

Phase 1 Project

Jane Mwangi

Part time 4

TM: Stella Waithera





Project Overview

In this project, we will employ exploratory data analysis to derive valuable insights for Microsoft's venture into the movie production industry.



Problem statement

Microsoft wants to enter the movie industry despite lacking expertise and seeks to find insights on making of successful films.

We are tasked with researching and providing actionable insights to guide their film creation decisions and maximize profitability.



Data used

movie budgets

tmdb.Movies

title.basics

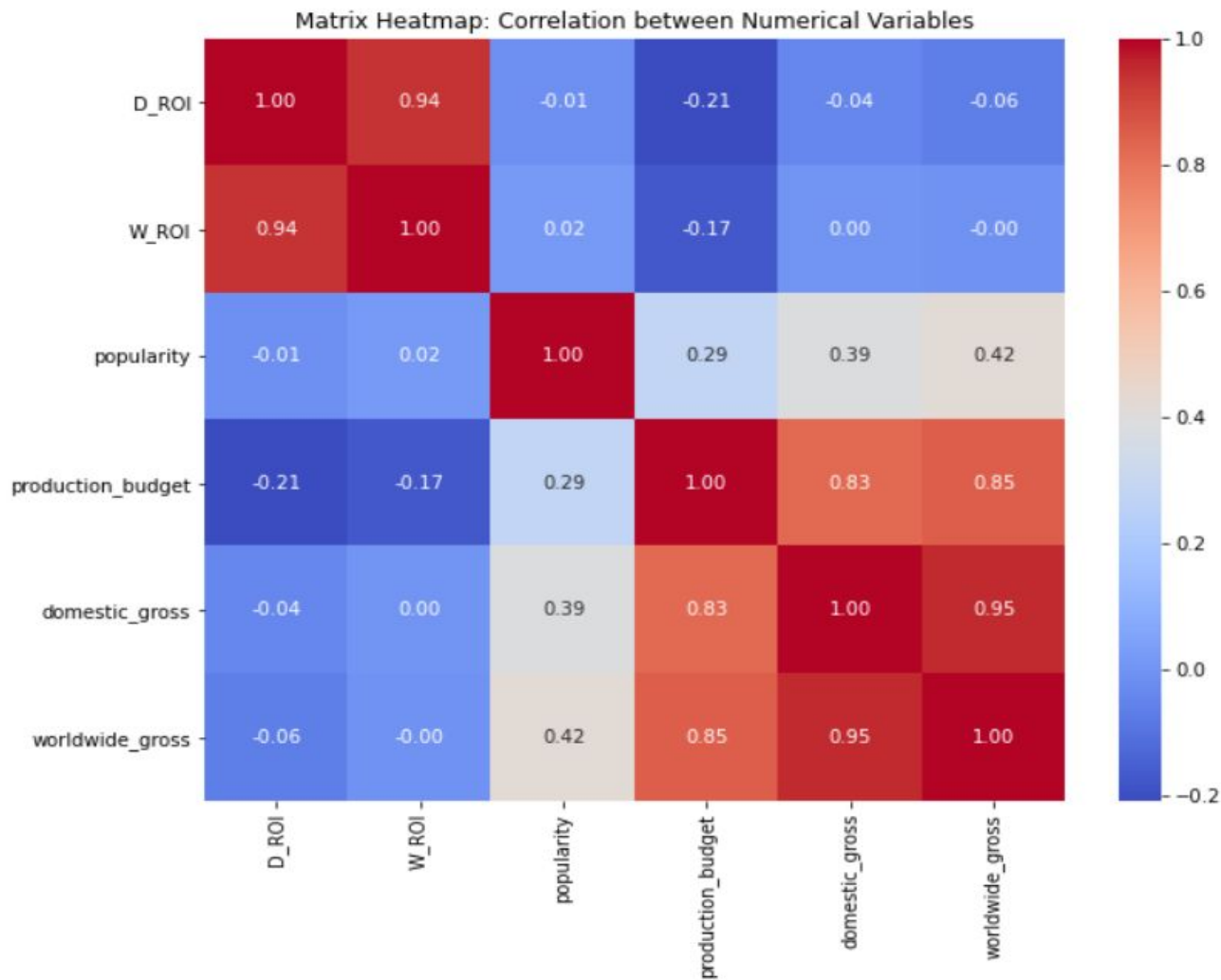


Method

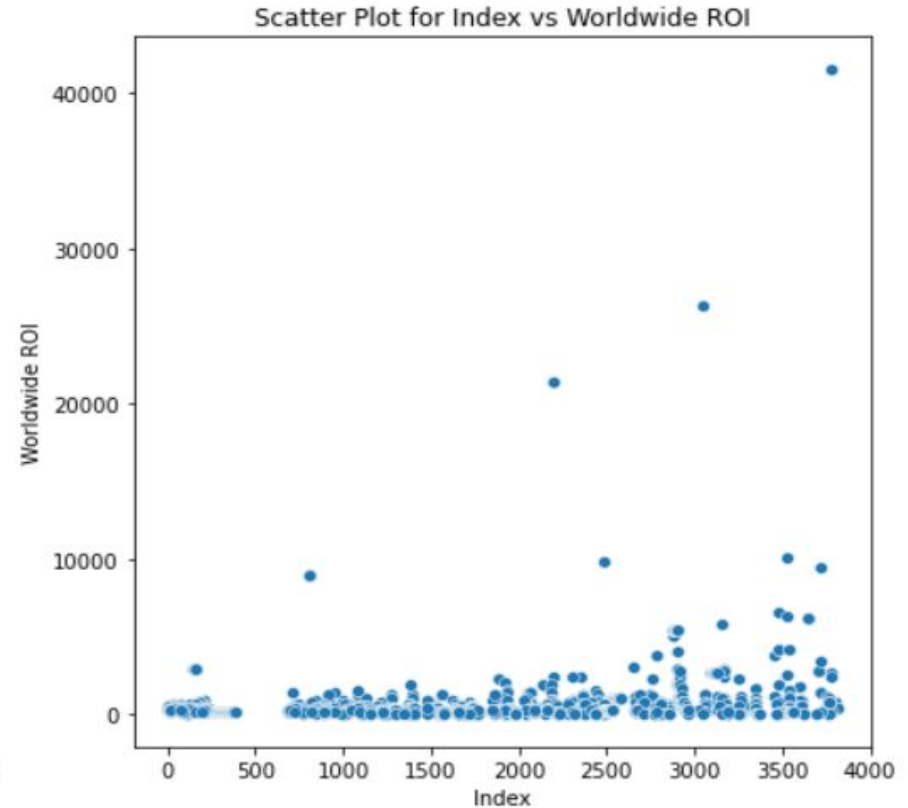
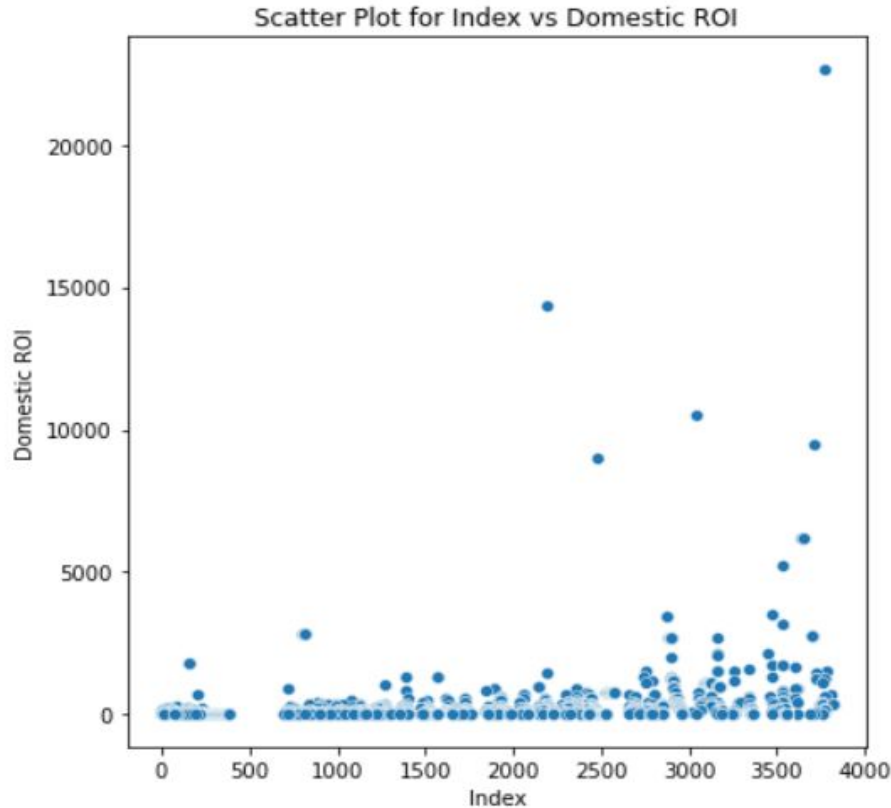
- 1.Data mining - Identifying data would answer my business question.
- 2.Loading of libraries
- 3.Data cleaning checking for missing values, duplicates, outliers, valid labels and use of up-to-date data to ensure maximum value of the results
- 4.Data Exploration (Exploratory data Analysis) highlighting patterns and relations in data. use of descriptive statistics such as mean.
- 5.Data visualization by creating plots such as histograms, scatter plots and line graphs identify trends and make insights.

Heatmap

correlation
between
selected
numerical
columns in
the
DataFrame

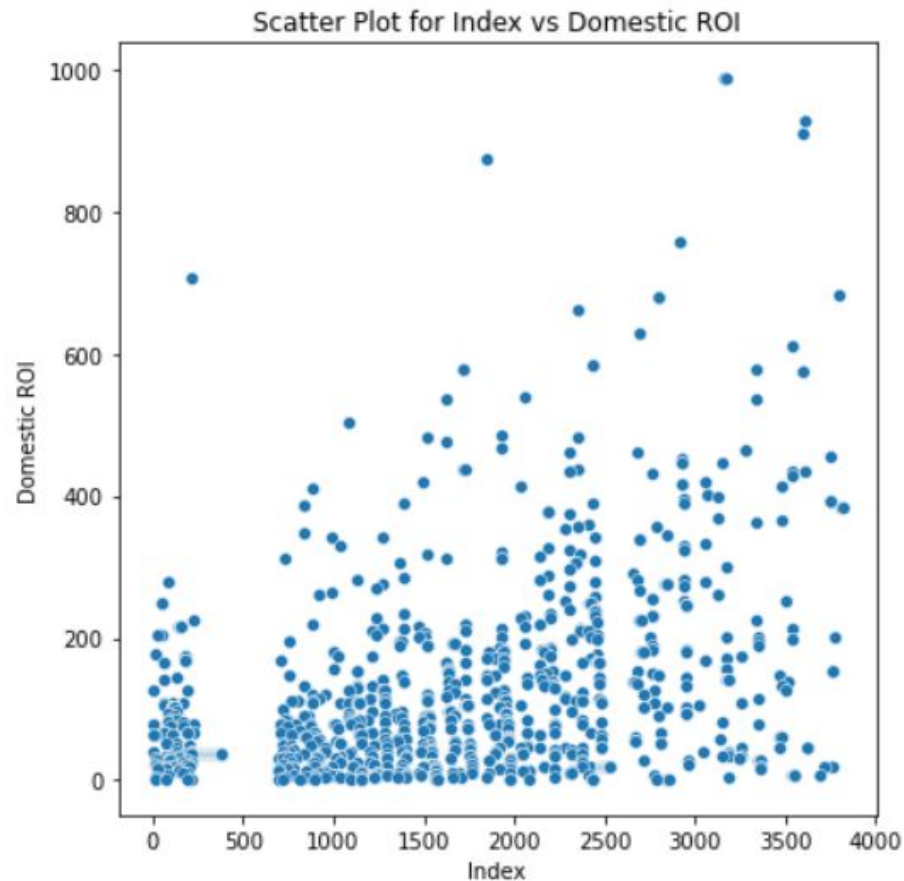


Scatter plots show a side-by-side comparison of the Domestic ROI (D_ROI) and Worldwide ROI (W_ROI) against the index values. We remove outliers to better visualise.

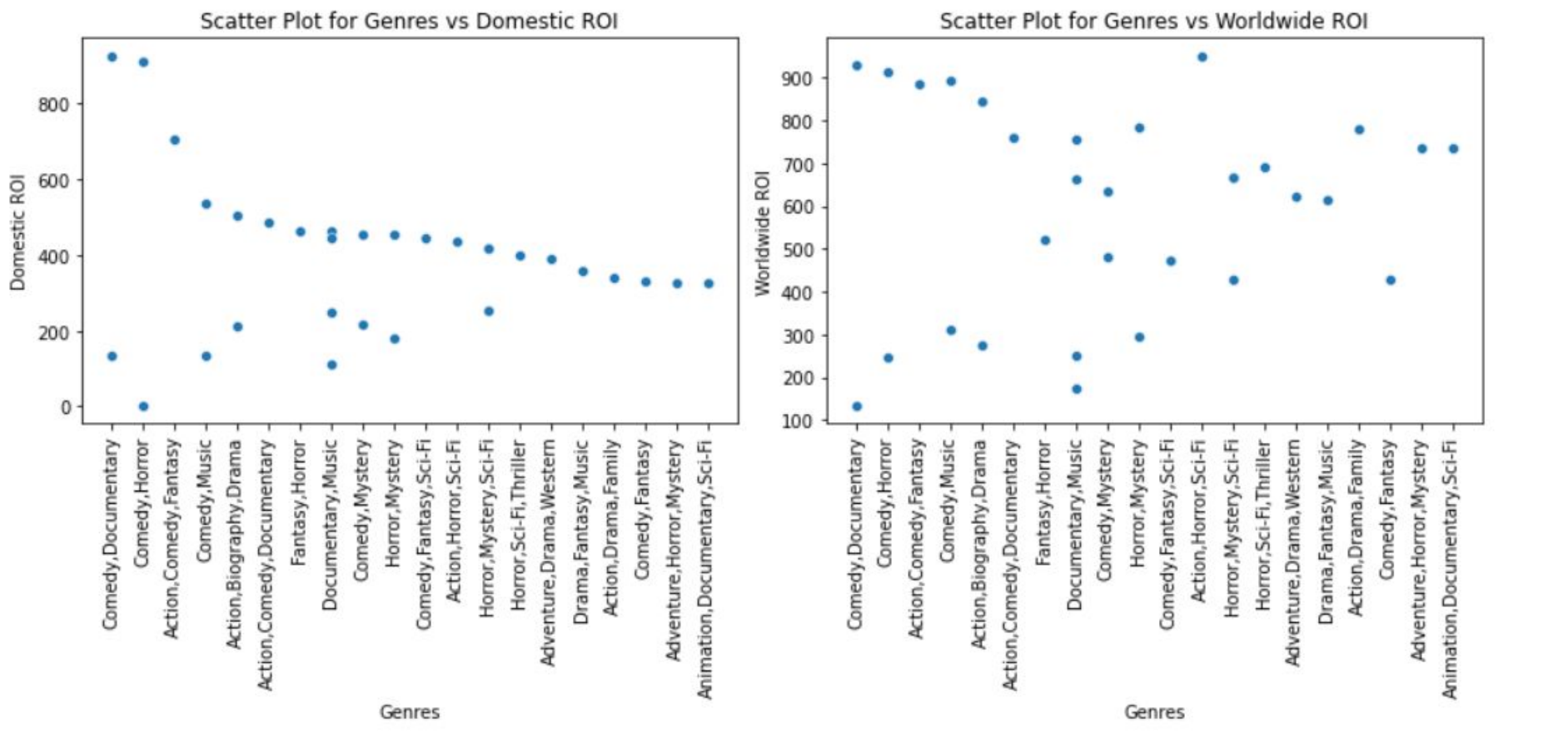


majority of our domestic ROI falls within the range of 0 to 200.

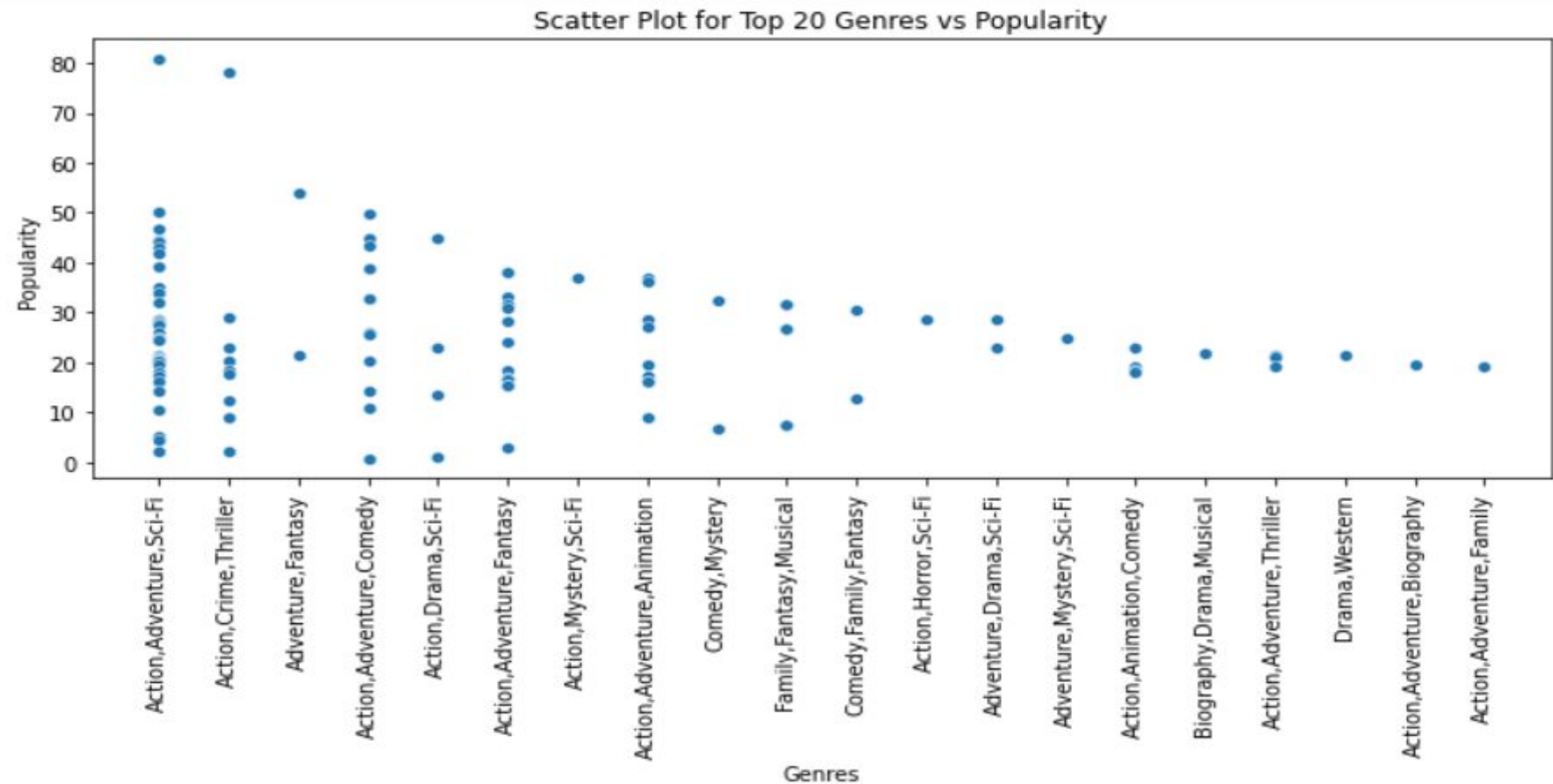
worldwide ROI shows a wide scatter. It is highly dispersed and not confined to a specific range.



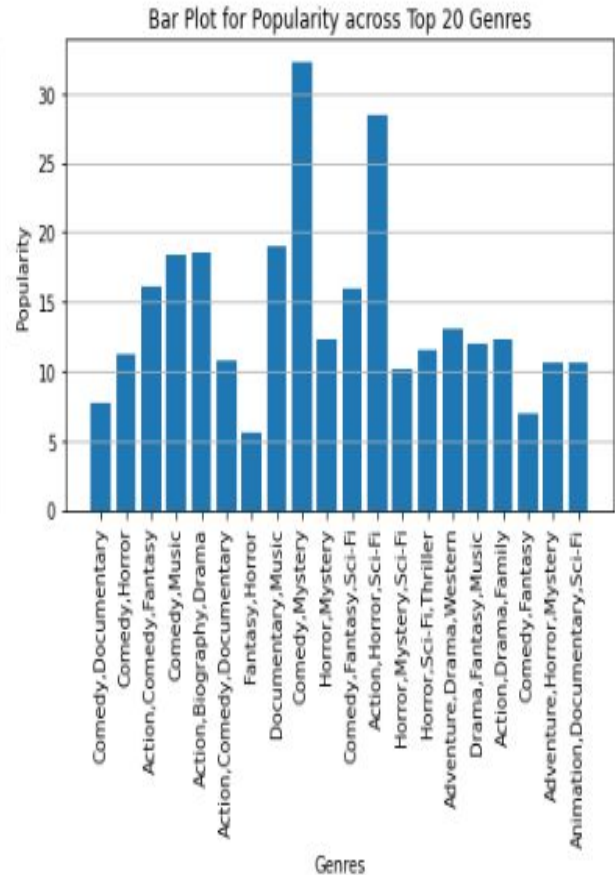
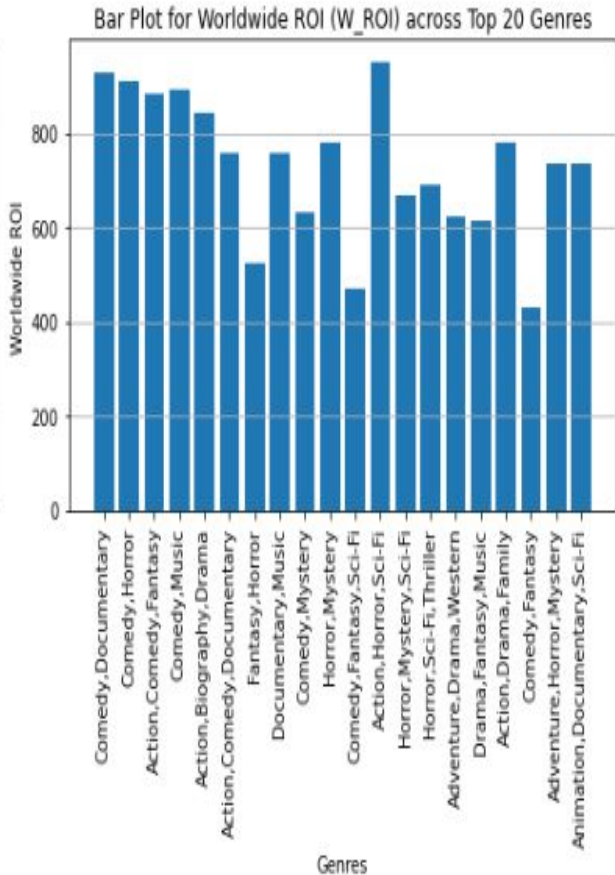
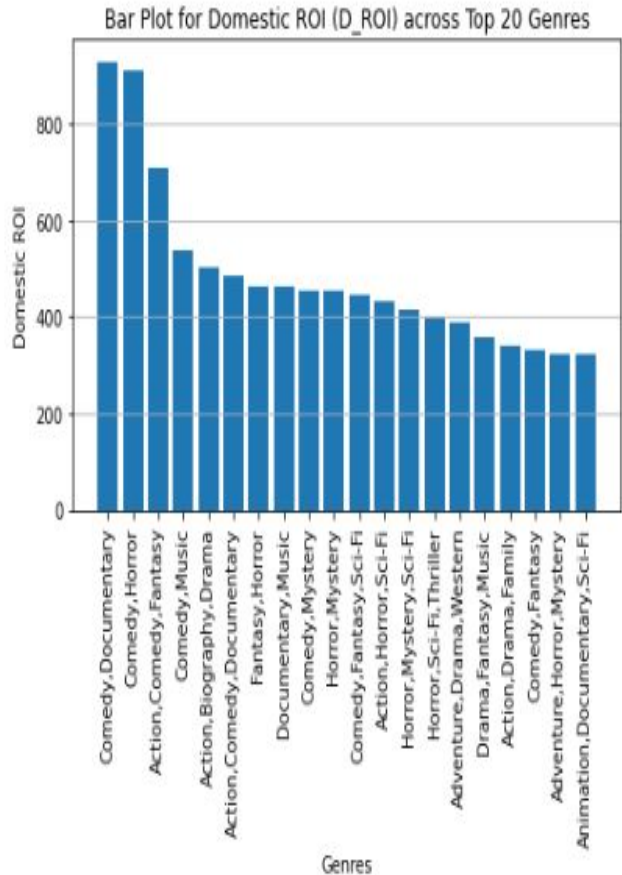
Scatter plots show a side-by-side comparison of the Domestic ROI (D_ROI) and Worldwide ROI (W_ROI) against the index values



Scatter plot depicts the relationship between popularity and genres.

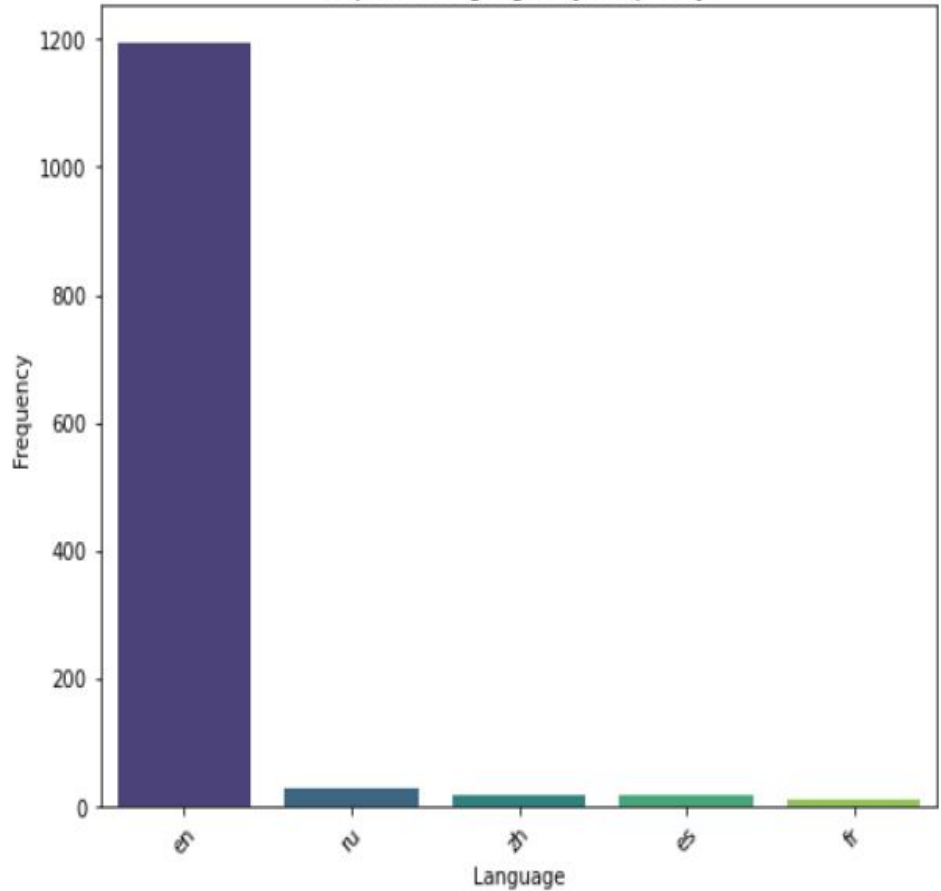


Top 20 Genres vs. Domestic , Worldwide ROI and Popularity

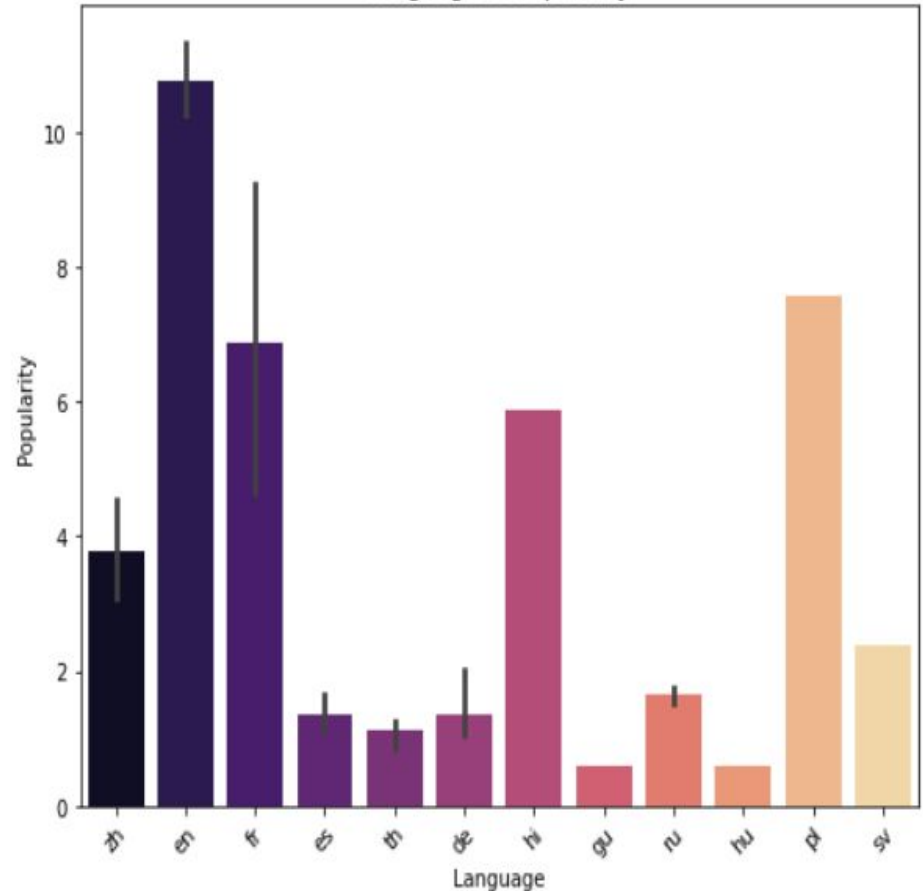


Languages vs. Frequency & Popularity

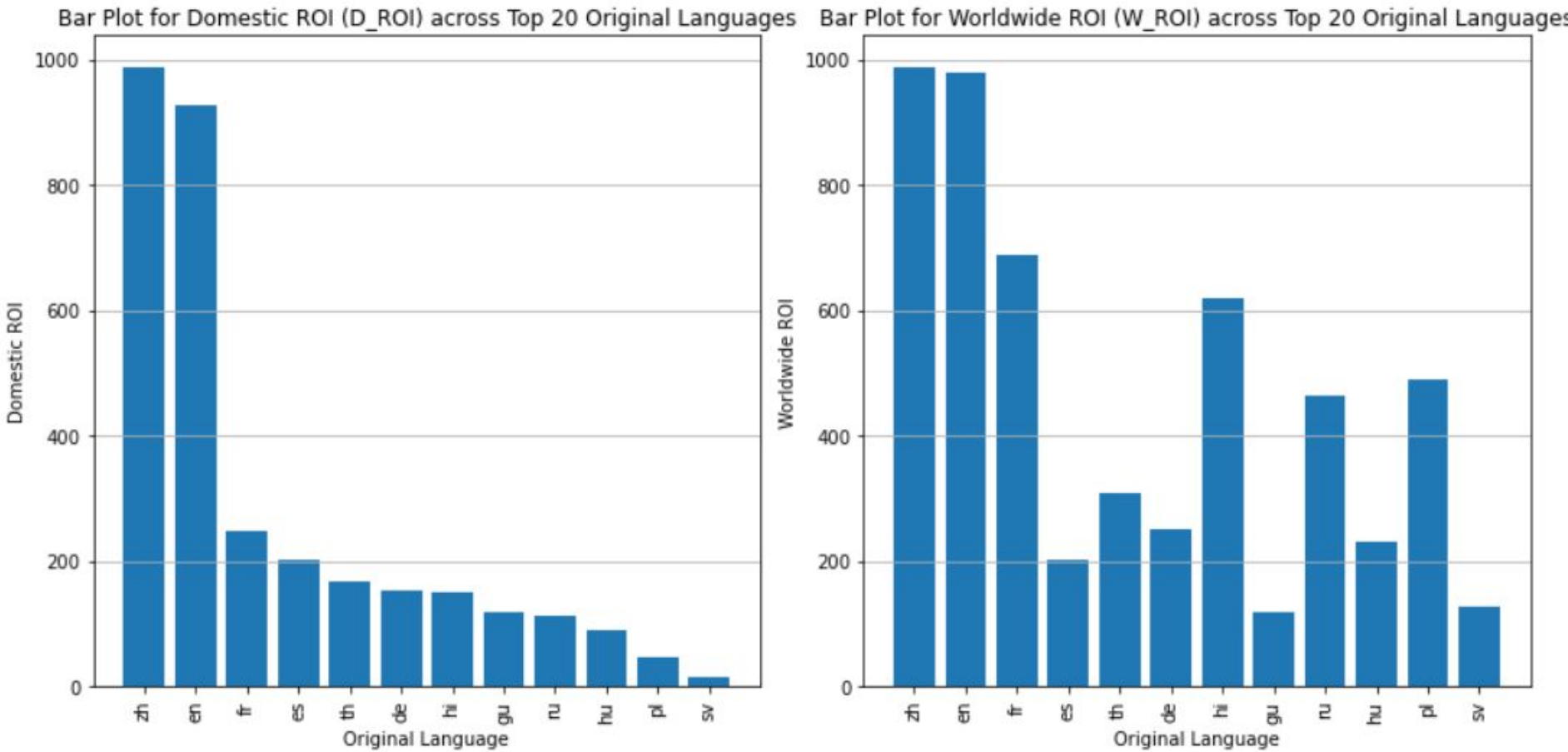
Top Five Languages by Frequency



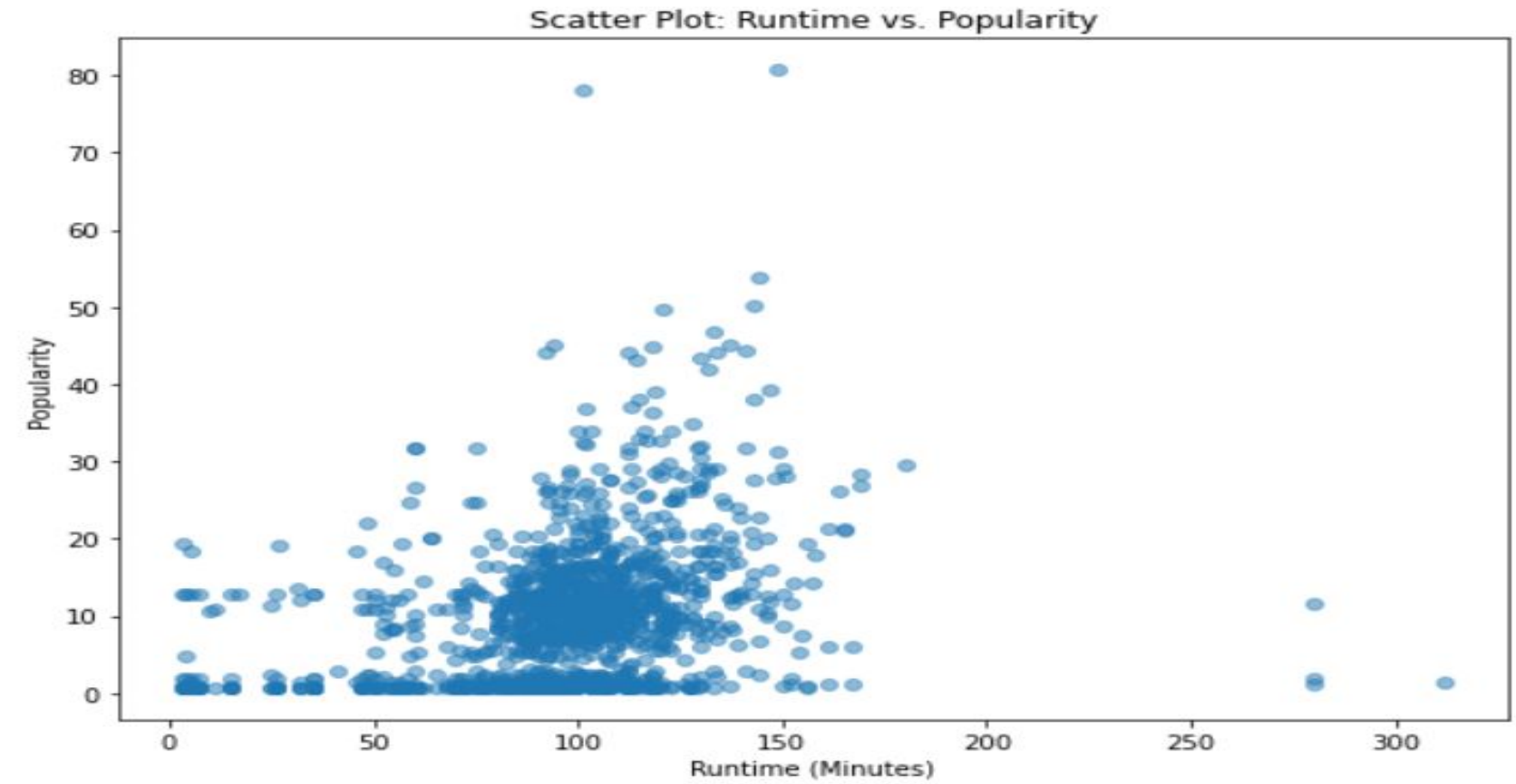
Language vs Popularity



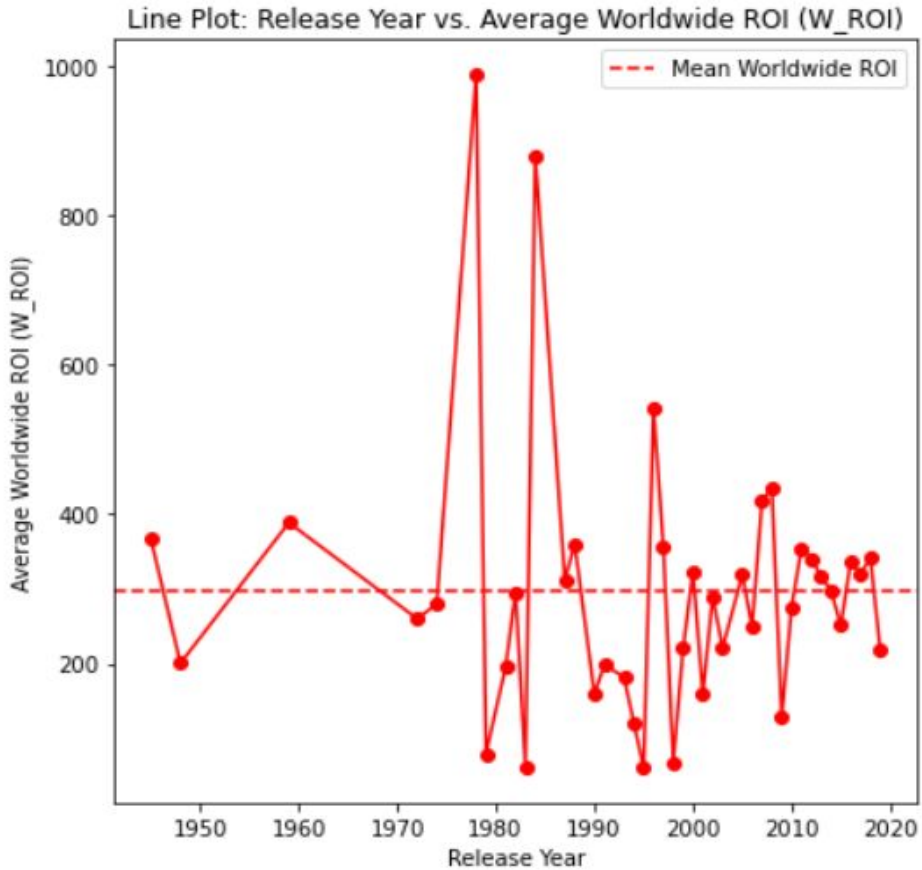
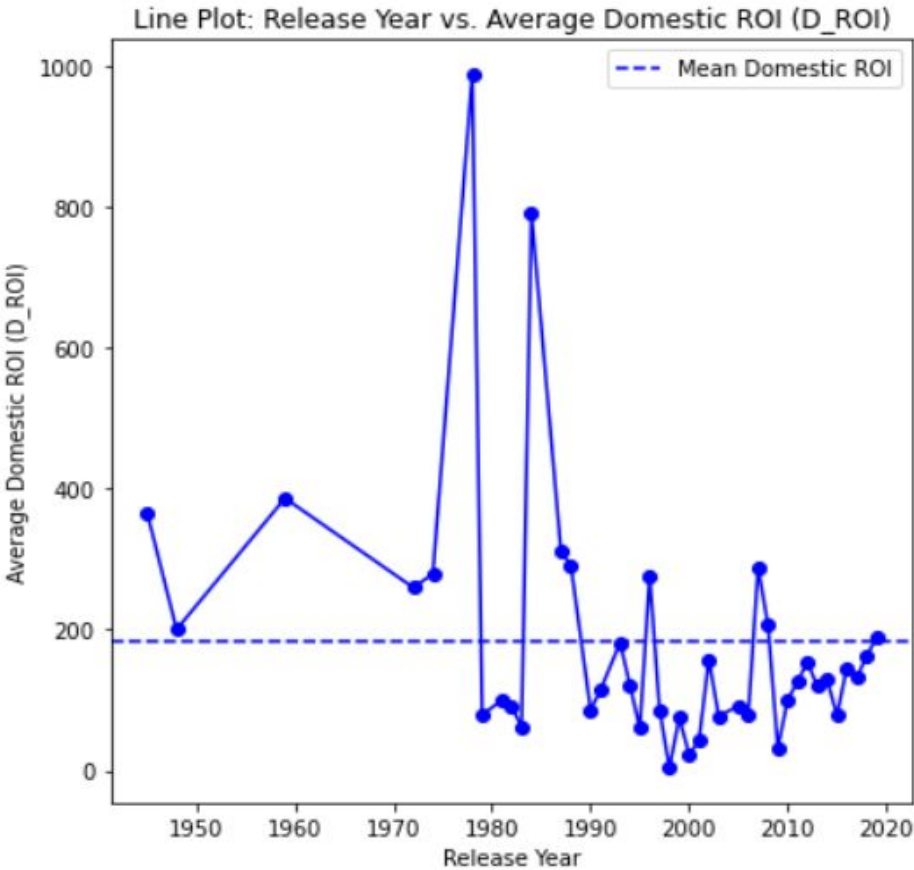
Top 20 Original Languages vs. Domestic and Worldwide ROI




Runtime vs. Popularity



Release Year vs. Average ROI



Recommendations

- 
- Focus on Genres with High Domestic ROI: Concentrate on genres like 'Comedy, Documentary,' 'Comedy, Horror,' and 'Action, Comedy, Fantasy' that have historically shown high performance in the domestic market.
 - Explore Genres with High Worldwide ROI: Consider genres like 'Action, Horror, Sci-fi' that have demonstrated exceptional worldwide performance for potential global success.
 - Leverage Popularity of Action Movies: Incorporate action elements into various genres to capitalize on the high popularity of action films.
 - Language Selection: English and Chinese movies have consistently achieved high ROI worldwide. Expanding into the Chinese market can enhance global reach.
 - Caution with Production Budget: Be mindful of budget allocation, as higher production costs do not guarantee a high ROI. Prioritize engaging storytelling and effective marketing.
 - Balance Domestic and Worldwide Focus: As the domestic market's average ROI declines, maintain a balanced approach between domestic and international film productions to ensure overall success and diversity in the movie portfolio.

Conclusion



The analysis of the movie industry data offers valuable insights for Microsoft's new movie studio. By identifying genres with high domestic and worldwide ROI, the studio can strategically focus on producing successful films for both markets. Action movies, being popular, present promising investment opportunities, and exploring genres like 'Action, Horror, Sci-fi' can lead to global success.

However, caution is necessary regarding the relationship between production budgets and ROI, as other factors influence a movie's financial success. Understanding market trends and making informed decisions based on historical data can increase the chances of creating profitable films.

Acknowledging the declining domestic ROI trend while recognizing higher worldwide ROI, the studio should balance domestic and international film releases for optimal success in the competitive movie industry. A data-driven approach and understanding of market dynamics can position Microsoft's movie studio for positive impact and growth in the entertainment landscape.