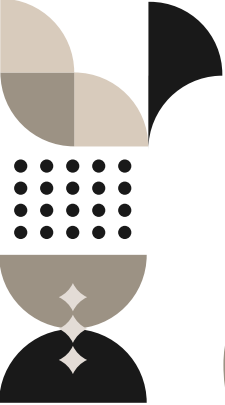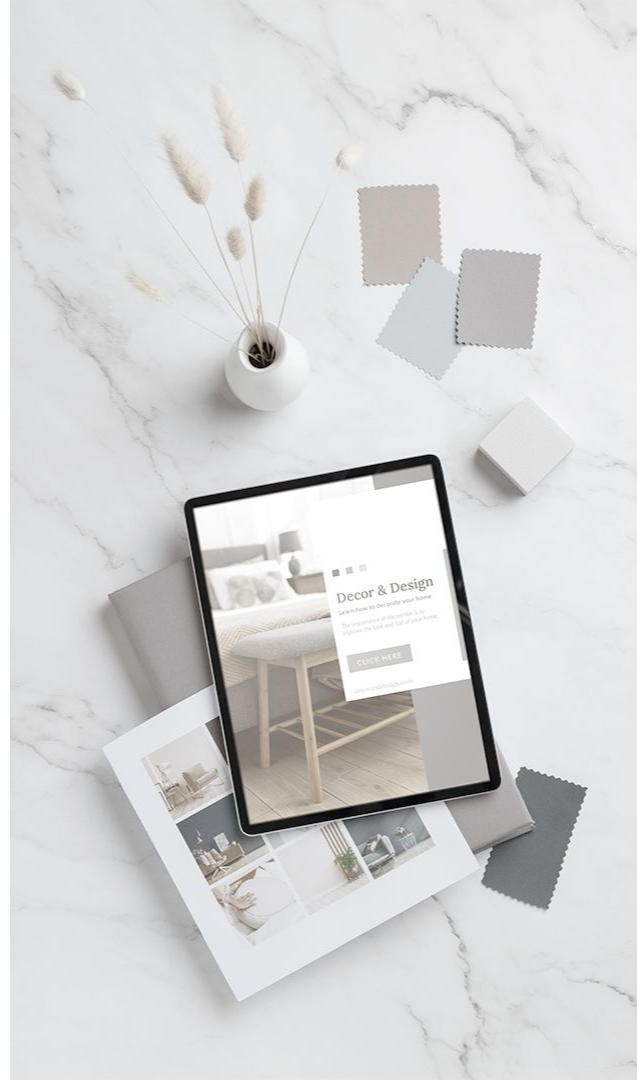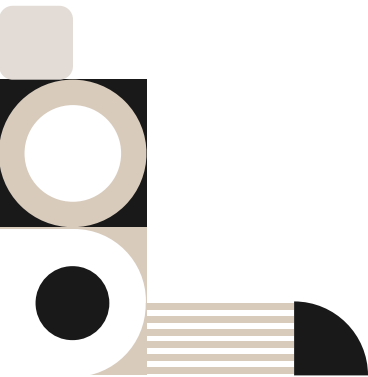# Data Analysis for Microsoft's New Movie Studio
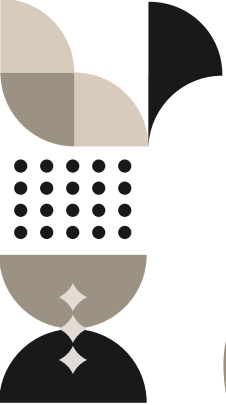
01

Overview

Overview

The project for Microsoft's new movie studio aims to get insights from various movie datasets to inform strategic decision-making.
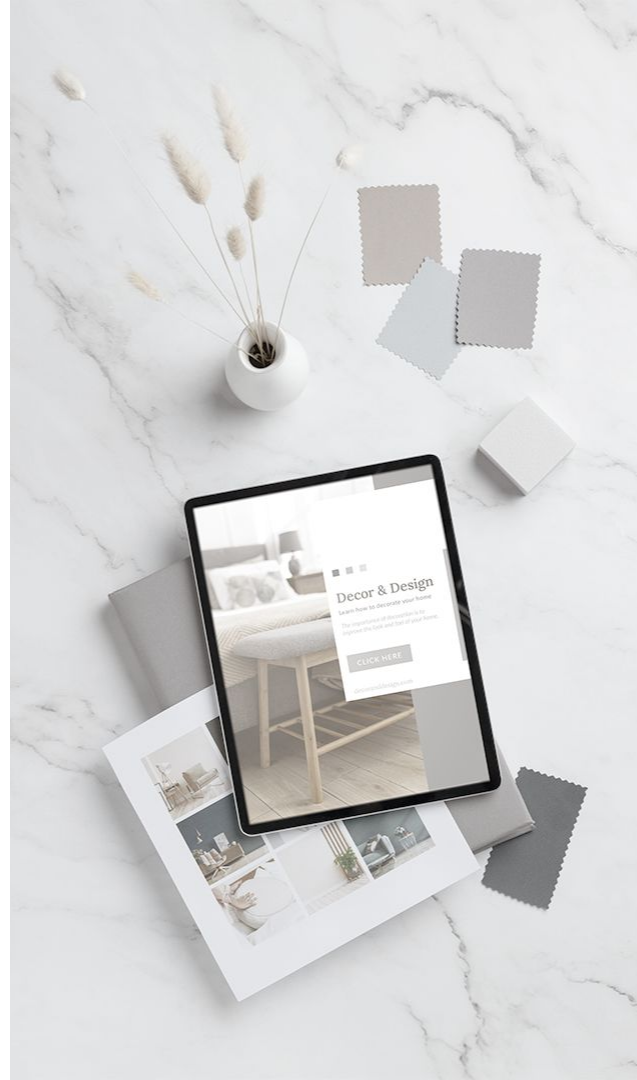
Microsoft seeks to enter the film industry by understanding the types of films currently succeeding at the box office. This project will delve into datasets sourced from platforms like Box Office Mojo, IMDB and TheMovieDB.

By conducting thorough analysis, we aim to uncover trends, patterns, and preferences within the movie industry to provide actionable recommendations for Microsoft's movie studio.

02

Business Understanding

Project Objective

The primary objective of this project is to equip Microsoft's new movie studio with comprehensive insights into the film industry.

By conducting thorough data analysis, we aim to provide recommendations that will enable Microsoft to make well-informed decisions regarding the types of films to produce.

Business Problem

While Microsoft has expertise in a variety of technology disciplines, it lacks firsthand familiarity with the film industry's many complexities.

Recognizing the growing demand for original video content and the success stories of other major companies in this space, Microsoft sees an opportunity to use its resources and capabilities to stake out a place in the entertainment industry.

Business Questions

- **What movie genres are currently dominating the box office?**

- **What trends can be detected within the dominating movie genres?**

- **What are the primary factors influencing the success of recent film releases?**

- **How do audience demographics affect movie tastes and box office performance?**

- **Are there any developing trends or possibilities in the film business that Microsoft can leverage on?**
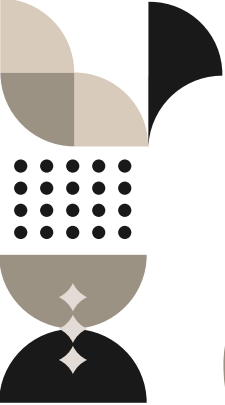
03

Data Understanding
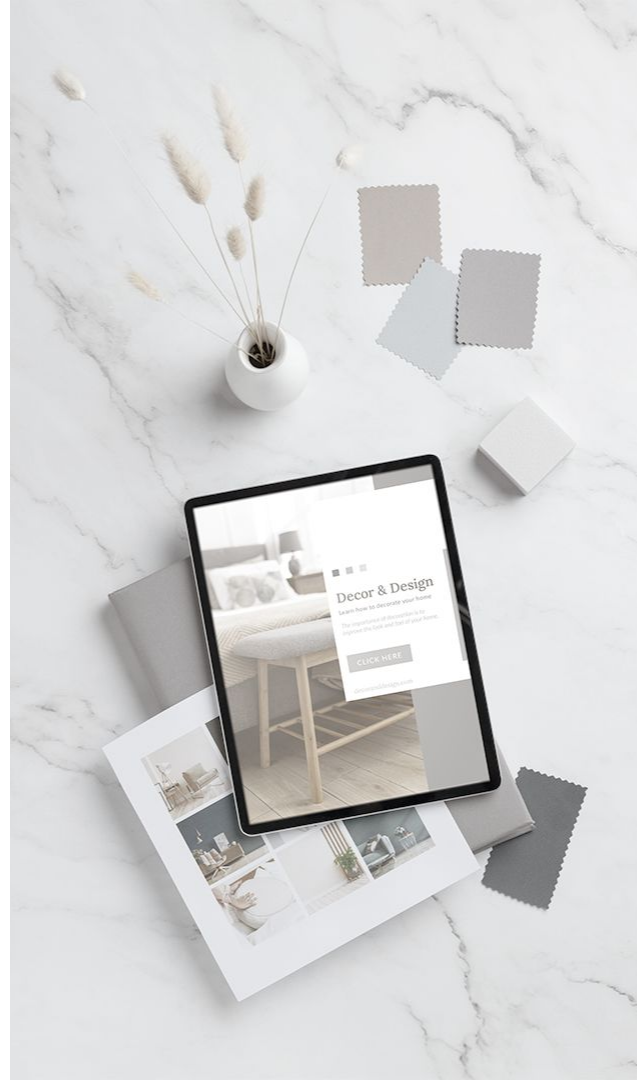
Box Office Mojo Dataset

This dataset contains 3387 entries (movies) and 5 columns. Each row represents a movie, and each column provides specific information about that movie.

Column names are title, studio, domestic_gross, foreign_gross and year

The dataset includes a combination of data types, including objects (strings), floats, and integers.

IMDB Dataset

The IMDB dataset consists of several tables including 'movie_basics', 'directors', 'known_for', 'movie_akas', 'movie_ratings', 'persons', 'principals', and 'writers'.

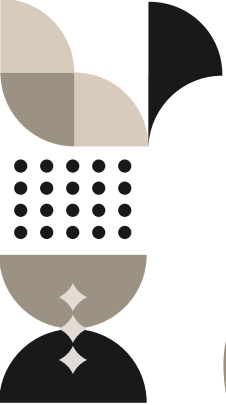The specific data types are a mix of object (string) and numeric data types.

The tables used were movie_basics and movie_ratings.

The MovieDB Dataset
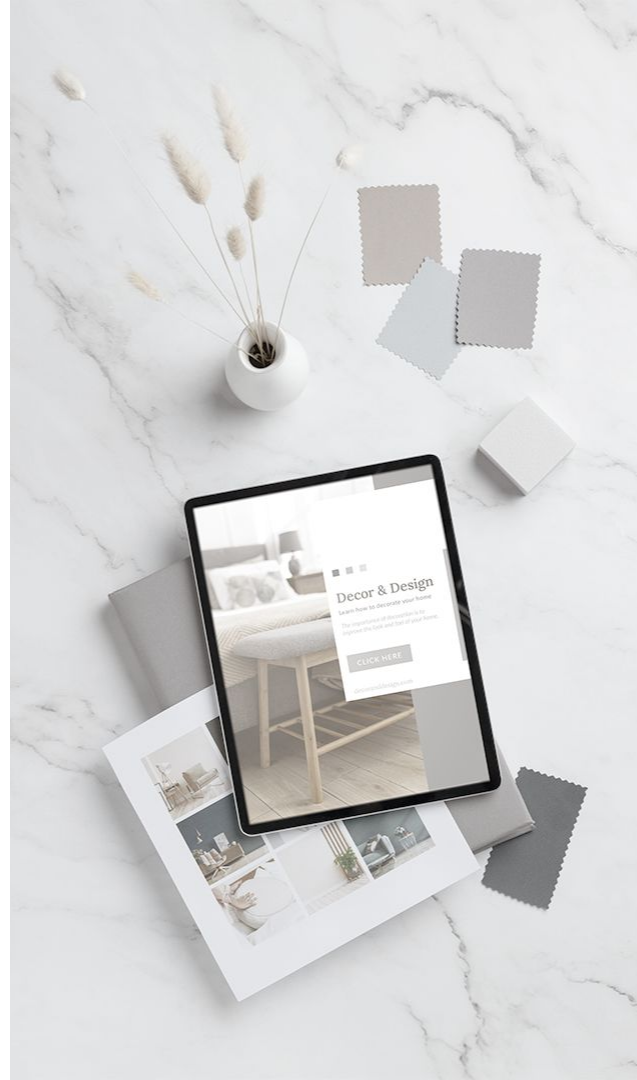
There are 26517 entries in the dataset and 9 columns.

The column names are 'genre_ids', 'id', 'original_language', 'original_title', 'popularity', 'release_date', 'title', 'vote_average', and 'vote_count'.
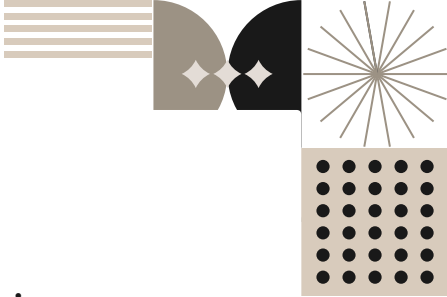
Data types are object(str), float and integer

04

Data
Analysing

- Data was prepared, ensuring its accuracy and completeness, by removing any inconsistencies or errors to ensure reliable insights.
- Afterwards both data visualizations and statistical analysis techniques were used to uncover meaningful patterns and trends within the datasets.
- Through visualizations like graphs, the data was in an easy-to-understand format.
- Statistical analyses was conducted to delve deeper into the data, identifying correlations, trends, and significant relationships between variables.
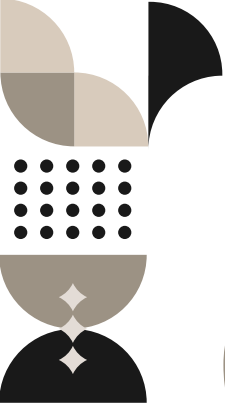
05

Recommendations
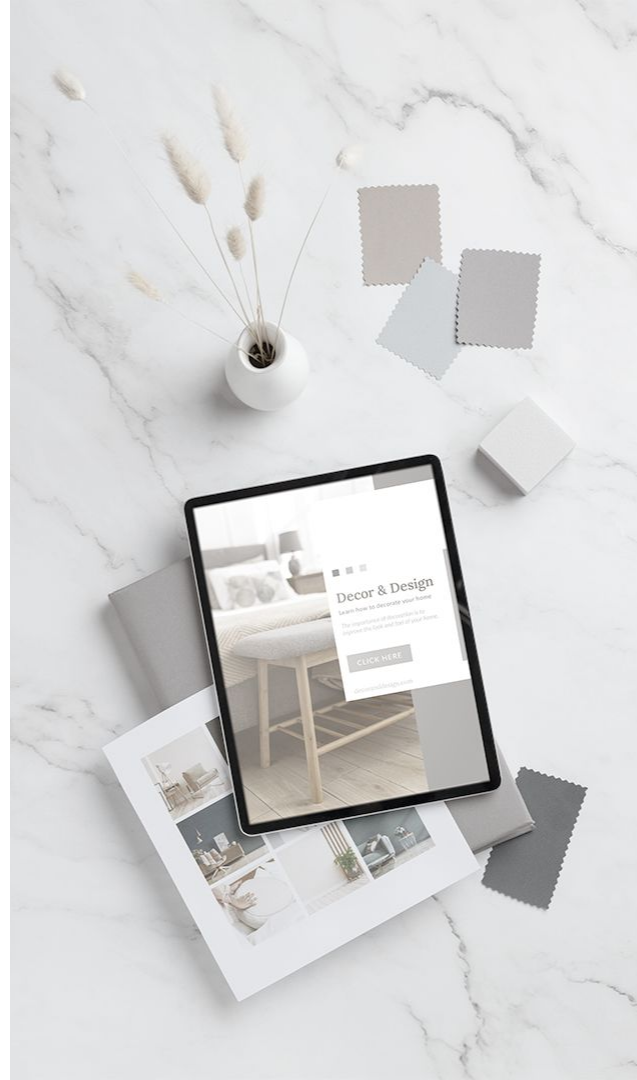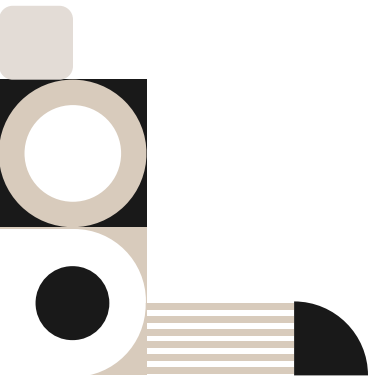
Genre Strategy

Microsoft's movie studio should consider pursuing a diversified content approach across high-performing genres such as drama, documentary, comedy, and horror. They can acquire a larger market share by branching out into new sub-genres.

Microsoft can also experiment with combining genres to create unique and innovative film experiences. They can provide unique and compelling storytelling approaches that appeal to a wide range of audience interests by combining aspects from different genres, such as drama and suspense or horror and humor.

Top 15 Movie Genres
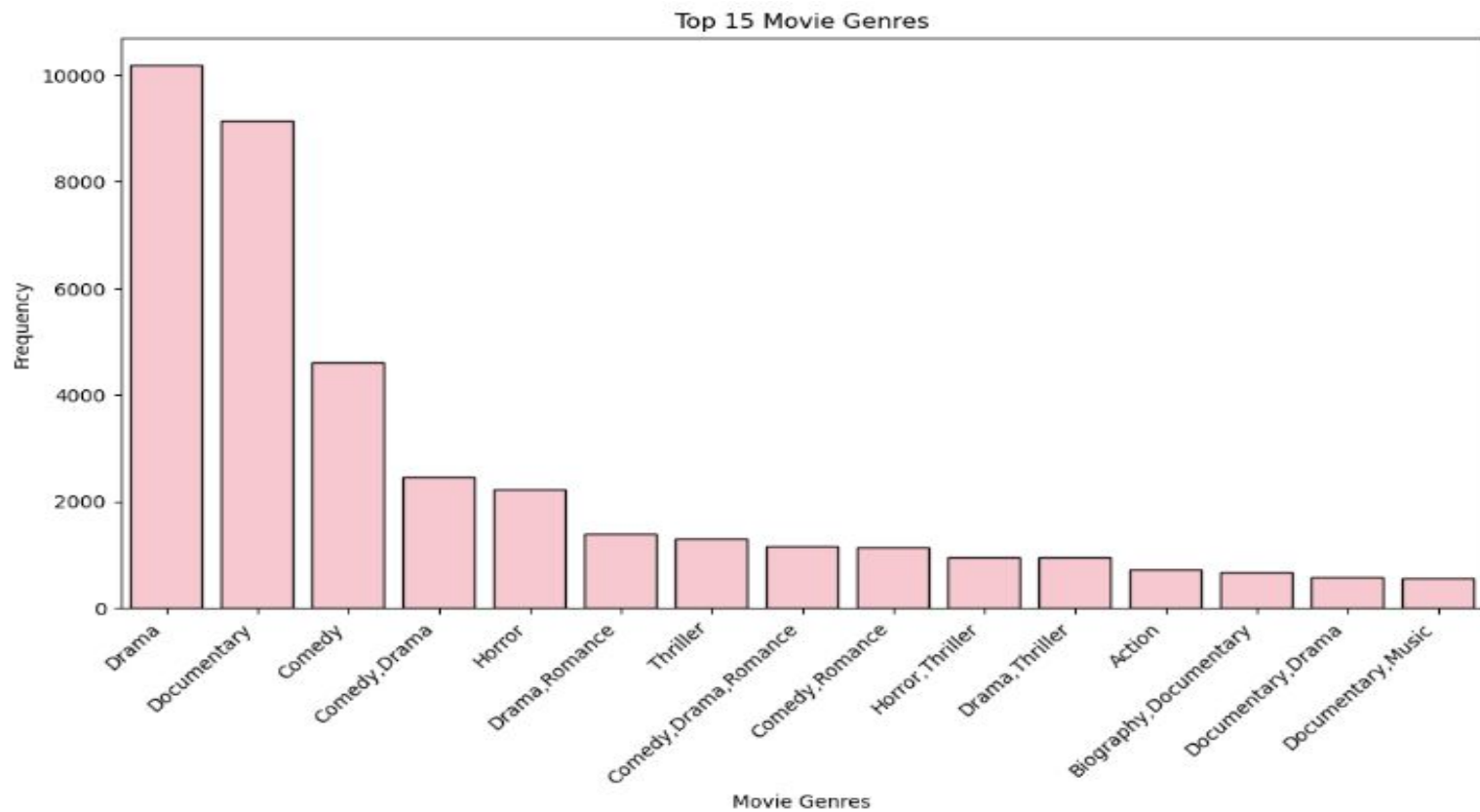
Optimizing Movie Time Release

Understanding the monthly distribution of movie releases guides Microsoft's strategic timing for movie launches. By identifying peak release months and avoiding crowded periods, Microsoft can optimize visibility and maximize audience engagement.

This data-driven approach ensures their films hit the screens when demand is high, increasing the chances of box office success amidst competition.

Distribution of Movie Releases by Month
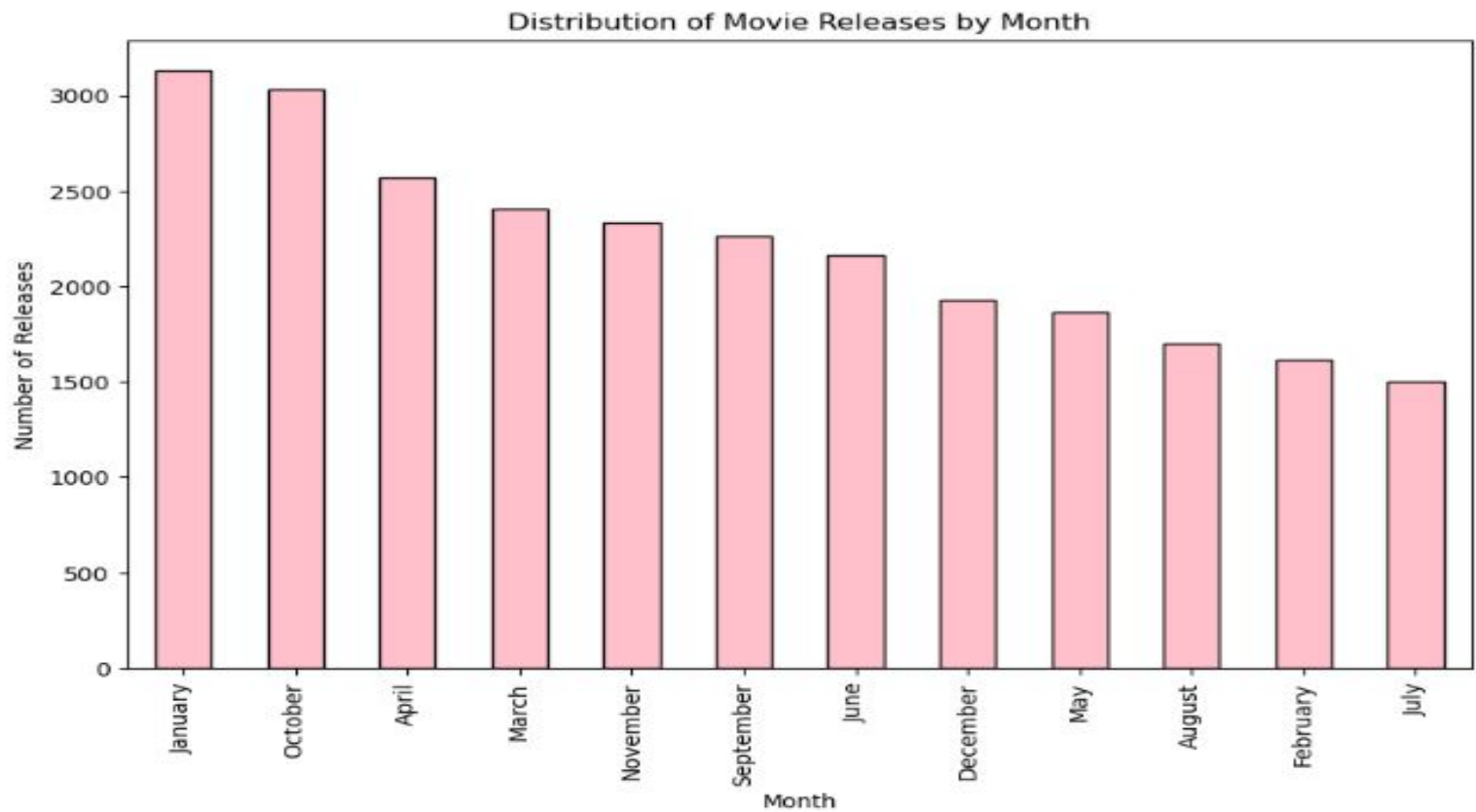
Global Revenue Strategy

The scatter plot shows the correlation between domestic and foreign gross revenue offering Microsoft invaluable insights into the performance of their movies on a global scale.

By analyzing this relationship, Microsoft should consider refining their international distribution strategies and tailoring their marketing efforts to target specific regions more effectively. Understanding how each movie performs in different markets will empower Microsoft to allocate resources wisely and maximize revenue potential worldwide.
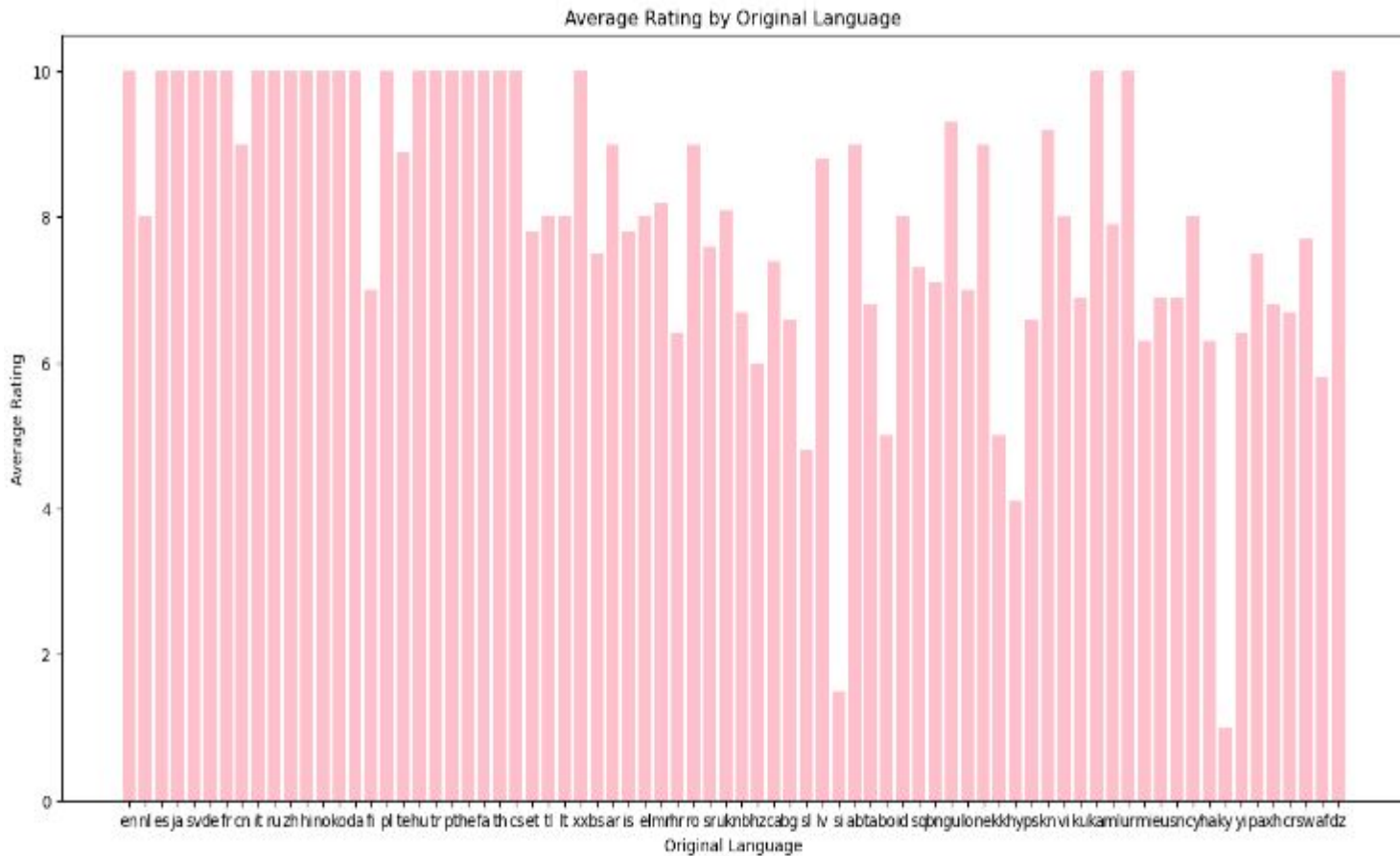
Domestic vs. Foreign Gross
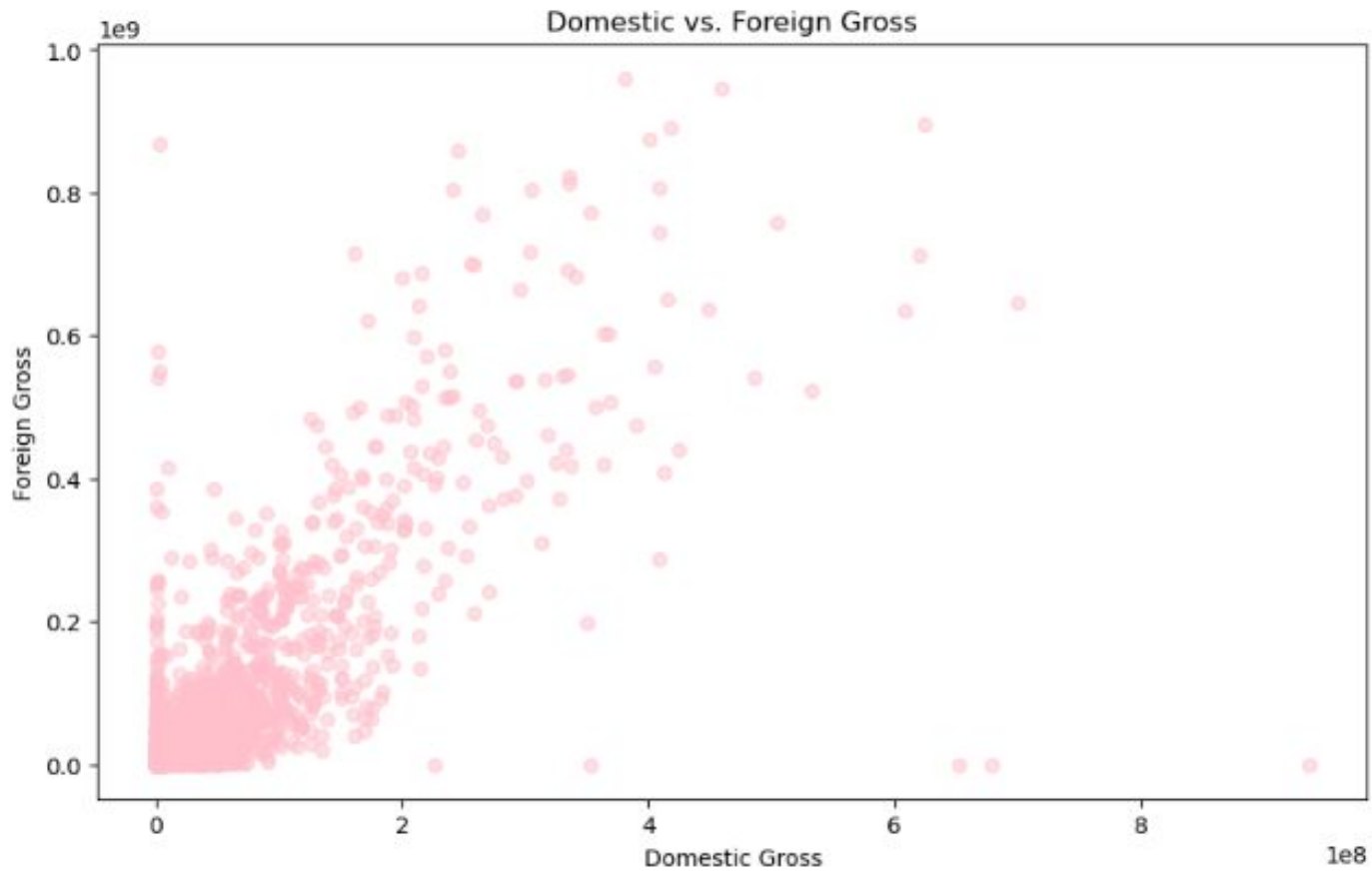
Language Preference Insights

Microsoft should prioritize languages with higher average ratings to tailor their content strategy and enhance audience engagement. By focusing on languages with consistently higher ratings, Microsoft can optimize content selection, forge strategic partnerships with filmmakers from relevant regions, invest in subtitling their movies, and expand marketing efforts in target regions.

This approach will ensure that Microsoft's movie studio venture resonates with diverse audiences worldwide, maximizing the potential for global success.

Average Rating by Original Language

Thanks!

Any Questions ?

Abigail Mwendwa Muthuri

https://www.linkedin.com/in
/abigail-muthuri-b64092261/