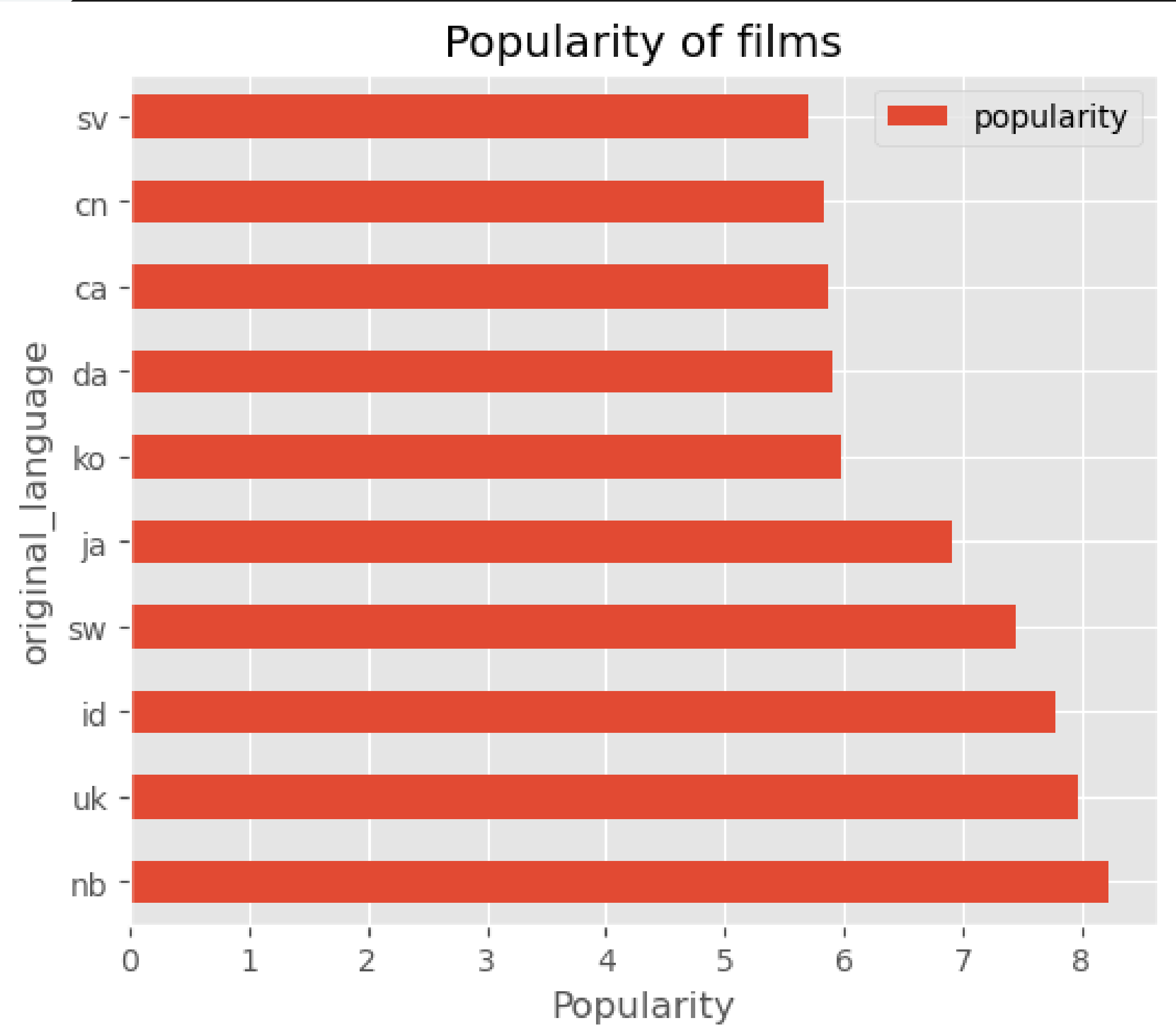
- 10749 ROMANCE
- 10751: FAMILY
- 18 DRAMA
- 10752 WAR
- 16 ANIMATION
- 36 HISTORY
- 10402 MUSIC
- 99 DOCUMENTARY

GENRE AND RATING ANALYSIS



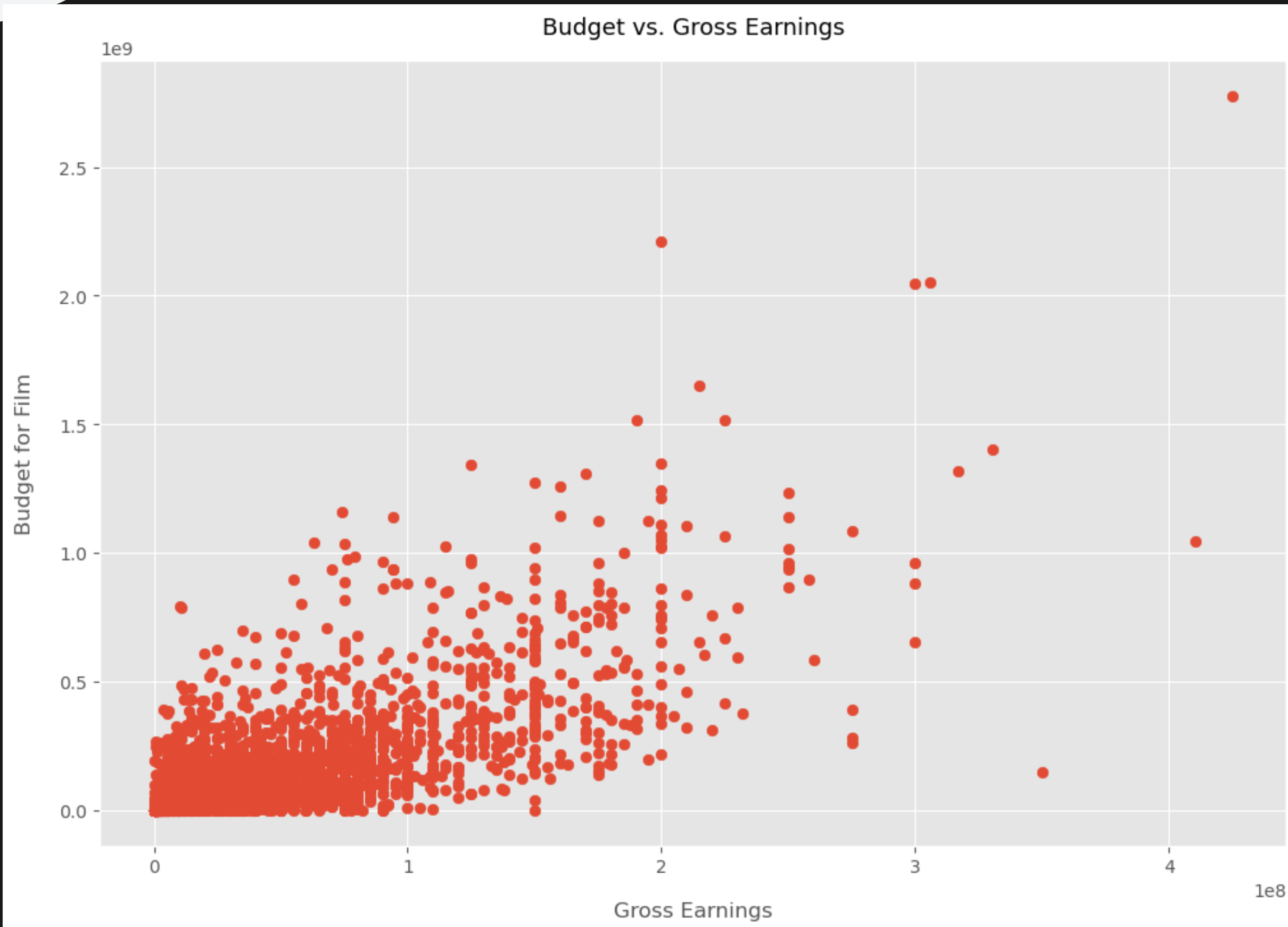
THE GRAPH SHOWS CORRELATION
BETWEEN POPULARITY OF MOVIES
AND THE RATINGS

GENRE AND RATING ANALYSIS



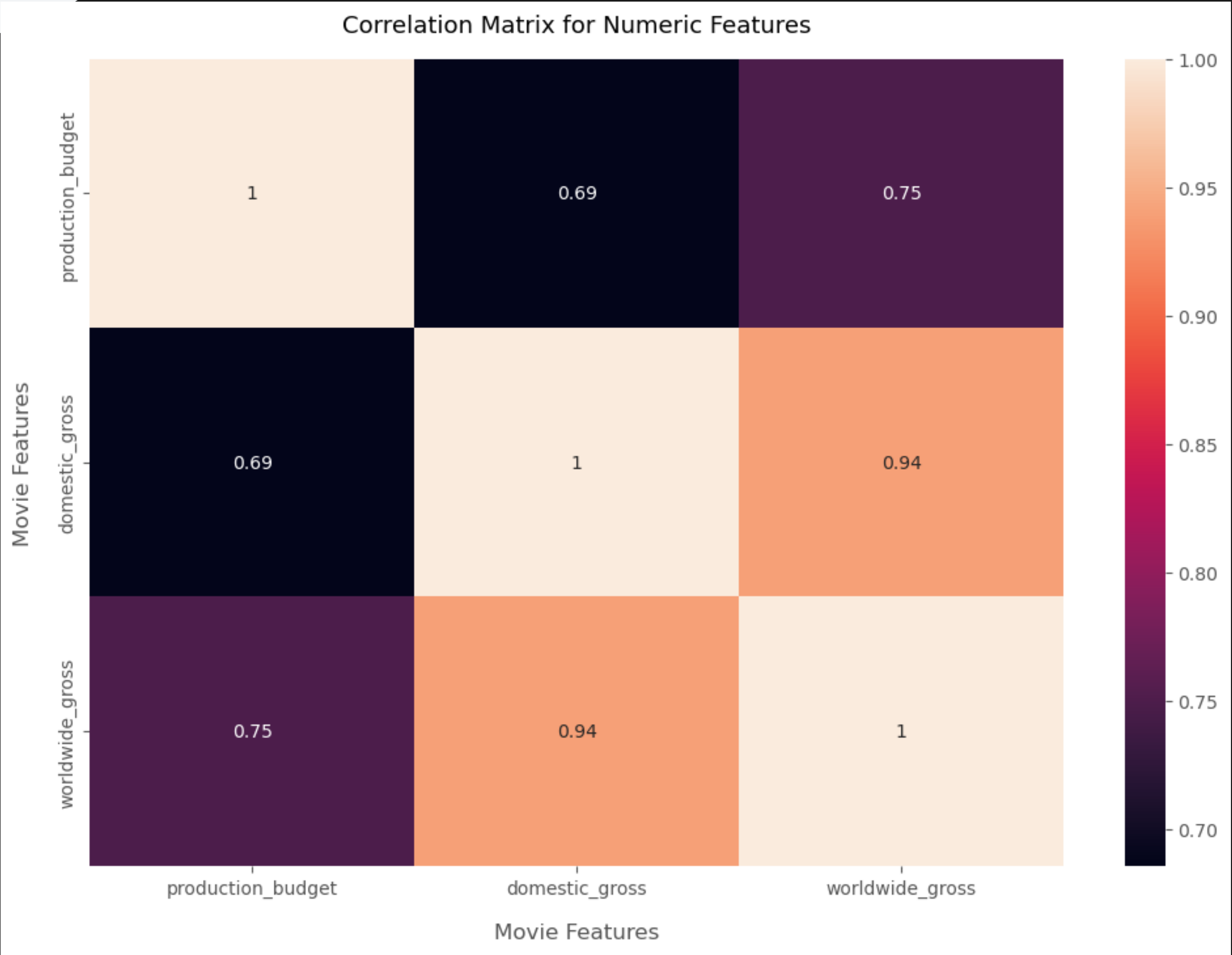
THE GRAPH SHOWS HOW
POPULAR MOVIES ARE BASED ON
THEIR ORIGINAL LANGUGE.

REVENUE AND BUDGET ANALYSIS



THE SCATTER PLOT SHOWS THE RELATION SHIP BETWEEN THE BUDGET AND THE GROSS EARNING OF MOVIES.

REVENUE AND BUDGET ANALYSIS



THE CORRELATION MATRIX EMPHASIZES ON HOW PRODUCTION BUDGET AND GROSS OF MOVIES RELATES TO EACH OTHER.



RECOMENDATIONS

As per the analysis conducted on the movie dataset, I have come up with three recommendations:

- Microsoft should put in money in the production as statistics have shown that the budget has a direct relationship on how much the movies grosses.
- Microsoft should venture in films of genres(Romance, family, drama, war, animation ,history, music, documentary) as they have proved to have better ratings.
- Microsoft should invest in movies originally in languages Norwegian, British, Indonesian, Swedish as they are more popular. The analysis has also proves that popular movies have better ratings.