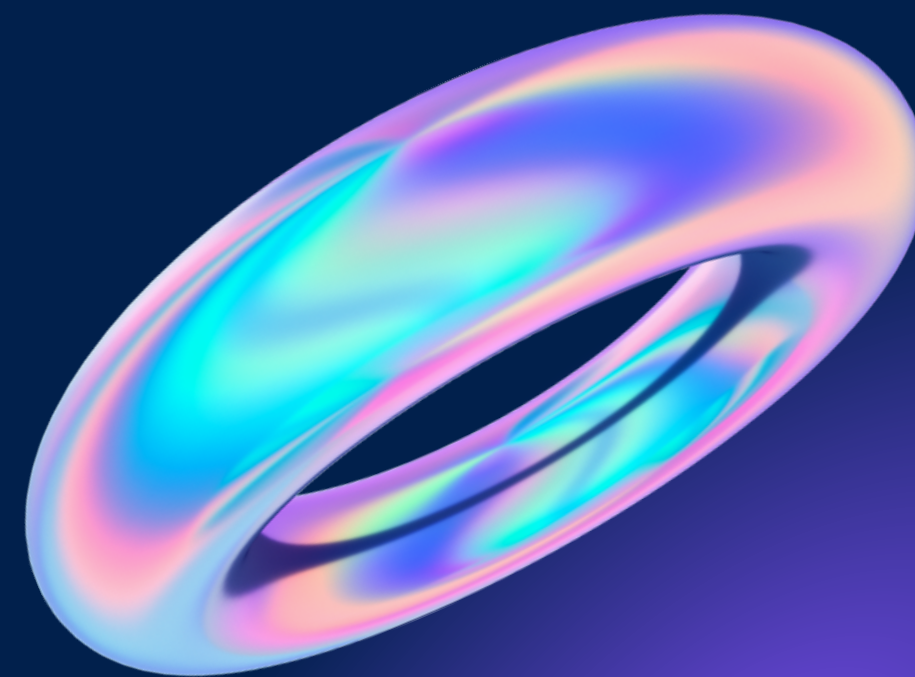# MICROSOFT'S MOVIE ANALYSIS

EDWIN CHERUIYOT

OVERVIEW

The movie industry is a dynamic and competitive sector encompassing the production, distribution, and exhibition of films. It is driven by creativity, diversity, and the desire to captivate audiences with compelling stories.

Movie studio are committed to fostering creativity, nurturing talent, and embracing diversity in all aspects in the entertaining industry and supporting emerging filmmakers, telling diverse stories that reflect the richness of our world, and contributing positively to the art of cinema.
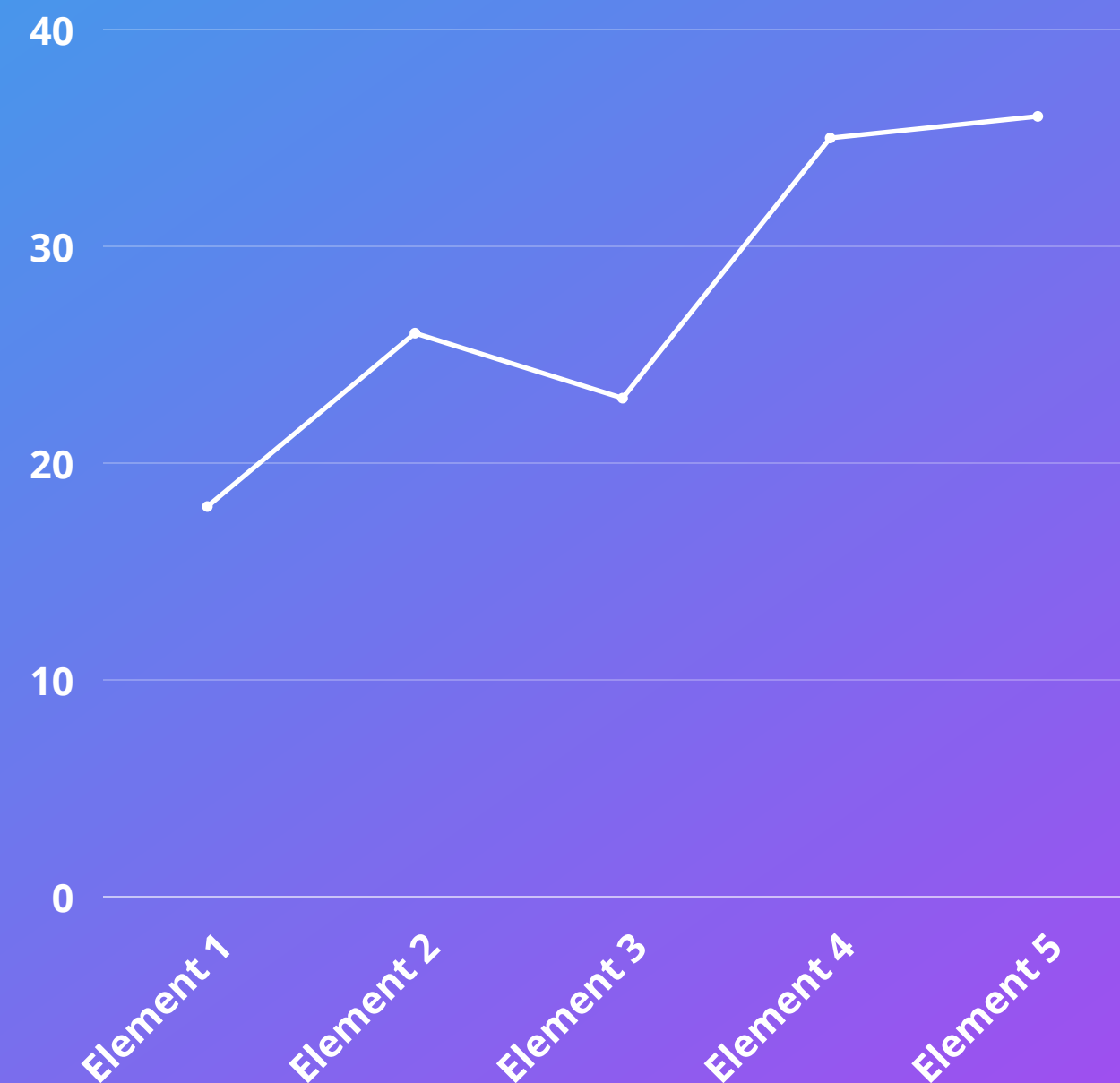
PROBLEM STATEMENT

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.
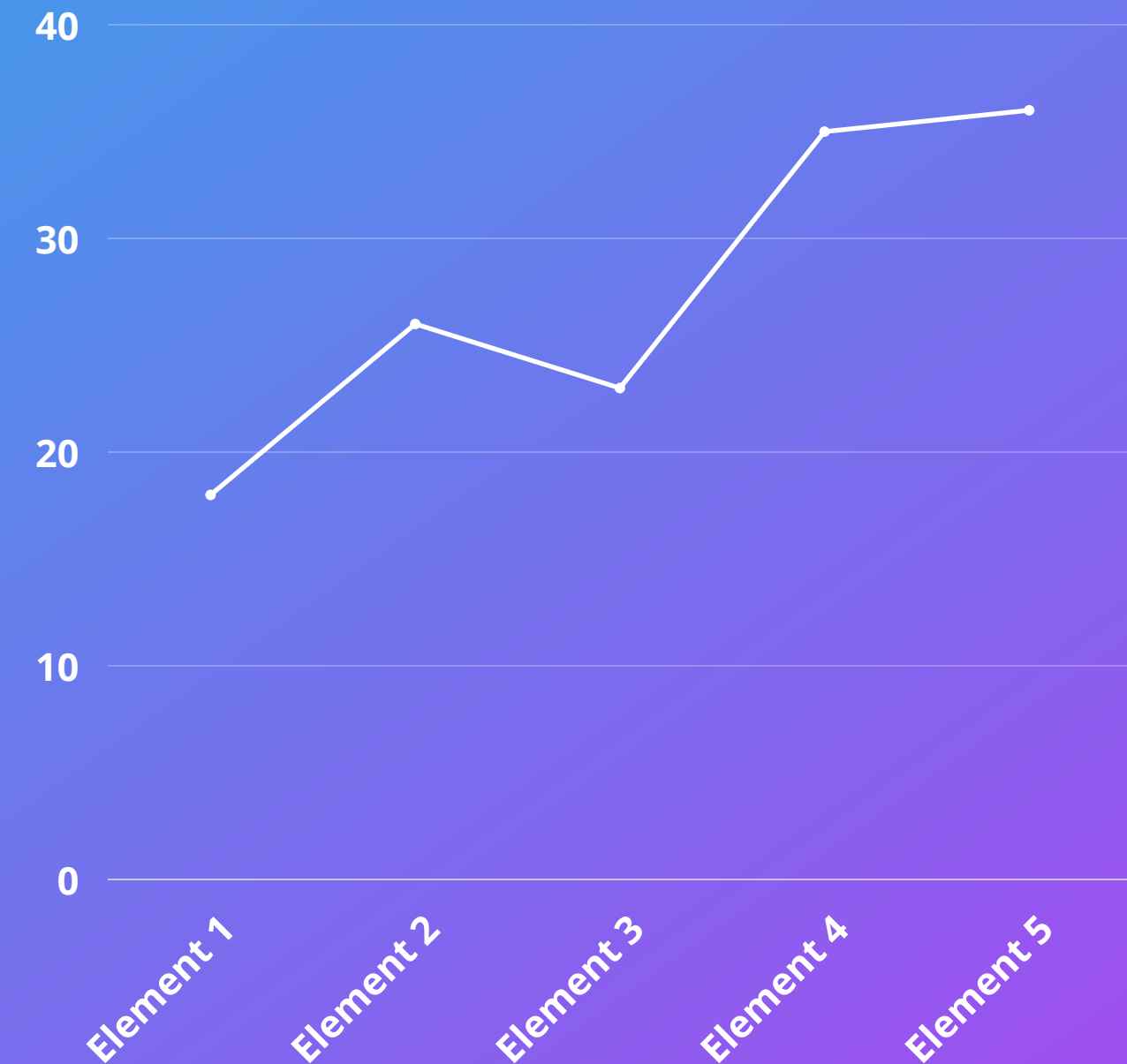
MAIN OBJECTIVE

The explore the types of films that have been performing exceptionaly well in the box office
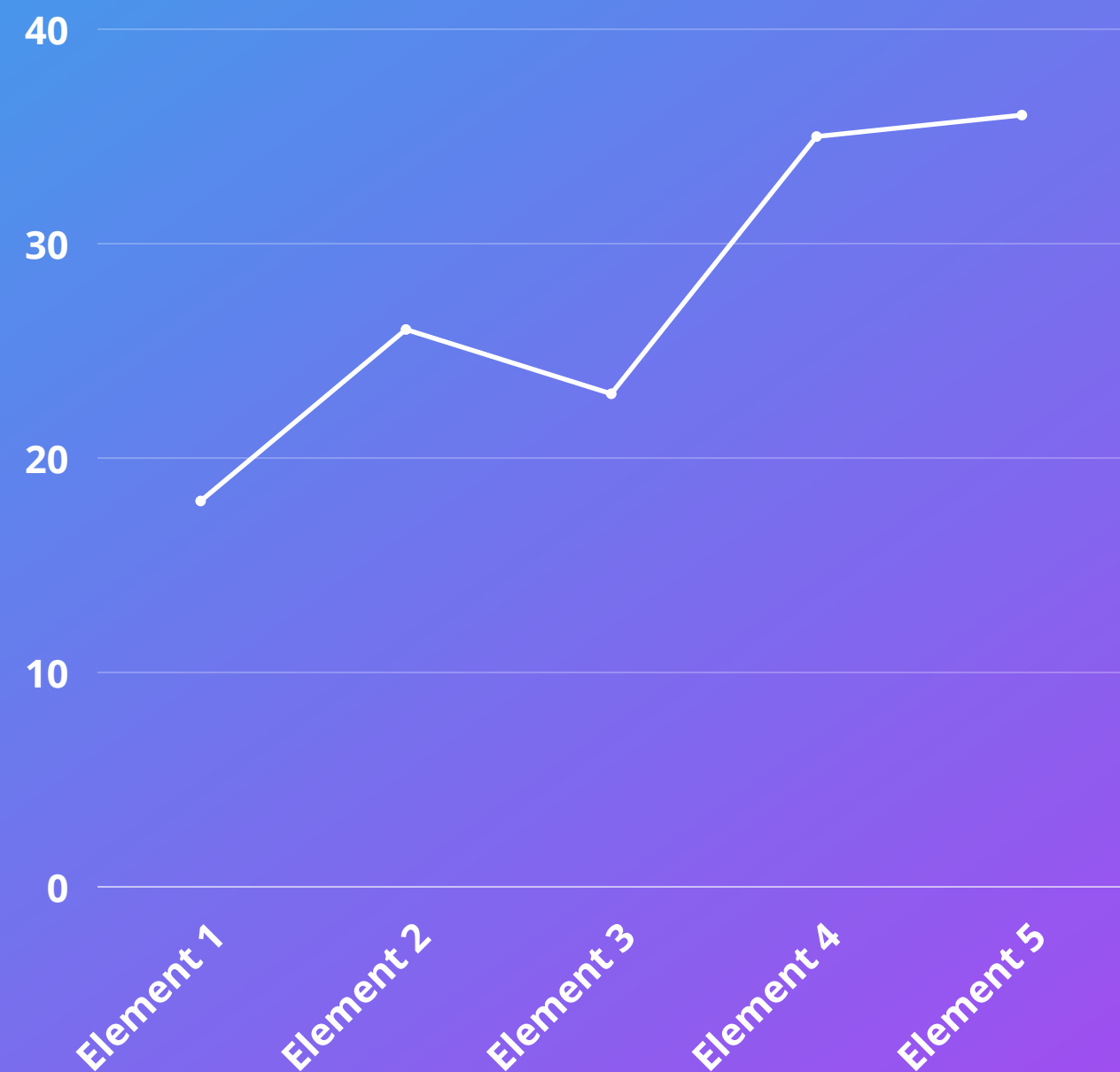
Specific Objectives

The primary objective of this project is to extract valuable insights that can effectively inform the decision-making process for Microsoft's newly established movie studio.

By analyzing relevant data and conducting thorough research, we aim to equip the management team with the necessary information and recommendations that will guide their decision-making process and help drive the company's success.
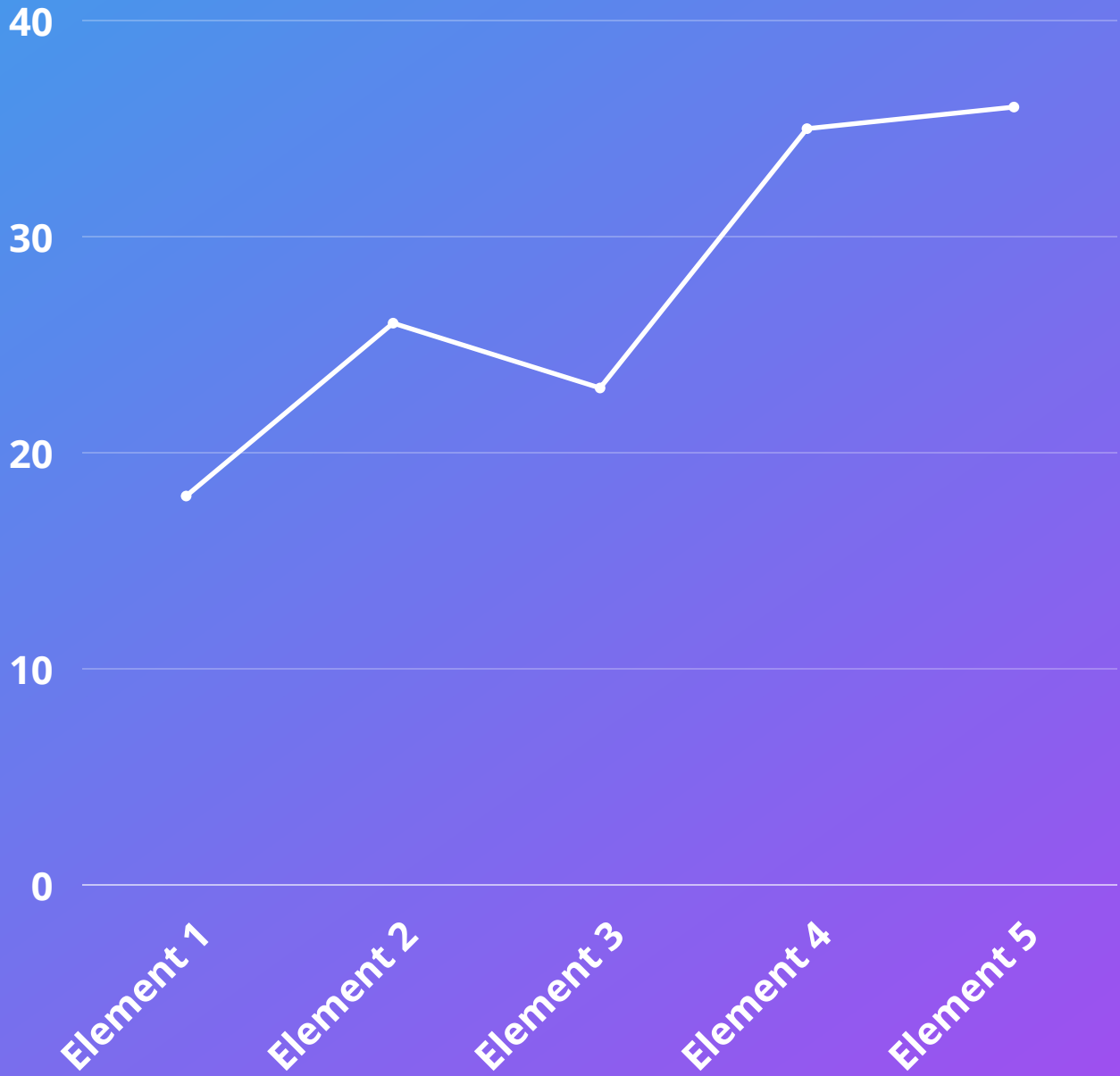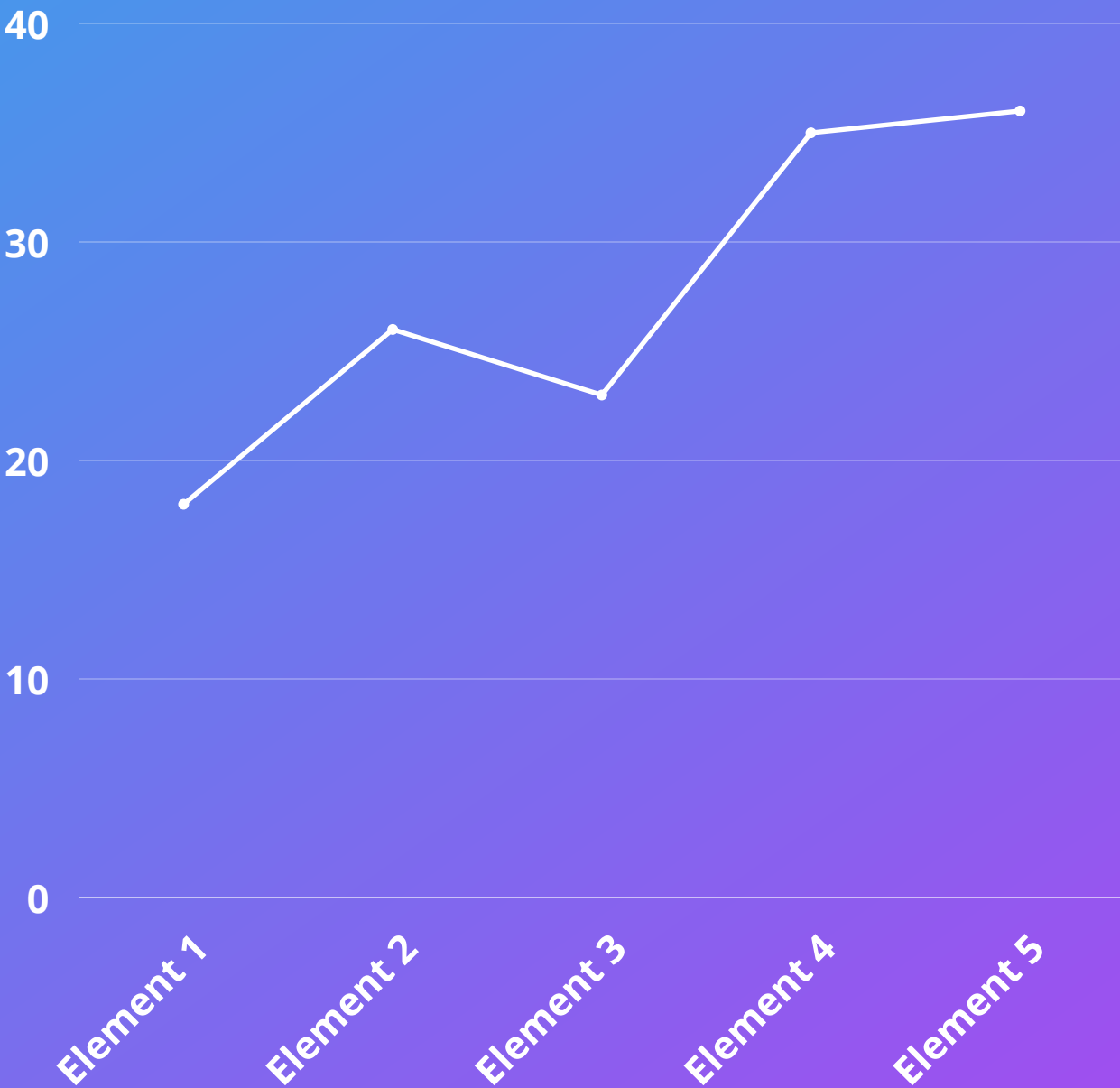
Notebook Structure

Data Collection

Read and check the data

Cleaning the data

 Exploratory Data Analysis

Conclusions and Recommendations

Data Understanding

The data we used was collected from various locations, the different files have different formats and Some are compressed CSV (comma-separated values) that can be opened using spreadsheet software or Pandas DataFrame using pd.read_csv, while the data from IMDB is located in an SQLite database.

Analyzing the dataset will determine what contributes to the "success" of a movie.
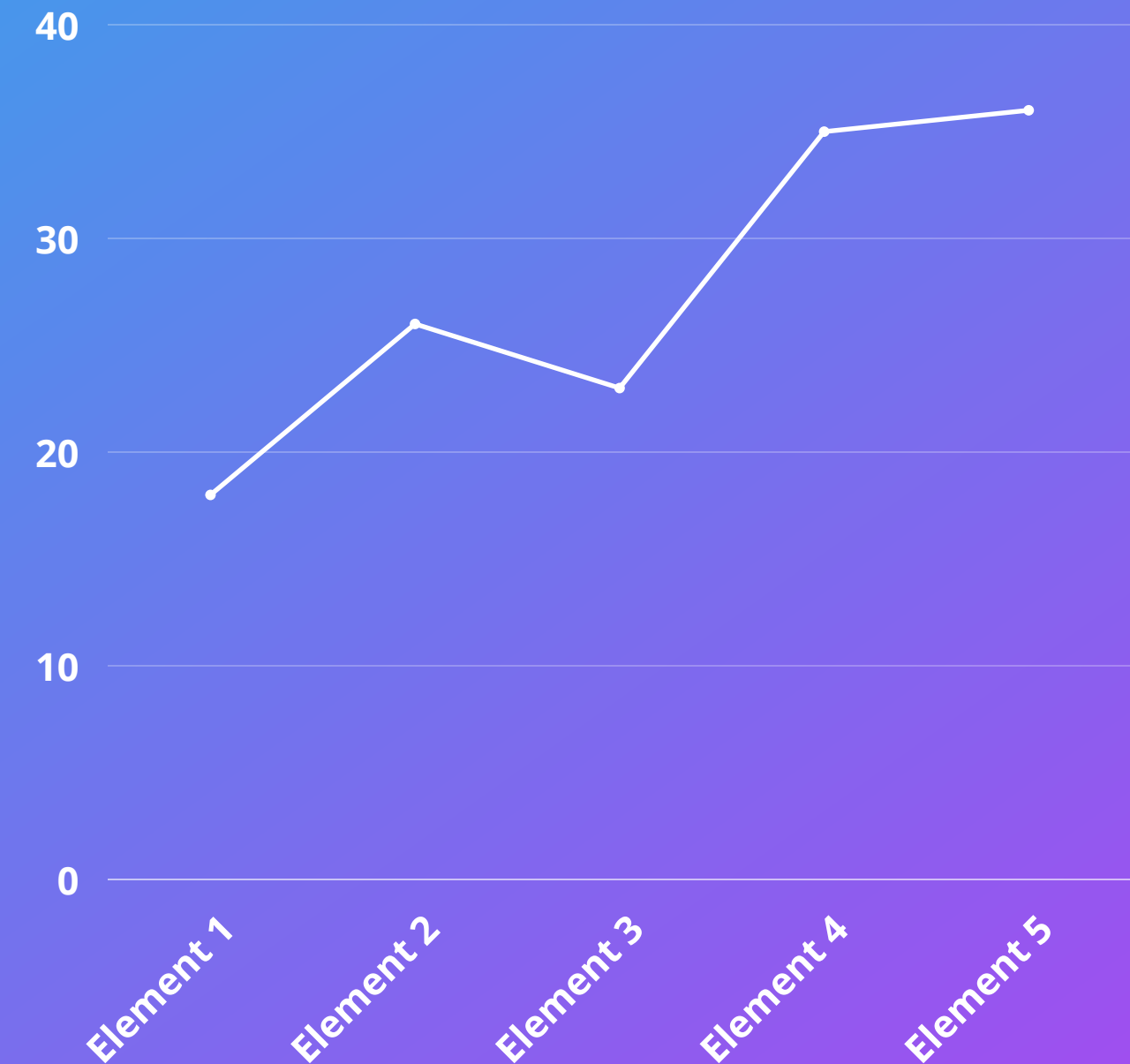
Data Wrangling

Selecting the relevant features and eliminating unnecessary information from the collected datasets.

This involved dropping the irrelevant features and rows from each dataset to streamline the data and the remaining datasets were merged or joined together to create a consolidated dataset for further analysis and exploration.

This step ensured that only the essential data was retained, reducing complexity and facilitating efficient analysis.
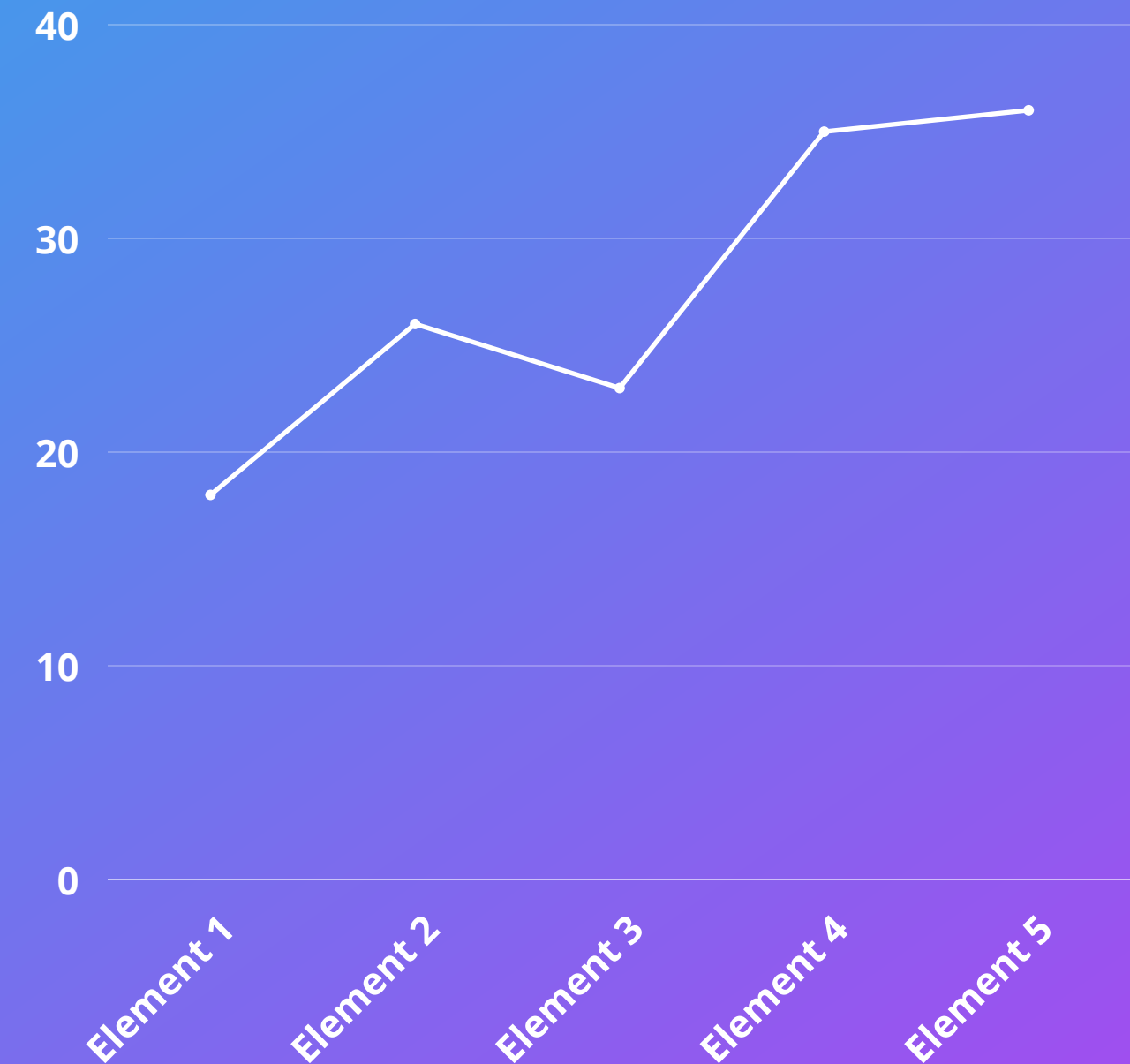
DATA ANALYSIS

Based on the above analysis plot, the top 20 most successful genres in terms of average income ratio are Action, Drama, and Romance. These genres consistently generated the highest average income compared to other genres.
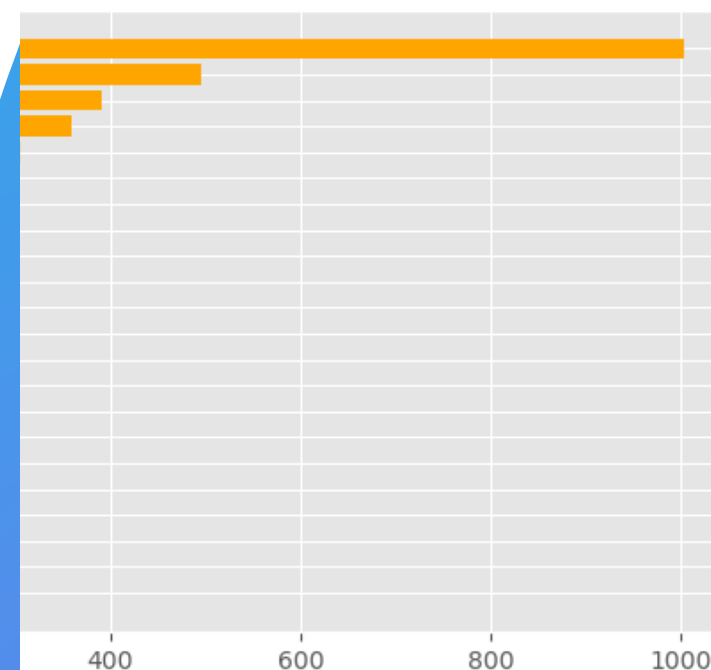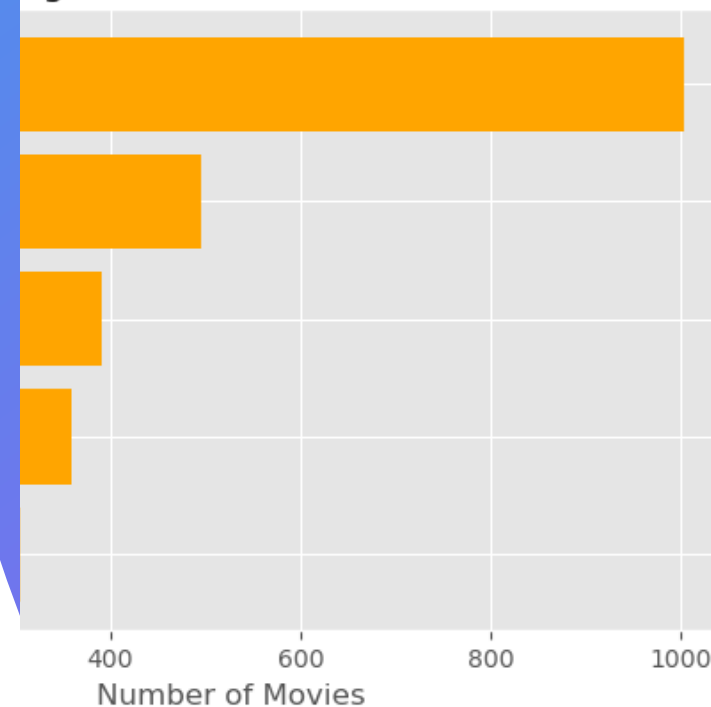

TOP 20 Succesful Genres in terms of Income Ratio
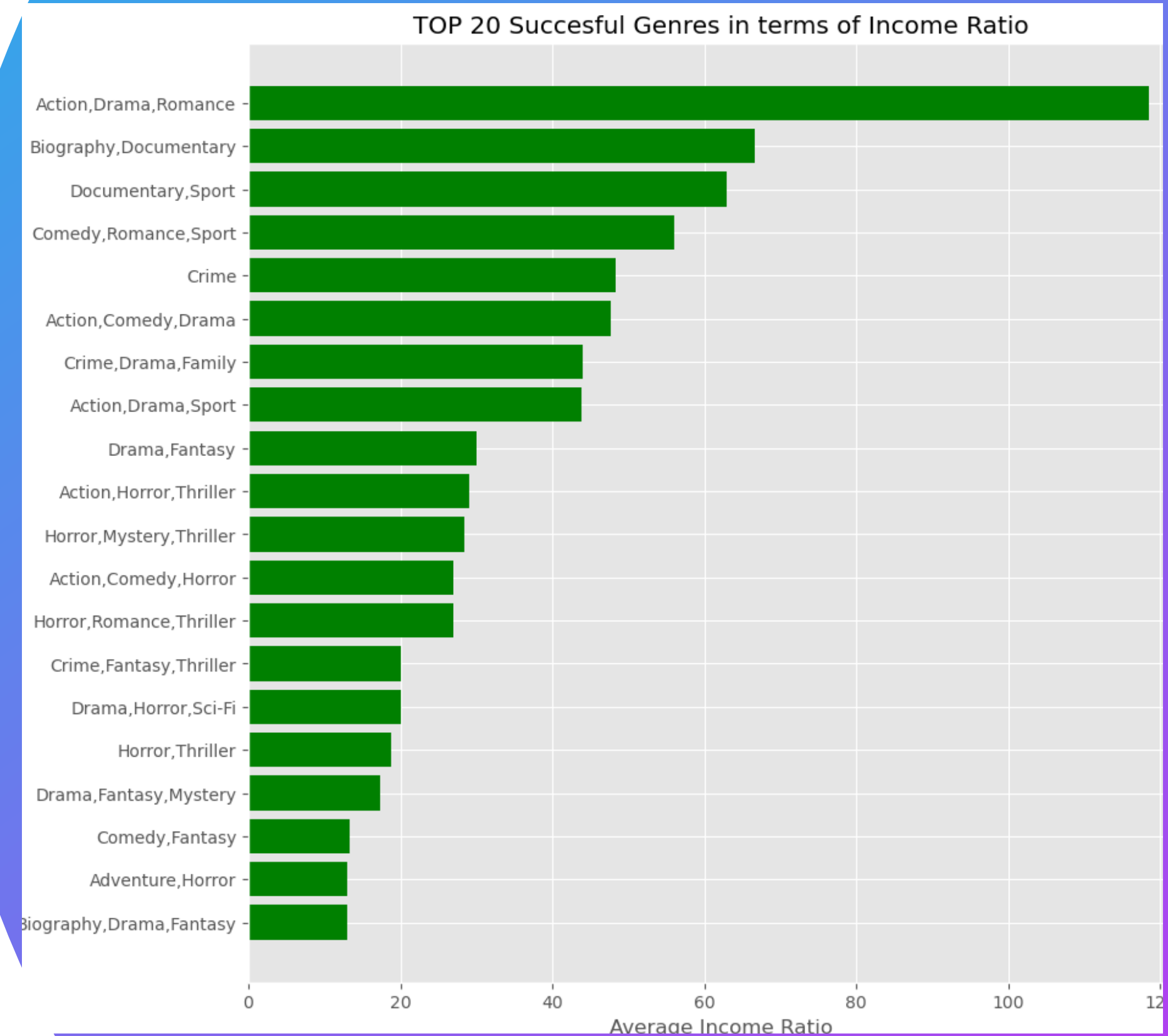
DATA ANALYSIS

Most common genres in the Top 100 and the successful movies and top successful genres
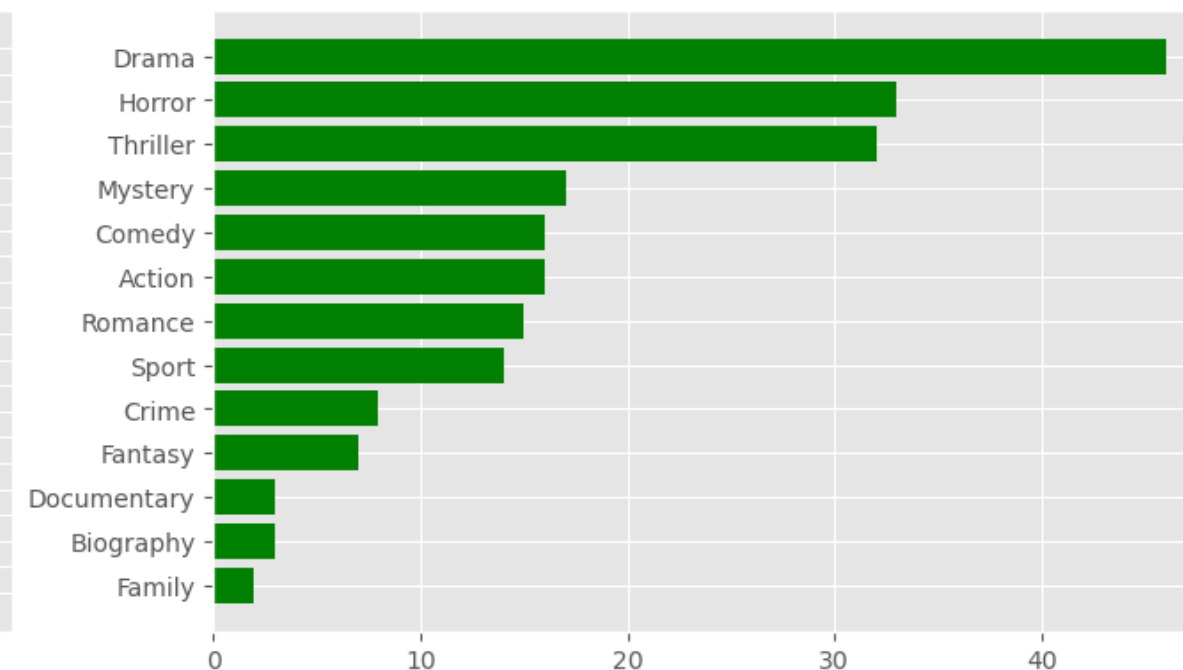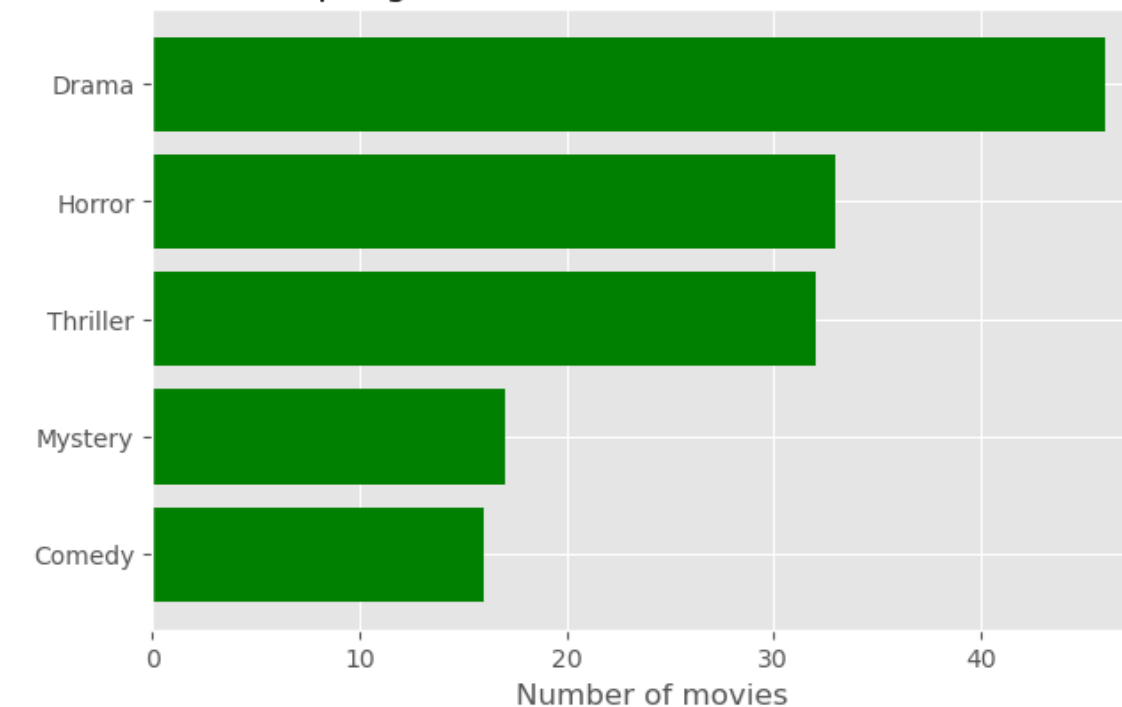
DATA ANALYSIS

From the observations, it can be inferred that while there is no specific standard required a number of votes for determining the average rating of movies, a higher number of votes tends to have a slight positive effect on the average rating.

This suggests that movies with a larger number of votes are more likely to have a higher average rating, indicating a positive reception from viewers.

Correlation between Number of votes and Average Rating
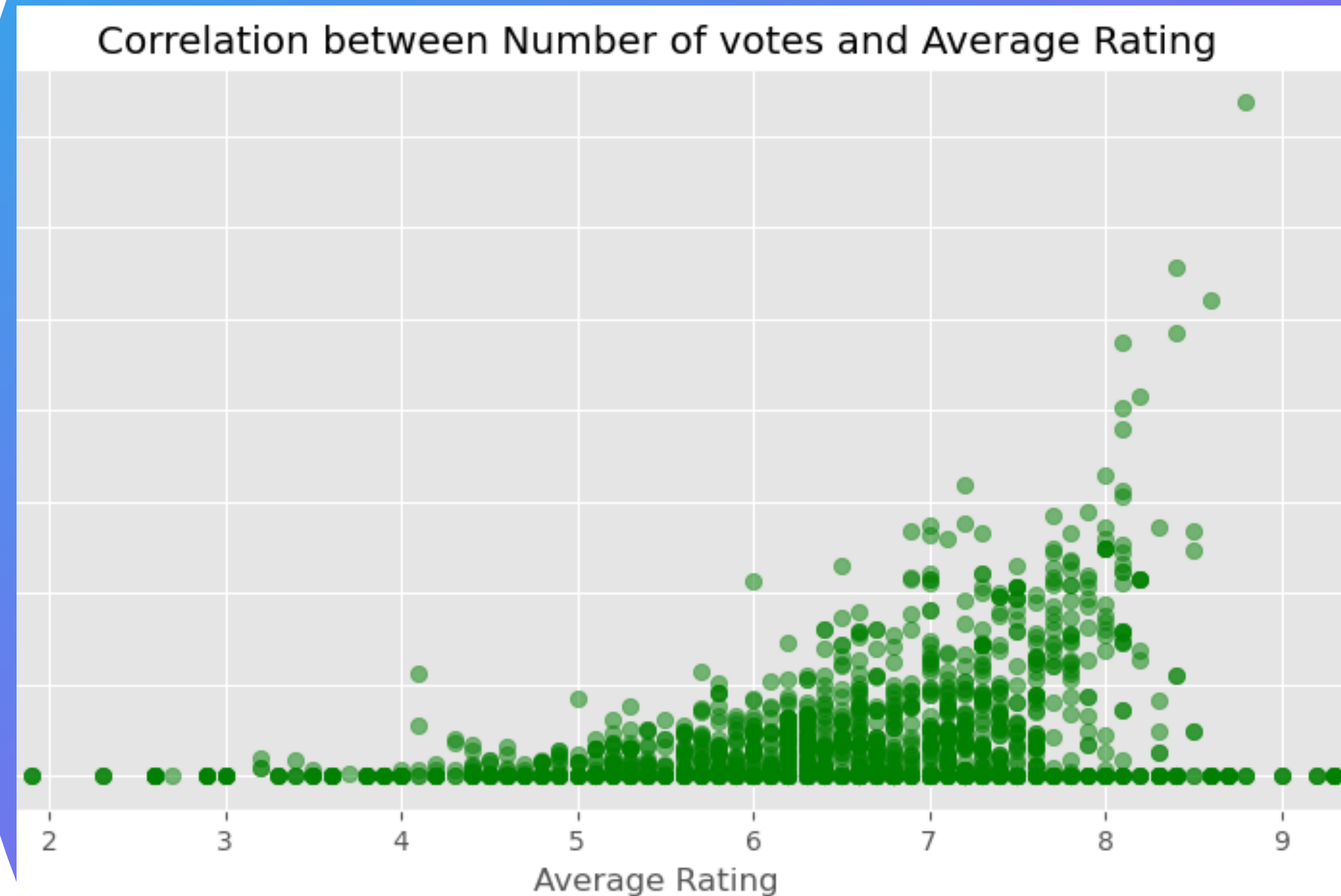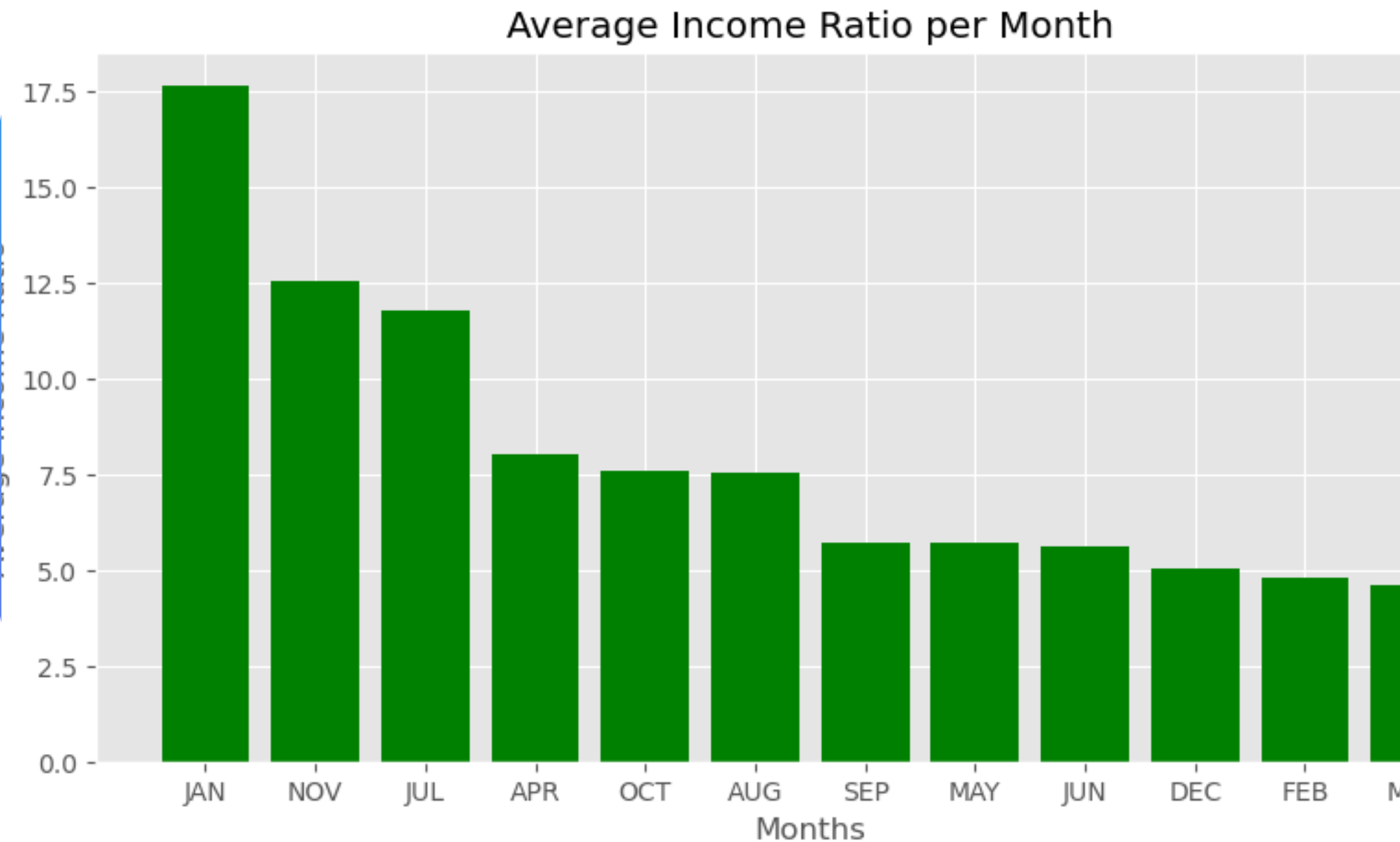
Average Rating

DATA ANALYSIS

Based on the analysis of the average income ratio for movies released in different months

It has been determined that the best months to release a movie, in terms of generating higher average income, are July, January, and November
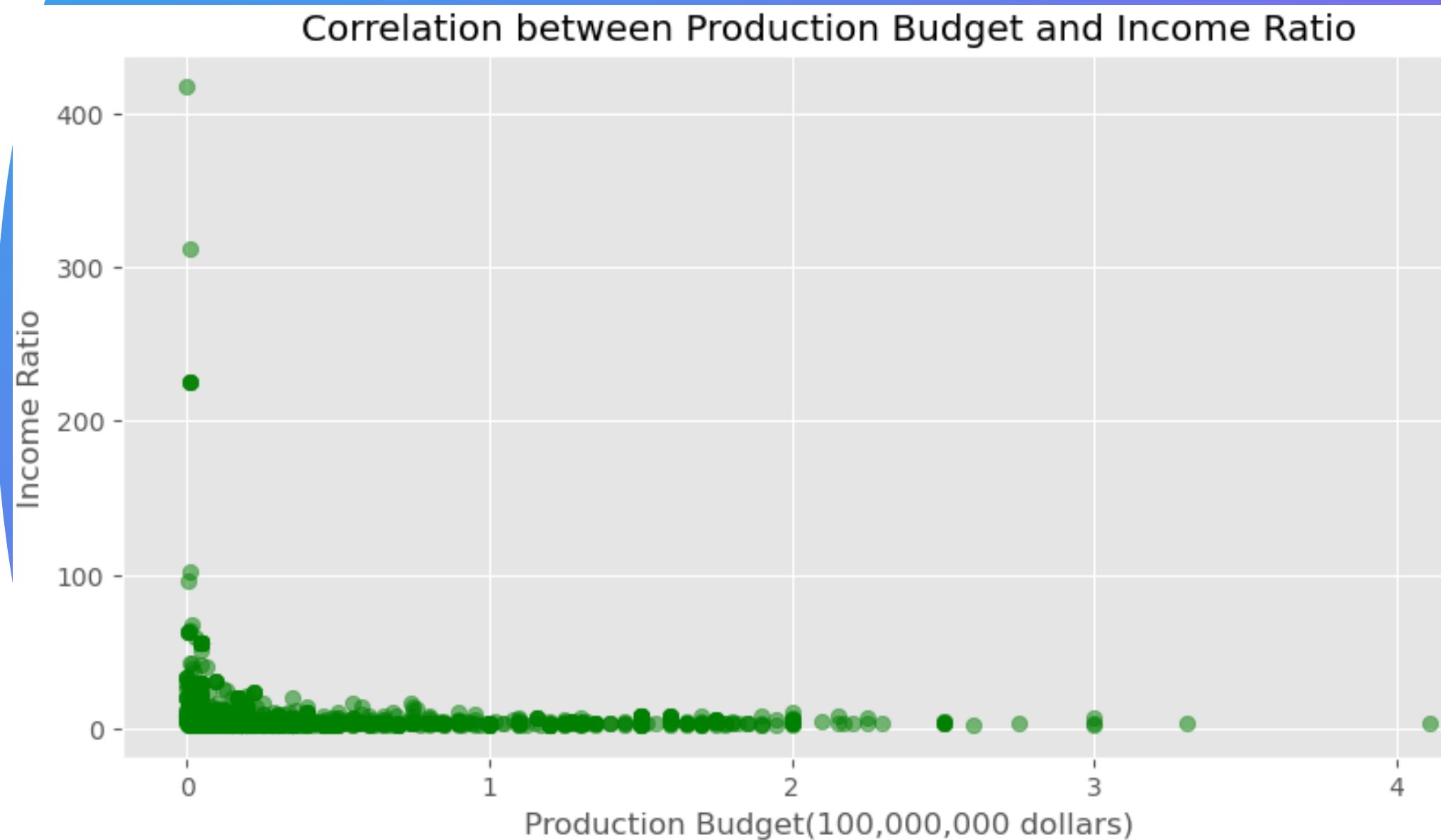

Average Income Ratio per Month

DATA ANALYSIS

The correlation coefficient indicates a low negative correlation

This means a slight inverse relationship and therefore the more the production budget the lesser the income ratio.

Correlation between Production Budget and Income Ratio

Conclusions / Recommendations

Film Genres: Most common and successful genres in movies are DRAMA, THRILLER, and COMEDY. Among these, DRAMA stands out as it has a high income ratio and is relatively cheaper to produce. Additionally, genres like ACTION, ADVENTURE, HORROR, and MYSTERY have also performed well and are worth considering.

Release Months: I recommend focusing on three months that have proven to be the best in terms of income generation: JULY, JANUARY, and NOVEMBER. Among these, JULY has shown the highest income ratio.

Budget Allocation: To optimize the budget allocation for a successful movie, it is recommended to allocate within 24 million dollars.