MICROSOFT MOVIE ANALYSIS

Phase 1 project

Introduction

This project aims to analyze movie data from various datasets, including IMDb, BOM, RT, TMDB, and TN. The analysis is conducted using Python and its data manipulation and visualization libraries, such as Pandas, Matplotlib, and Seaborn.

Business Problem

- What Genre Should Microsoft focus on?

- Marketing and distribution strategies to boost domestic gross revenue?

- Collaboration with movie studios with a strong distribution network

Objectives

- Examining movie genres distribution and ratings in IMDb data, domestic gross revenue trends in BOM data, movie studio distribution in RT info data, and other relevant insights

- Visualizations such as bar charts, line charts, and scatter plots were used to effectively communicate the findings of the data analysis.

- Come up with recommendations Based on the data analysis findings

The Sampled Data for this project

- This analysis is based on the following data sets
- bom.movie_gross.csv.gz as df_bom_csv
- Rotten tomatoes dataset rt.movie_info.tsv.gz
- The Numbers datset tn.movie_budgets.csv.gz

The Analysis

Top 10 Studios with Highest Average Movie Gross Revenue
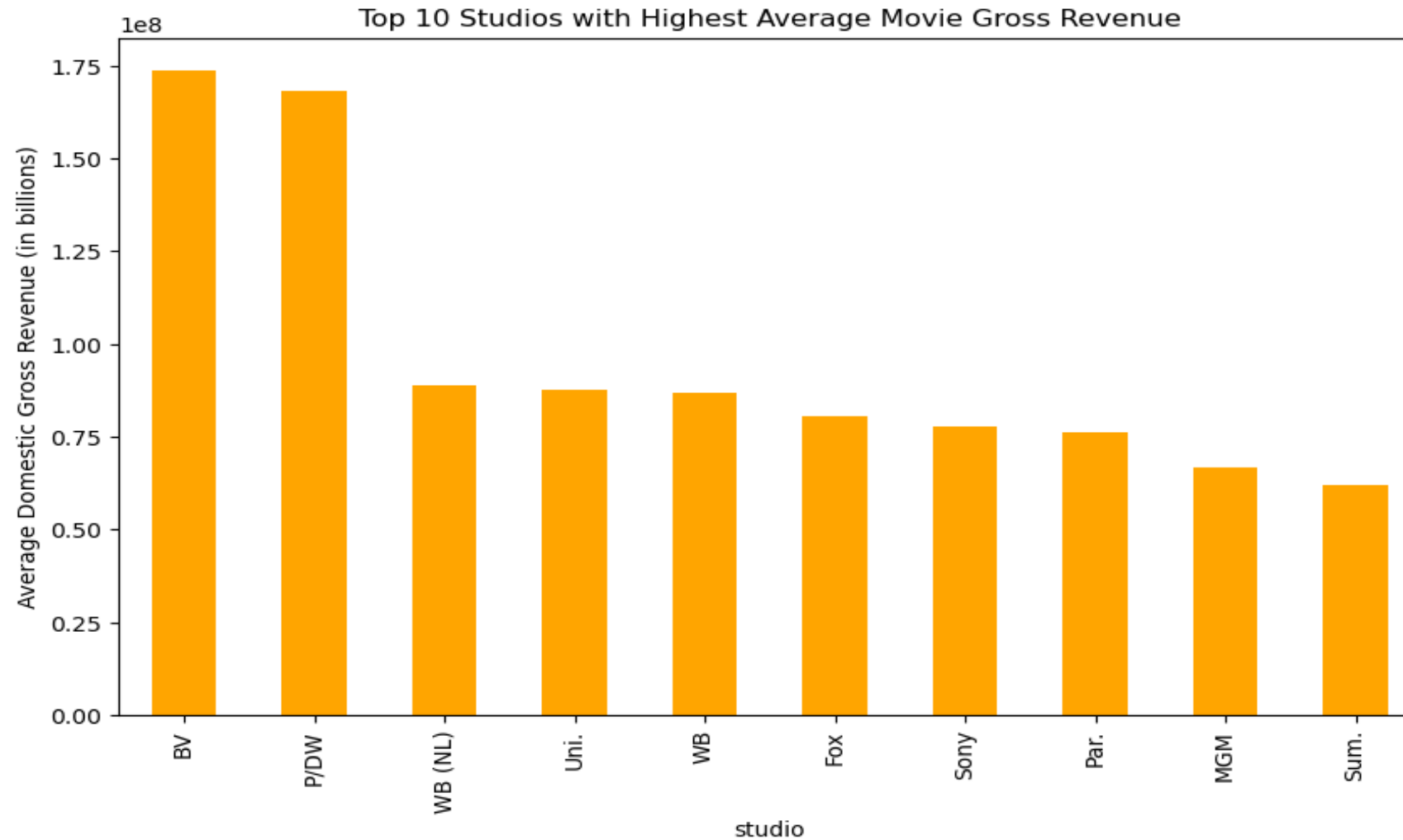
1. The bar plot visually presents the top 10 studios with the highest average domestic gross revenue. This provides insights into which studios, on average, have been most successful in terms of domestic box office performance. Studios with higher average gross revenue may indicate a higher level of success in producing movies that generate higher box office earnings.

2. Additional analysis will be performed on the top studios to understand the factors contributing to their success, such as genre, release date, marketing strategies, and critical acclaim. This can help studios identify patterns and make informed decisions in their future movie production and distribution strategies.

Gross revenue by Year

Domestic Gross Revenue by Year (BOM)
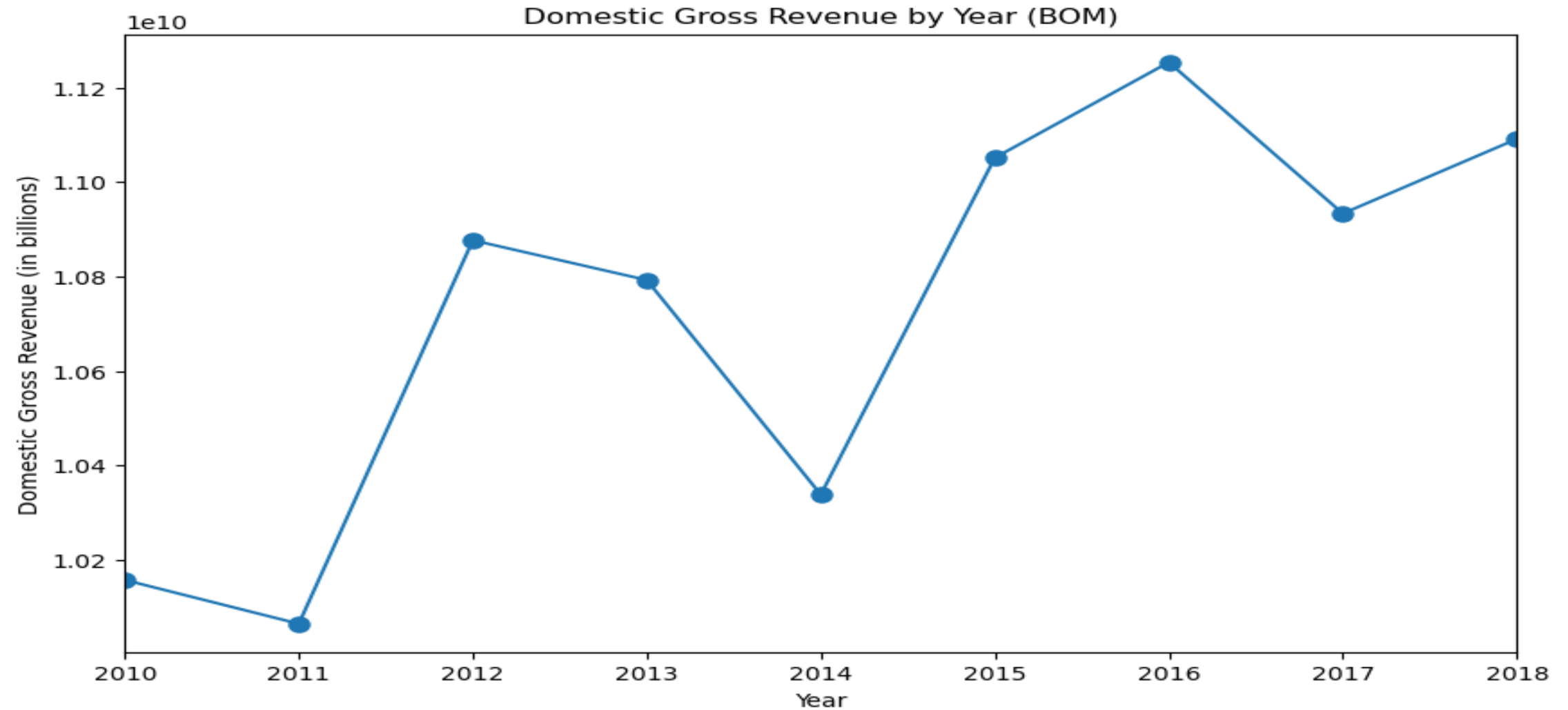
- The domestic gross revenue has changed over the years. line chart shows a line connecting circular markers representing the total domestic gross revenue for each year in the "df_bom_CSV" Data Frame.
- The x-axis shows the years, and the y-axis will show the total domestic gross revenue in billions of dollars. The chart provides a visual representation of how

Conclusion

- In conclusion, this data analysis project provides insights and recommendations for the movie industry based on the analysis of multiple datasets. The findings and recommendations can be used by Microsoft industry stakeholder for production , studios, and marketers, to inform decision-making and improve business strategies.

Recommendations

Based on the data analysis findings, the following recommendations are proposed for Microsoft for the movie industry:

1. Focus on producing movies in genres with high demand, as identified in the IMDb data analysis.

2. Invest in marketing and distribution strategies to boost domestic gross revenue, as suggested by the BOM data analysis.

3. Collaborate with movie studios with a strong distribution network, as revealed in the RT info data analysis.