A Data Analysis of Movie Popularity, Profitability, Genres, and Director Success.

Summary

- In this data analysis project, I explored the factors influencing movie success by examining the relationships between profitability, genres, director performance, and popularity.
- Through exploratory data analysis, data collection and statistical analysis, I gained valuable insights.
- Specific genres, such as action and adventure, demonstrated a higher likelihood of yielding profitable outcomes.
- Additionally, directors with a track record of success in the industry tended to create more popular and profitable films.
- The analysis revealed valuable insights, such as the positive correlation between movie popularity and vote counts
- These findings can guide data-driven decision-making, enabling Microsoft Studios to enhance their movie offerings and appeal to their target audience, overall leading to increased success in the film industry.

Outline

Business Problem

Data Understanding

Exploring the Data

Data Preparation

Data Modelling

Evaluation

Conclusions

Business Problem

- The business problem is to enhance Microsoft's movie selection and marketing strategies to optimize box office revenue.
- Analyzing the influence of movie genre, director, and ratings on popularity and financial performance is a key focus.
- The findings will help inform data-driven decisions for greater success in the film industry. revenue.

Data

- The data for this project comes from three different sources: Rotten Tomatoes, The Movie DB, and Box Office Mojo.
- The rt.movies_info.csv includes the target variables are 'Genre' and 'Director', and it contains both categorical and numerical data.
- The tmdb.movies.csv contains the target variables are 'popularity', 'release_date', and 'vote_counts', and it includes categorical, numerical, and time-based variables.
- The bom.movies_gross.csv from Box Office Mojo includes details such the target variable is 'domestic_gross', and it contains both categorical and numerical data.

Methods

The primary method used was Exploratory Data Analysis (EDA).

EDA provided valuable information for addressing the business problem and guiding decision-making in the movie industry.

The process involved data cleaning, handling missing values, and merging relevant datasets.

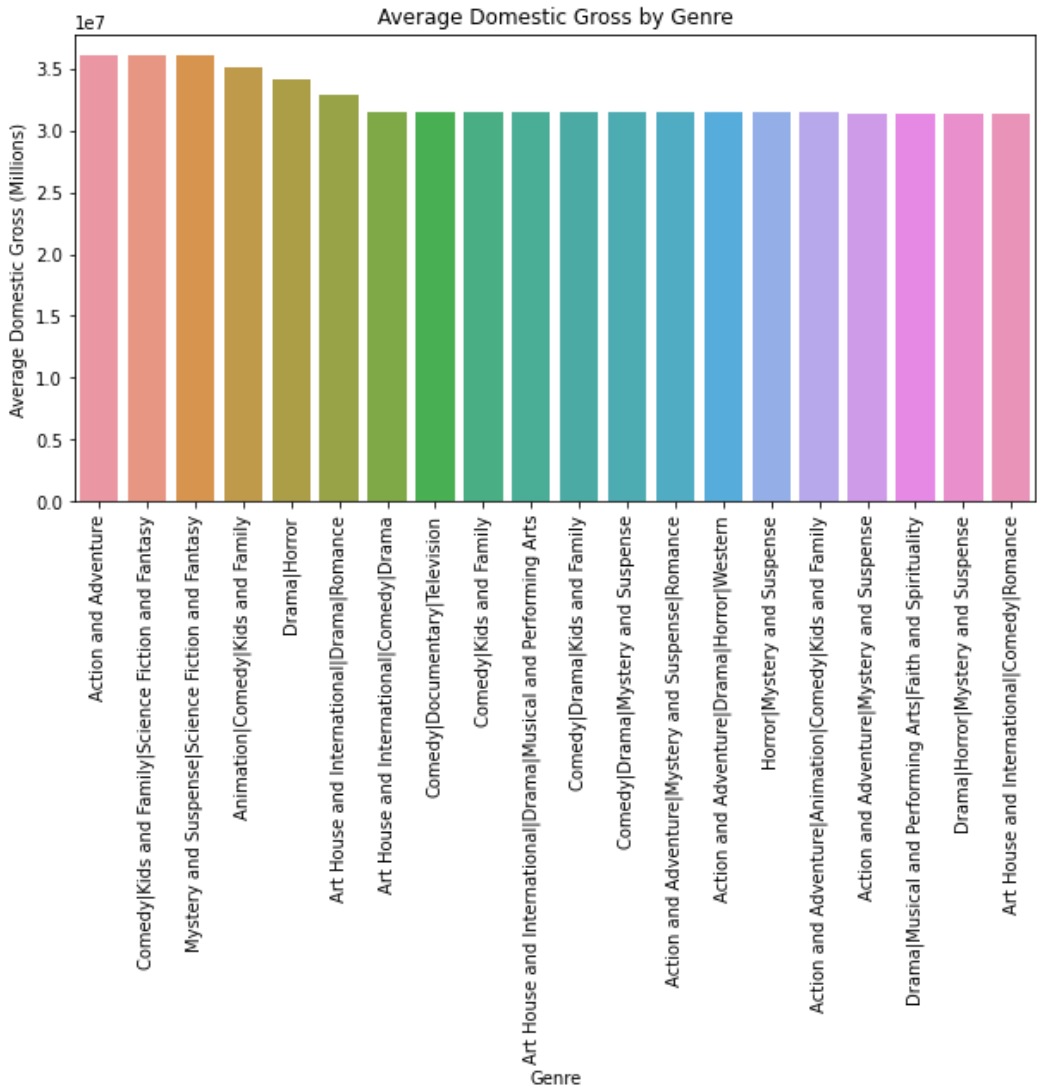
Various visualization techniques such as histograms, box plots, and bar charts were used to gain insights into movie profitability, genres, and director success.

RESULTS

The following are the key results after the analysis of the datasets to address the business problem:

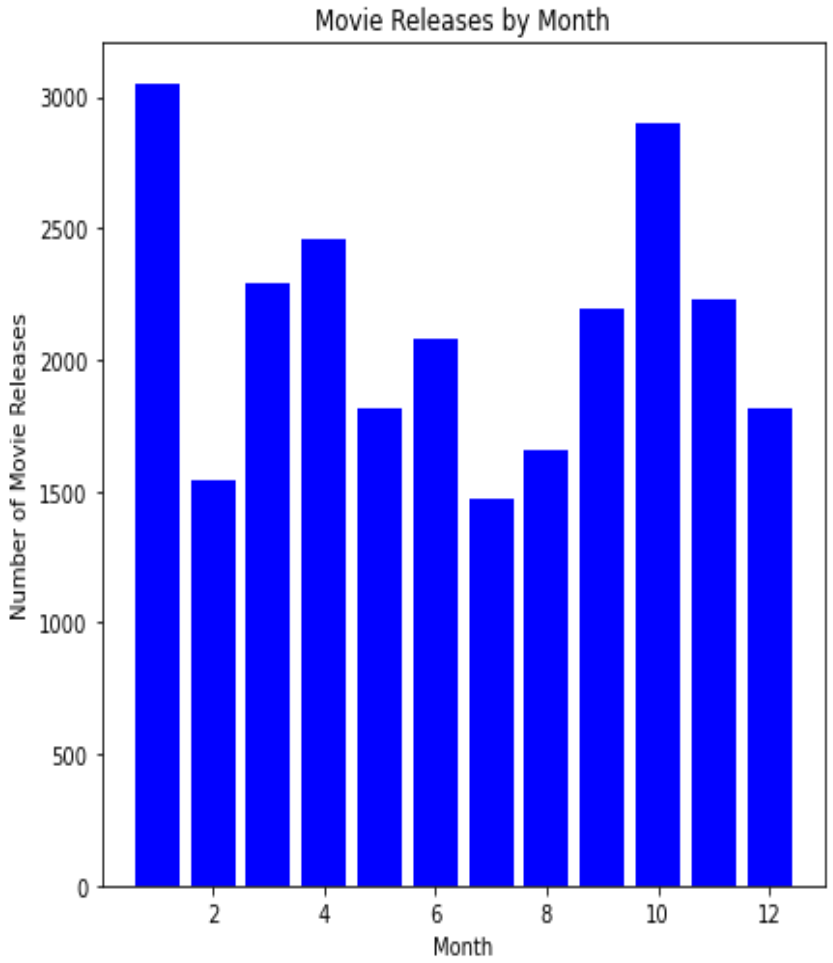
Genre Profitability:

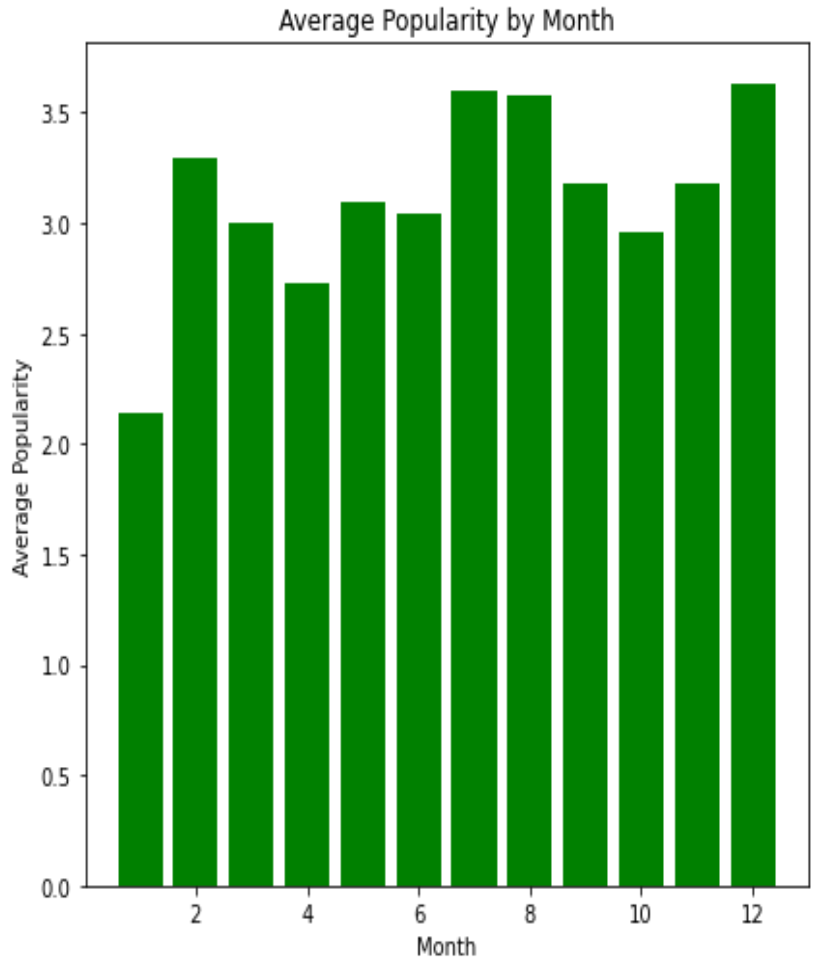
The analysis of average domestic gross by genre revealed that certain genres are more profitable than others. This information can guide the studio in focusing on genres that have higher average domestic gross, thus maximizing revenue.



Release Date and Popularity:

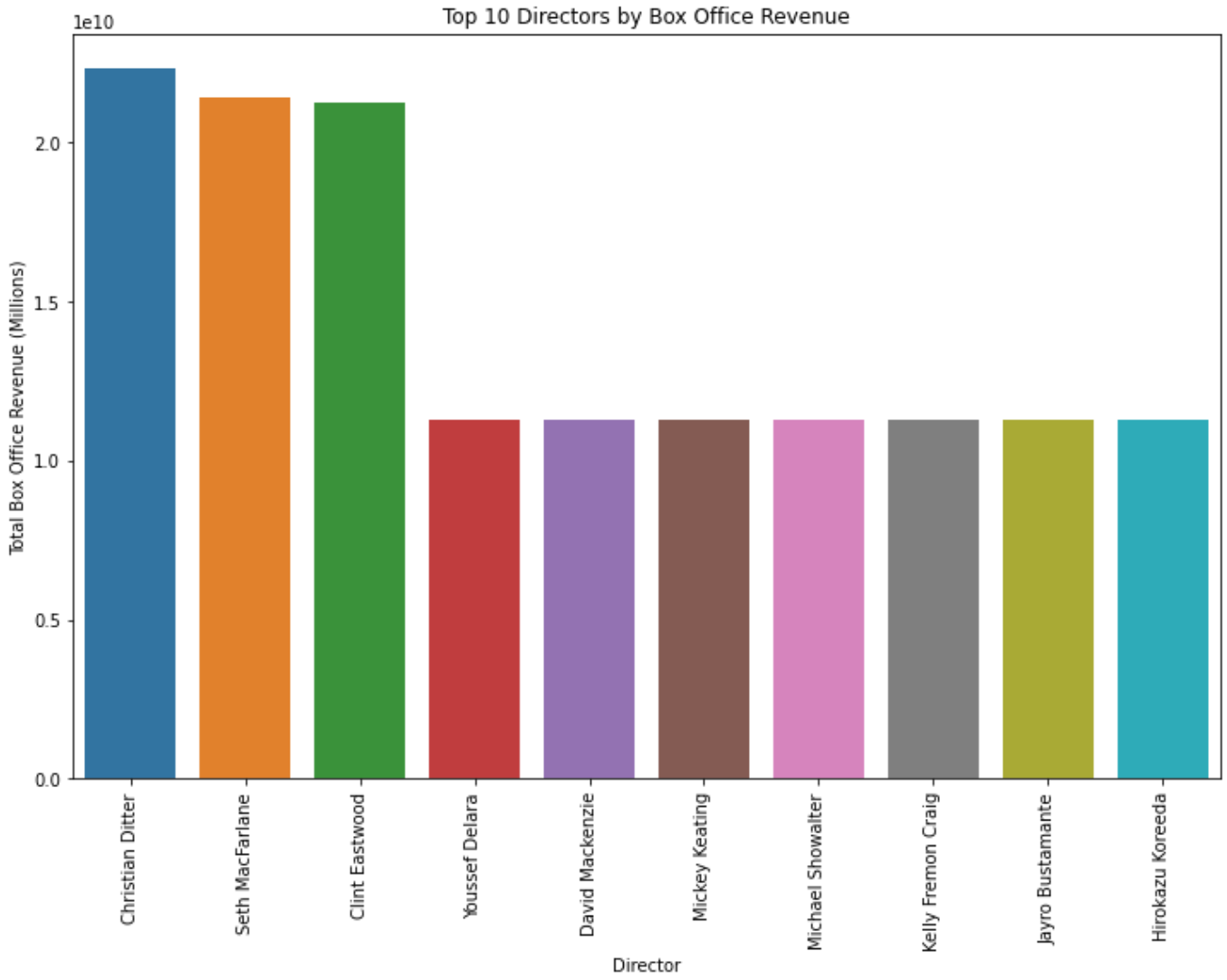
Understanding the relationship between movie popularity and release dates might help determine the best months or years to release movies. Strategic planning for the release of movies can benefit Microsoft Studio’.





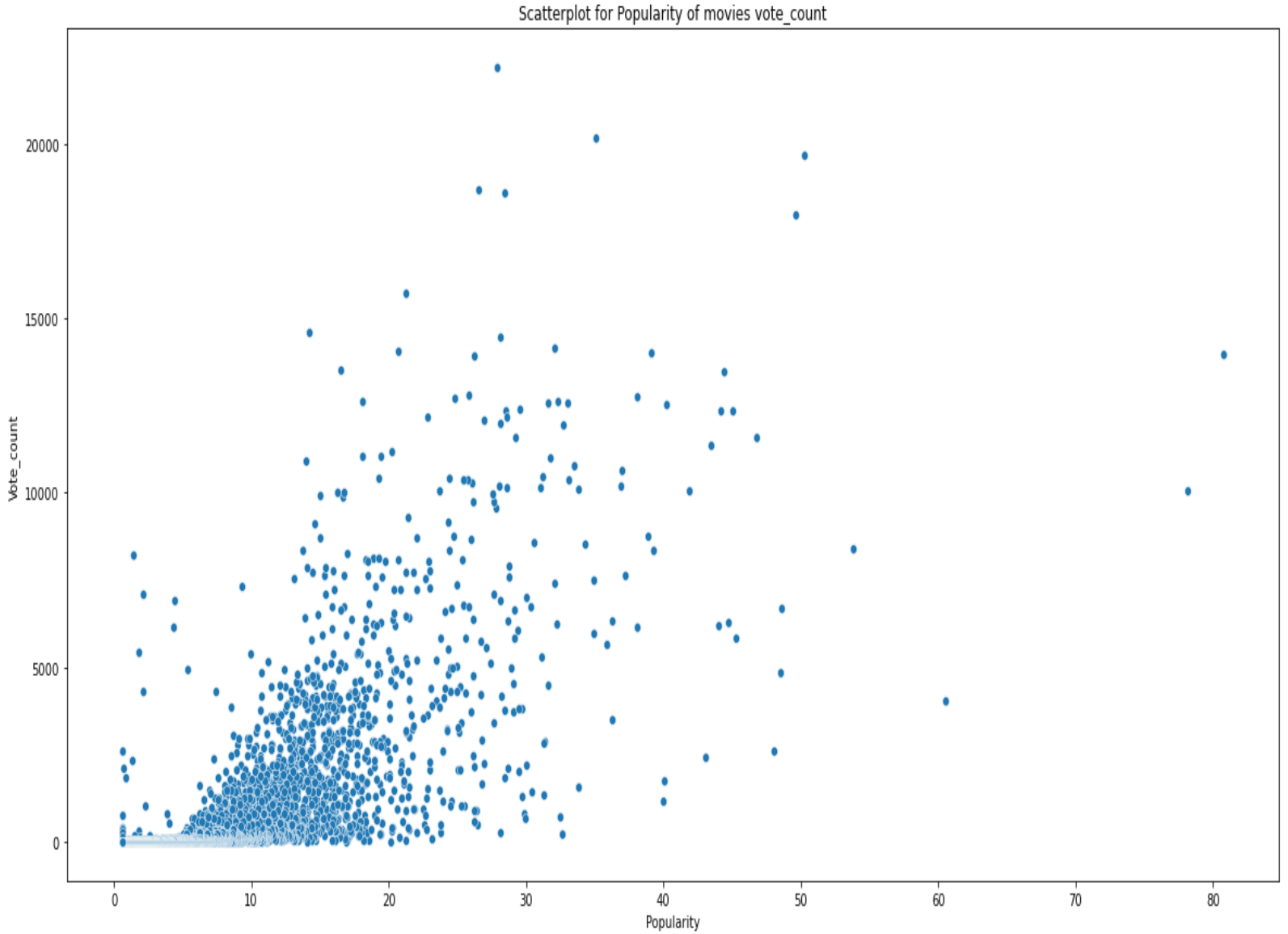
Most Successful Directors:

By analyzing the data on domestic gross revenue for each director, I identified the top 10 most successful directors in terms of box office performance. Collaborating more frequently with these directors could increase the likelihood of producing successful movies.



Positive Correlation:

The analysis indicates a positive correlation between movie popularity and vote counts. Movies that are more popular tend to receive higher vote counts, indicating that audience engagement significantly influences the number of votes a movie receives



Conclusions

- In conclusion, the data analysis project has provided valuable insights for Microsoft Studios to make informed decisions in the movie industry.
- By identifying profitable movie genres and successful directors, the studio can focus on producing content that the audience like and maximizes box office revenue.
- The analysis of release dates and their impact on popularity and reception allows for planning of movie releases.
- However the limitations of the project include missing data, which may have impacted the accuracy of the analysis, and the influence of external factors which was not considered.
- Future work could also involve predictive modeling to forecast movie performance and assist in strategic decision-making and investments in potential movies.
- Having additional relevant data, such as movie budgets and marketing expenses, would provide deeper insights into factors affecting movie success.

Thank You!

EMAIL: BERNADETTE WANJIKU NGANGA

GITHUB: BERNADETTE2023