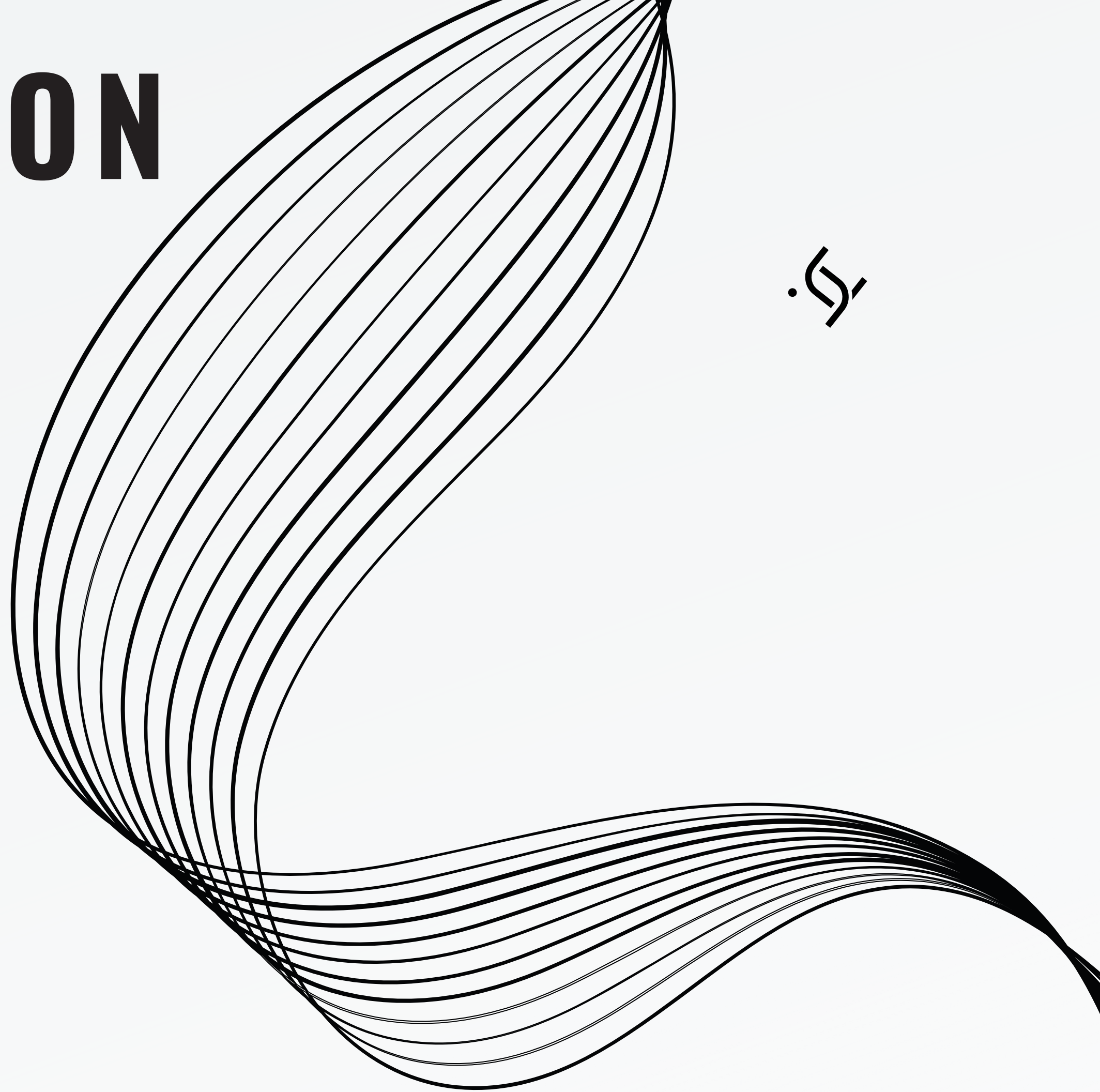# MICROSOFT'S MOVIE

JULIE CHEPNGENO

INTRODUCTION

In the realm of entertainment, the magic of movies has captivated audiences for generations.It is within this ever-evolving landscape that a new movie studio emerges, ready to make its mark on the industry.

LARANA, INC.

INTRODUCTION

The project will involve analyzing different movies with the aim of finding he most favourable one will in turn be of key value in making strategic decisions for our studio regarding production and financing.

PROBLEM STATEMENT

Microsoft sees all the big companies creating original video content and they want to get in on the fun. They have decided to create a new movie studio, but they don't know anything about creating movies. You are charged with exploring what types of films are currently doing the best at the box office. You must then translate those findings into actionable insights that the head of Microsoft's new movie studio can use to help decide what type of films to create.

MAIN OBJECTIVE

The goal of this analysis is to provide an insight for Microsoft regarding the latest box office films and what is popular and most successful so they can leverage this information to provide quality content in their new movie studio

SPECIFIC OBJECTIVE

To translate those findings into actionable insights that
the head of Microsoft's new movie studio can use to
help decide what type of films to create.

To equip the Microsoft's management team with the
necessary knowledge to determine the most succeful
genres, themes, and strategies for achieving success in
this competitive domain.

NOTEBOOK STRUCTURE

* Data Collection
* Read and check the data
* Cleaning the data
* Exploratory Data Analysis
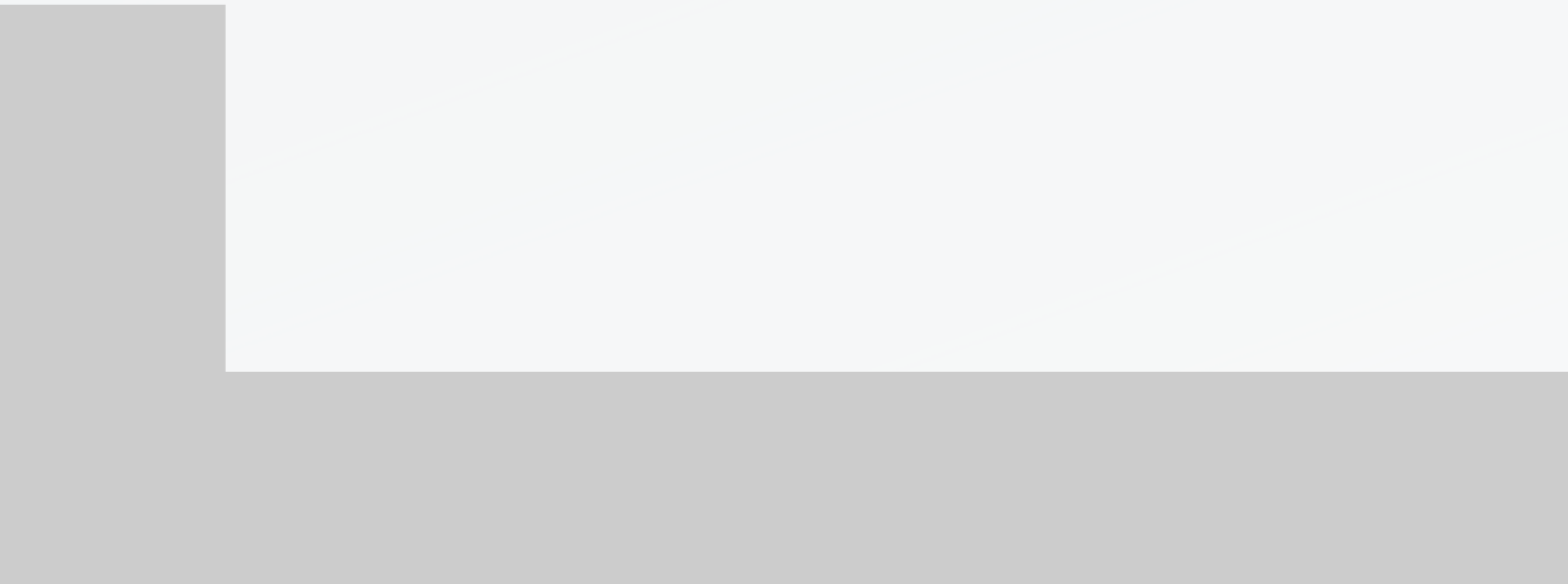* Conclusions and Recommendations

DATA UNDERSTANDING

The data was collected from various locations, the different files have different formats. Some are CSV or TSV files that can be opened using spreadsheet software or pd.read_csv, while the data from IMDB is located in a SQLite database.The data used in this project were downloaded from tn.movie_budgets.csv,title.basics.csv and title.rating.

DATA UNDERSTANDING

.
Analyzing dataset will determine what contributes to the "success" of a movie. In this analysis, I will define analysis in financial terms basing the success of a movie on the amount of money it earns in comparison to the budget.
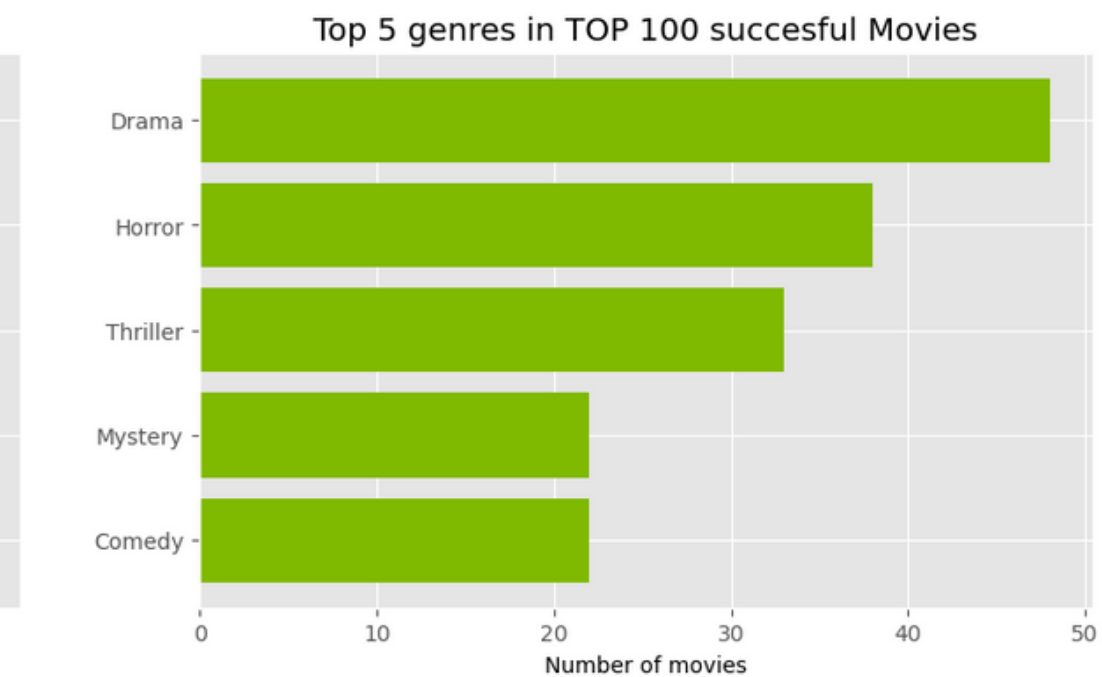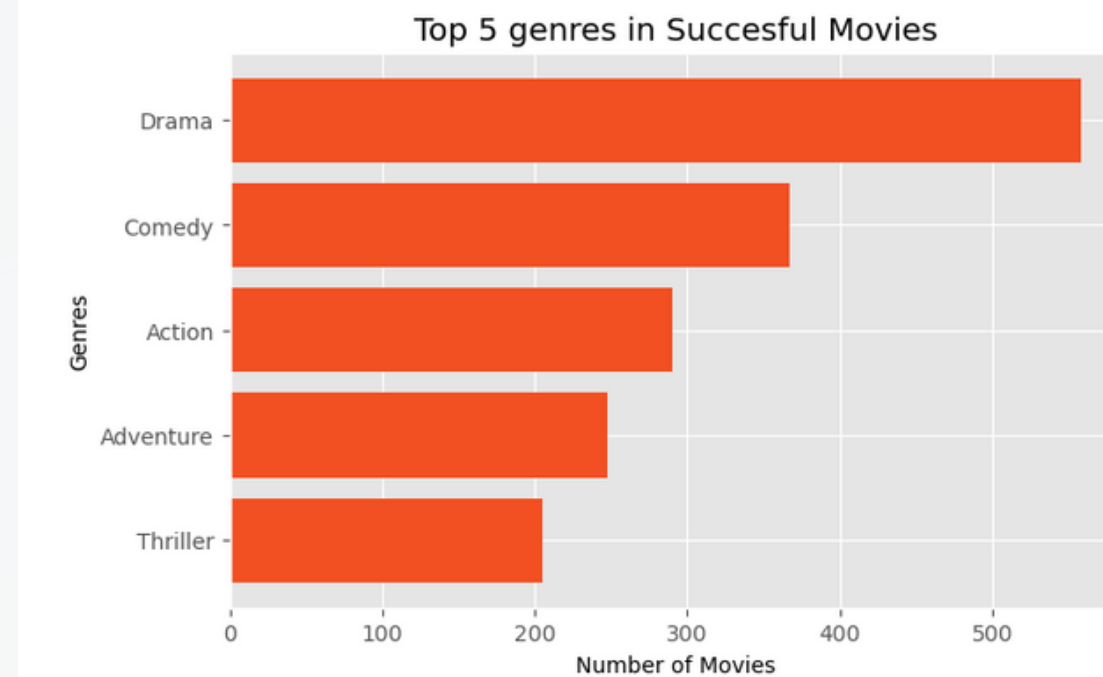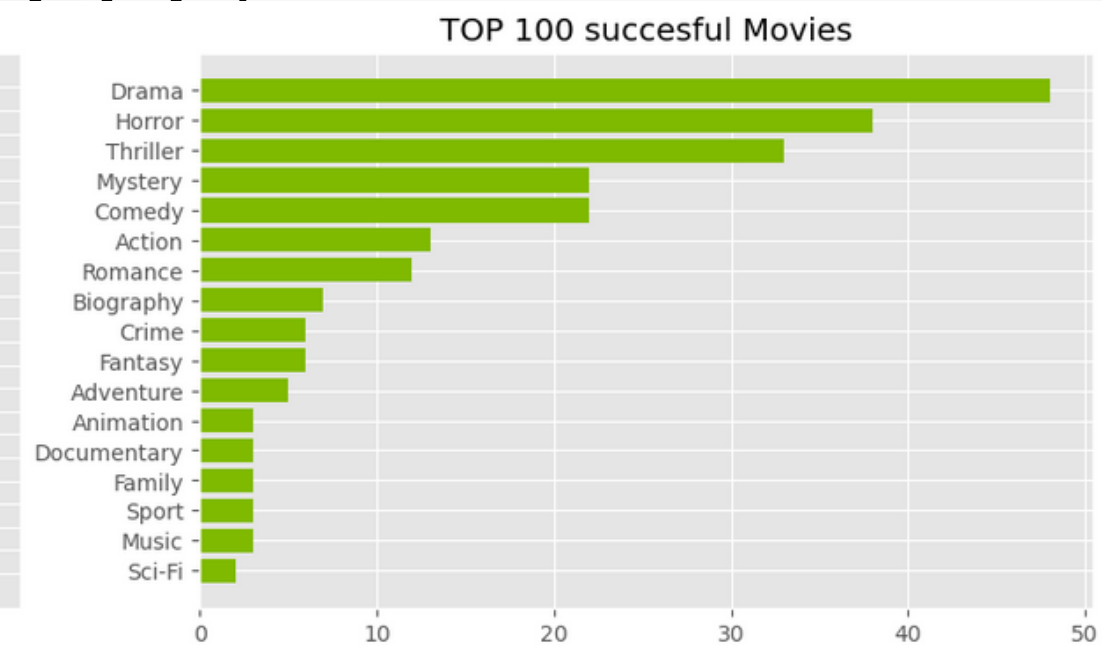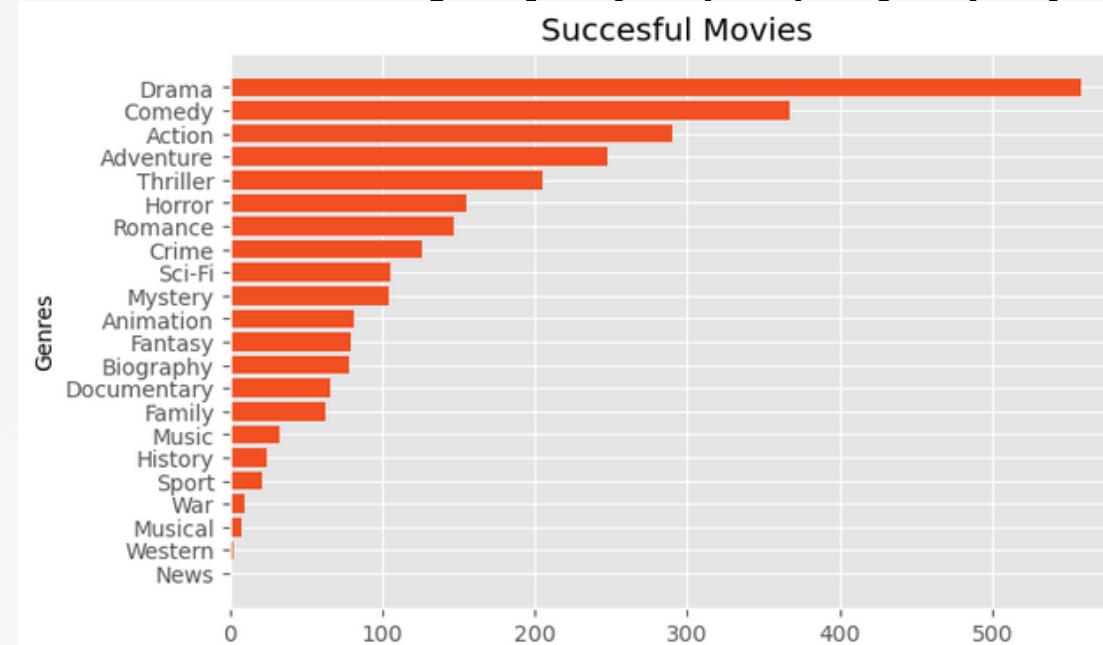
DATA CLEANING

.
Out of the several datasets that were collected, only
some features and rows are relevant to the process.
Therefore, in this step, the features that are not
required from each dataset were dropped.The
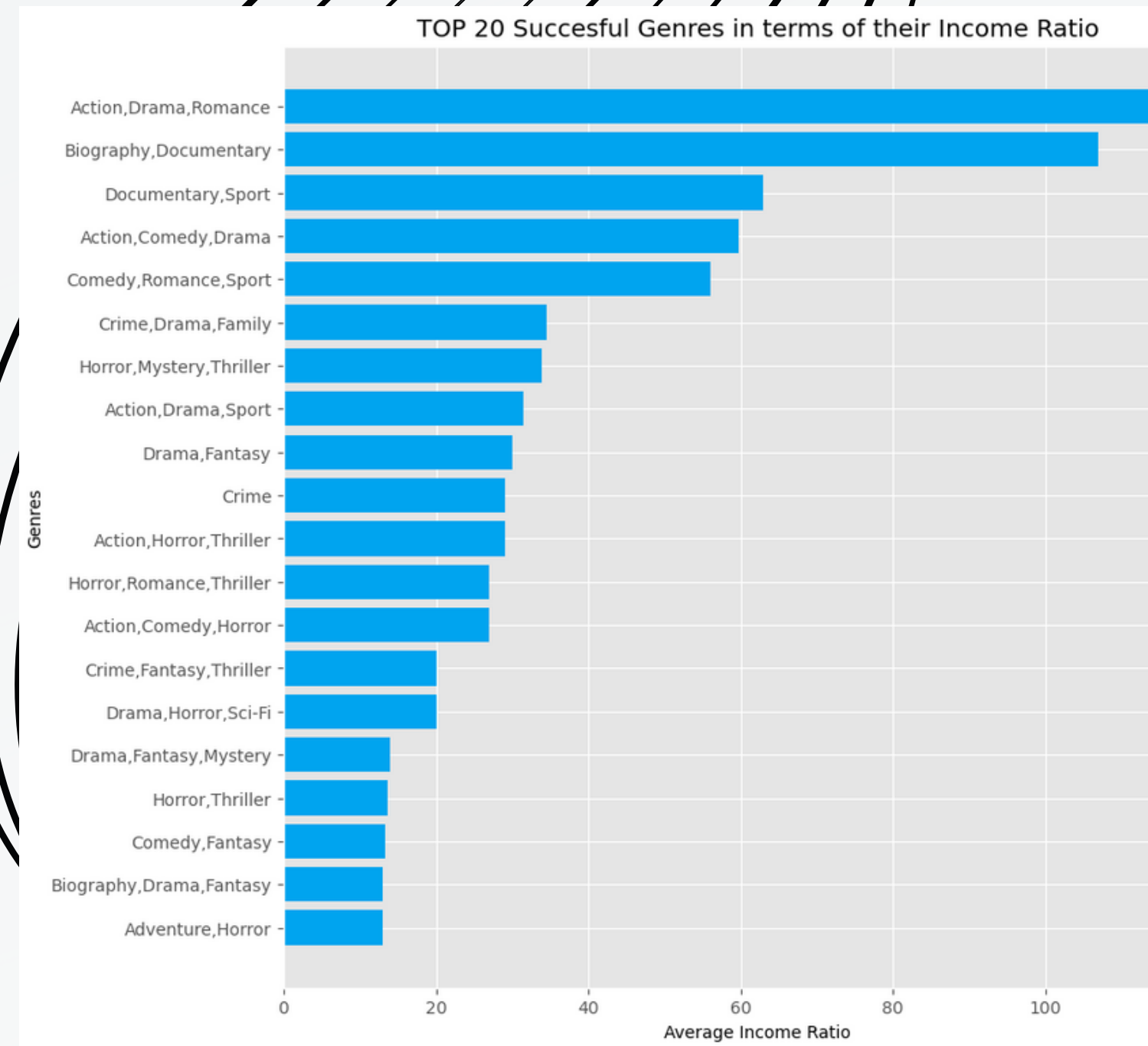remaining datasets were then joined

DATA ANALYSIS

.* From our analysis, we can see that in the top 20 most succesful genres ,the combination of genres; ACTION, DRAMA, ROMANCE had the most in average income ratio
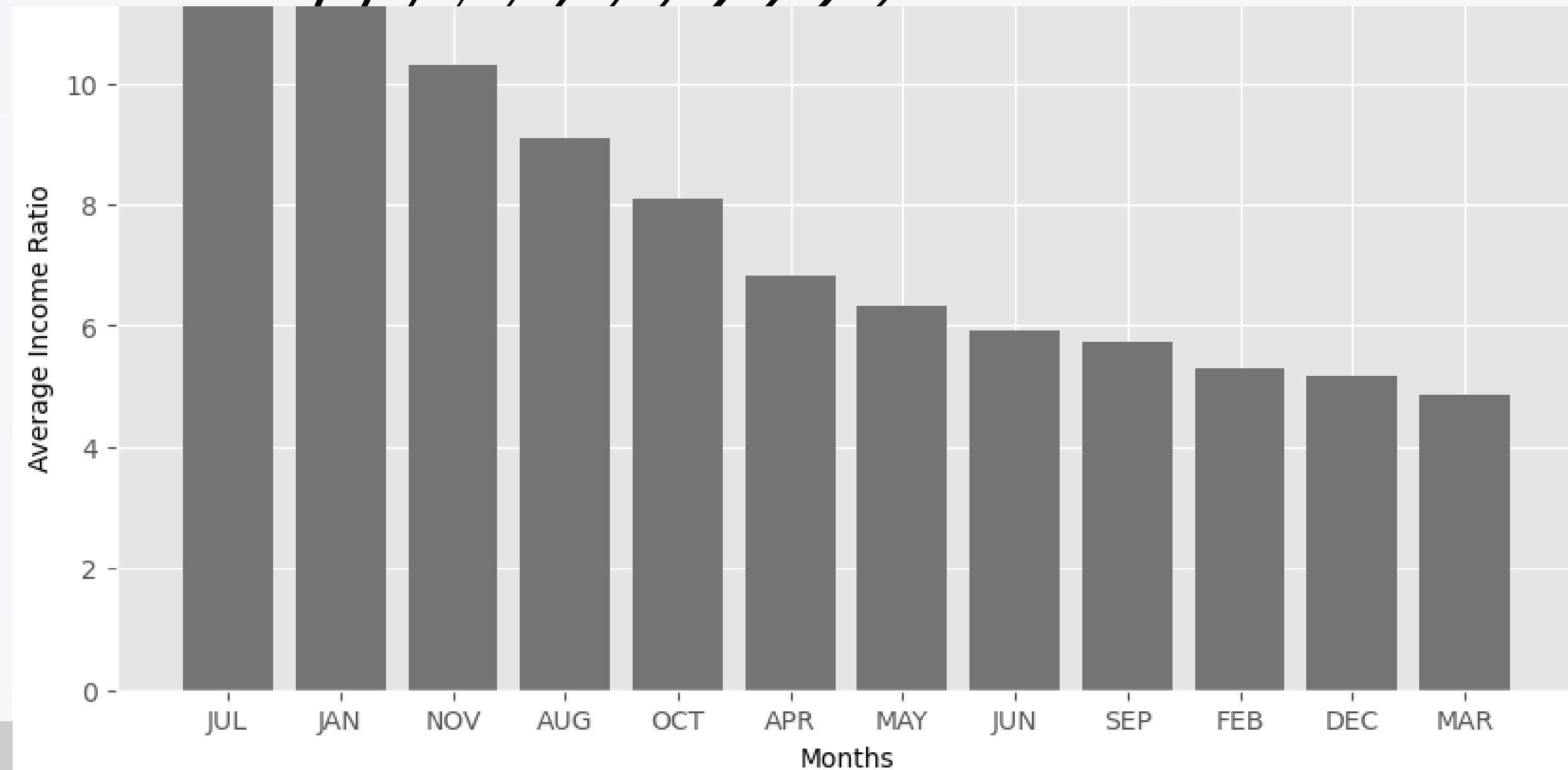
DATA ANALYSIS

* The data from the number of votes is skewed as we can see that the majority of successful movies have less than 100,000 votes with the average number of votes being 123,205 votes. There is a very huge variance in the number of votes (182,927 votes) due to the very many outliers above 400,000 votes. This tells us that there is no standard number of votes required for the average rating.

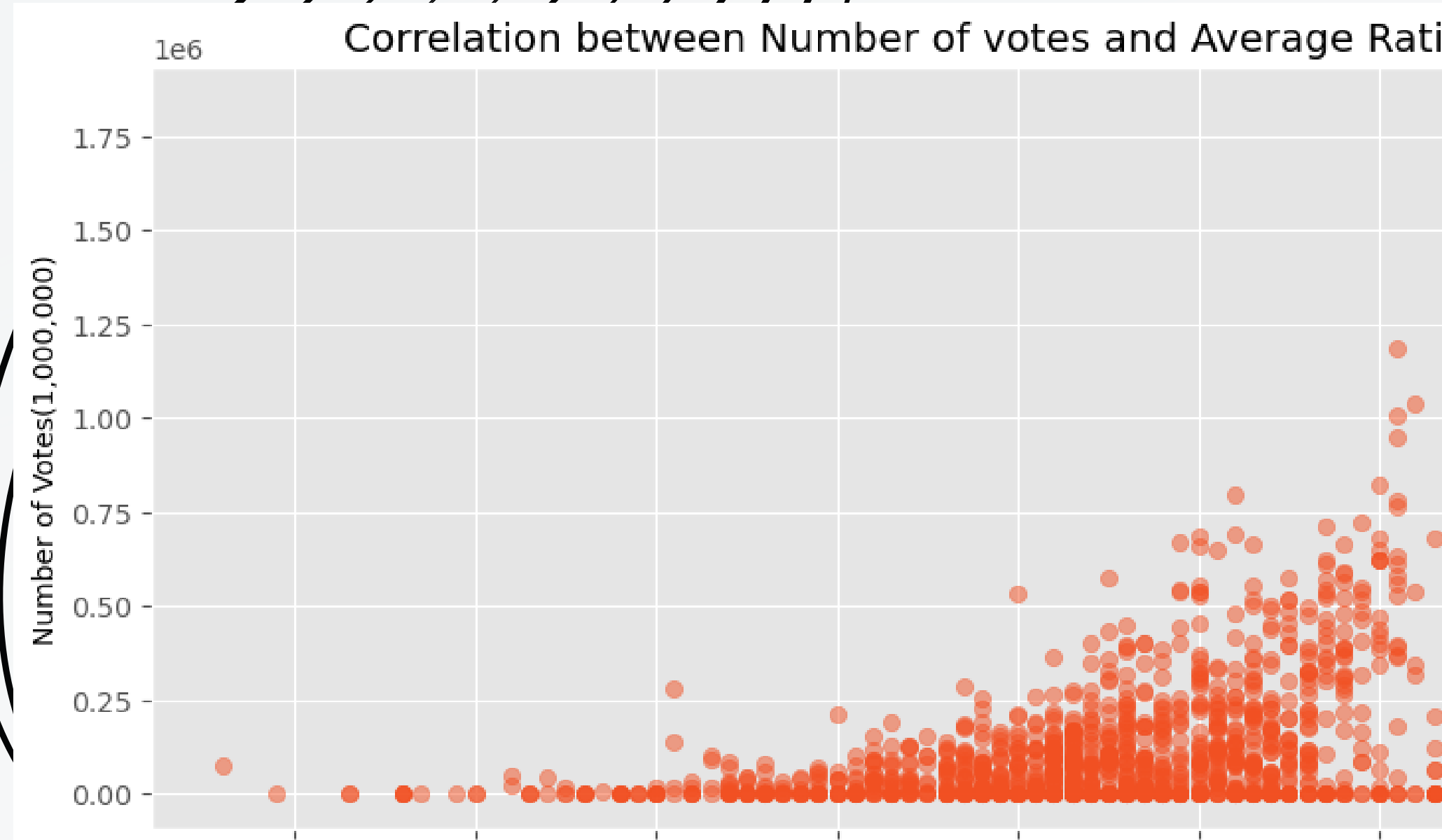## TOP 20 Succesful Genres in terms of their Income Ratio

DATA ANALYSIS

In the analysis above, we've found out that the best months to release a movie considering the average income ratio, are July, January and November
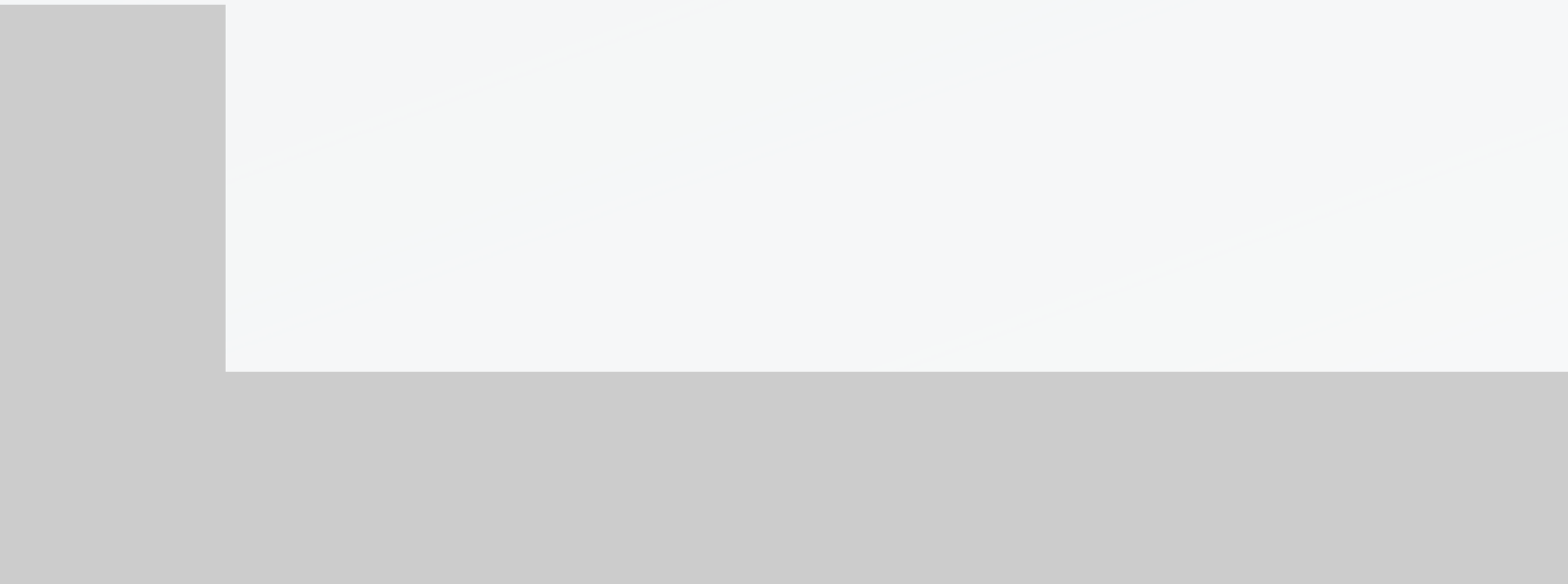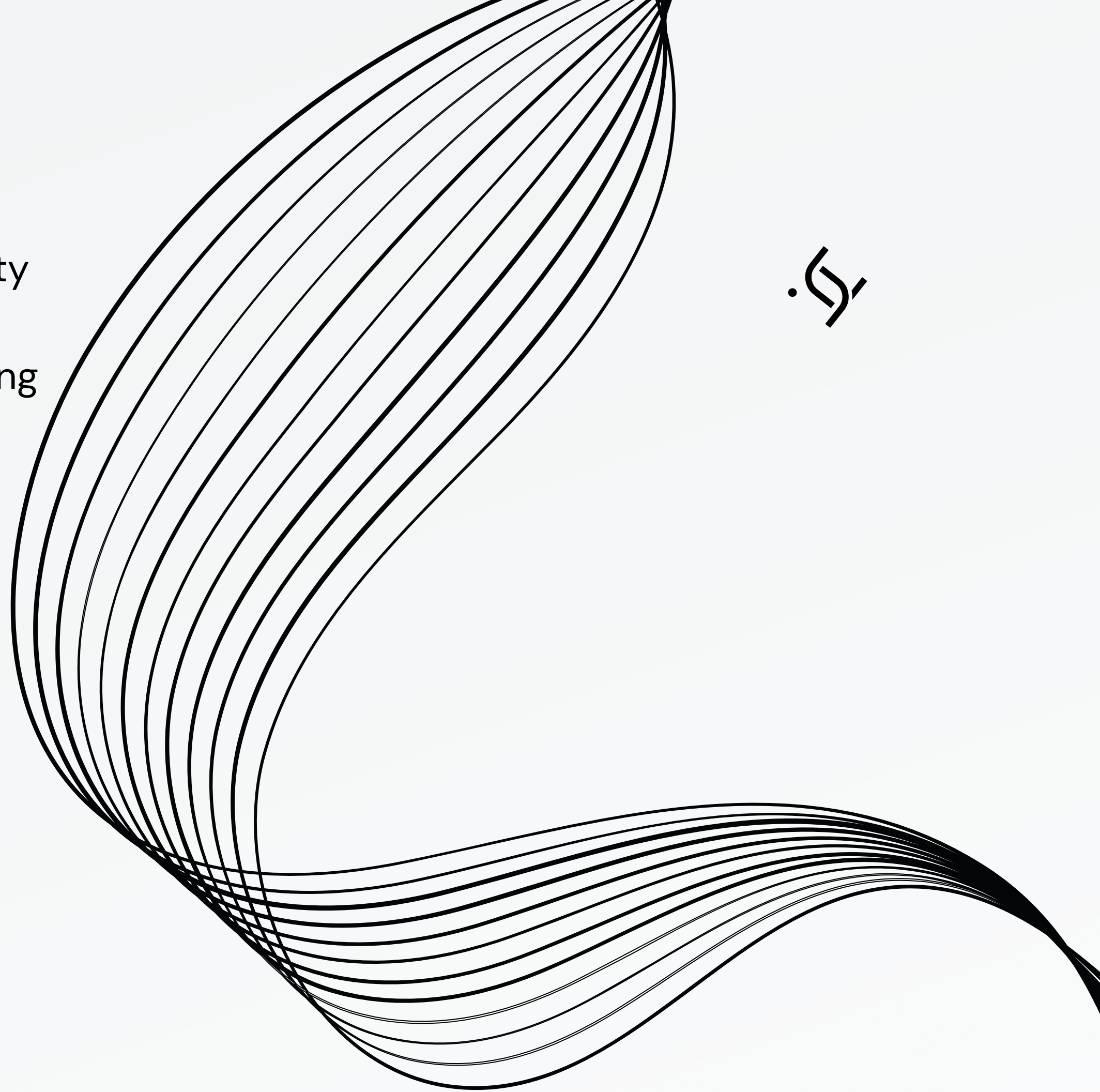
DATA ANALYSIS

* There is a correlation coefficient of 0.407 which indicates a low positive correlation. There is a low positive correlation between these two. We can see that the number of votes slightly affects the average rating

Correlation between Number of votes and Average Rati...

CONCLUSIONS

1. From our boxplot, we see there's dual seasonality of higher profits in July, January, and November. The highest profit in summer can be realized during the month of July.

2. Drama, thriller and comedy are likely to attract more viewers

RECOMMENDATIONS

1. From our findings, we can see that DRAMA, THRILLER, and COMEDY are the most common genres in successful movies. Therefore I highly recommend the three film genres. Most preferably DRAMA since it is cheaper to produce and has a high-income ratio.

RECOMMENDATIONS

2. The success of a movie also highly depends on the release month. I'd also want to recommend three months which proved best in terms of the income generated. They include **_JULY, JANUARY, and NOVEMBER_** with **JULY** having the highest income ratio.

RECOMMENDATIONS

I would recommend a budget allocation within 24 million dollars which proved to be a good allocation for a successful movie