Microsoft's Movie Studio Exploration: From Data Driven To Silver Screen.

Caroline Njeri Njoroge.
Github: https://github.com/CarolineNjorog3/dsc-phase-1-project
November 11 , 2023.

Summary

In a rapidly evolving entertainment landscape, Microsoft is embarking on a bold venture by establishing a new movie studio. The success of this venture hinges on understanding the intricate nuances of the film industry, from market trends to audience preferences. The primary objective of this project is to analyze and leverage movie data effectively to inform decision-making at Microsoft's movie studio.

Outline

- Business Problem
- Data
- Methods
- Results
- Conclusions

Business Problem

The business problem at hand revolves around Microsoft's venture into the entertainment industry by establishing a new movie studio.

Microsoft faces several challenges and the main points in this venture, include: Understanding Market Dynamics, Language Diversity,Audience Preferences ,Budget vs Performance,Financial Success,Optimal Release Timing,Highly Rated and Popular Movies.

By addressing these questions, Microsoft will be better equipped to navigate the complexities of the film industry, produce successful movies, and maximize profitability while providing an enriching cinematic experience to their audience.

Data

The data used for this project were collected from various reputable sources:Box Office Performance,Genre Trends,Audience Preference, Market Opportunities.The datasets cover thousands of movies, providing a substantial and diverse sample. The time period covered varies across datasets, with information spanning from historical records to more recent data.

The variables used in this analysis have diverse properties, including numerical, categorical, and ordinal data. The sources had movie title,genre,release date,audience preference,critical reception,market opportunities. This information is essential for framing the context and scope of the subsequent data analysis

Methods

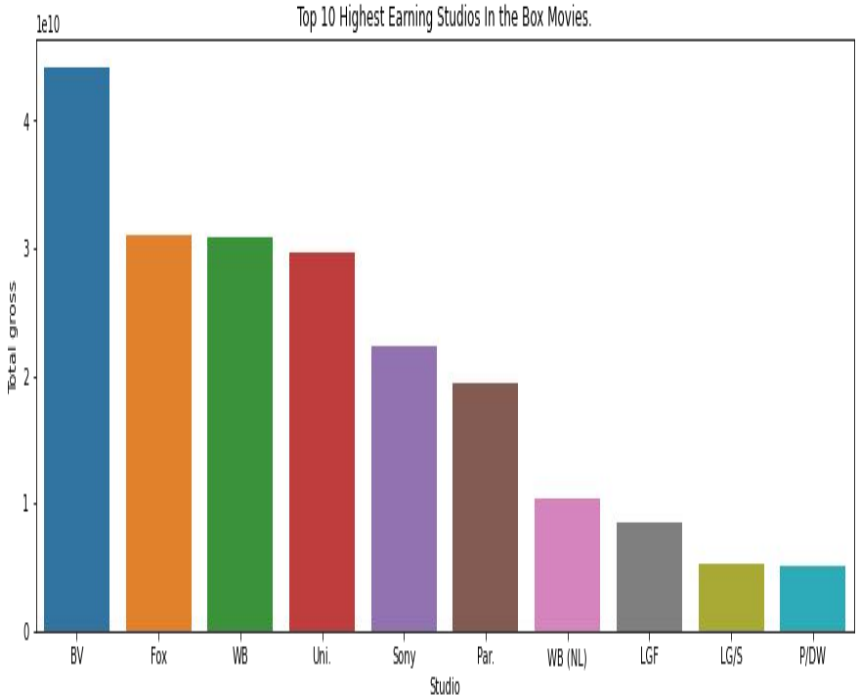
Data Extraction. We extracted data from the different sources into formats we could use in our analysis.

Data Cleaning. Here we selected data required from each sources to help us achieve our objective. We removed unnecessary rows, columns, also null values and data transformations were done example creating revenue fields stored as string to usable float or integer data.

Data Analysis.Under this some data frames were merged and used aggregation methods to extract aggregate data.We then derived the necessary statistical information such as mean, mode and extra.

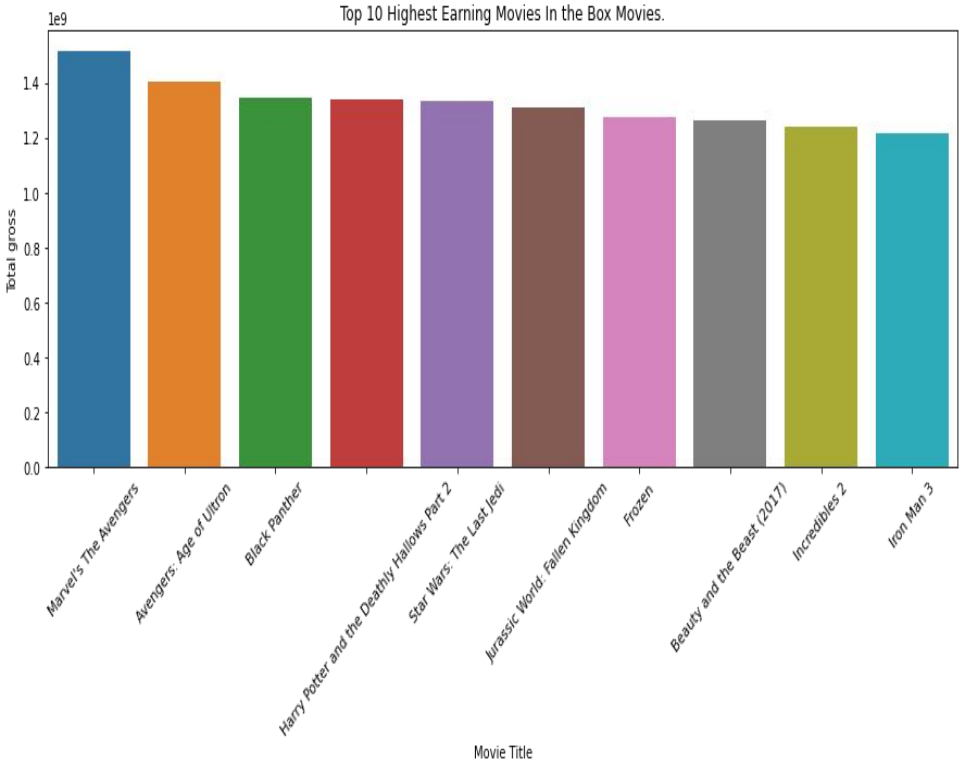
Top 10 Highest Earning Studios in Box Movies.

The graph sheds light on major movie studio and their total gross earnings..Disney's Buena Vista(BV) leads the industry with 44 billion and so forth.



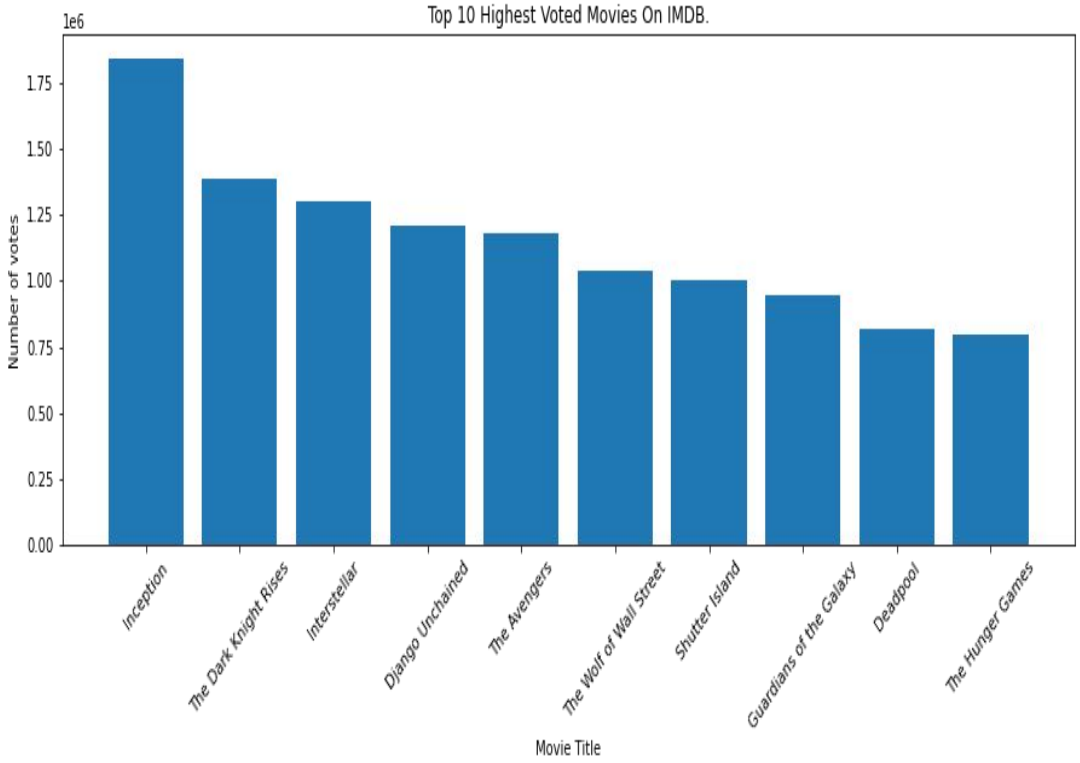
Top 10 highest Earning Movies in Box Movies.

Notable observations from this dataset include the highest-earning movies, such as "Marvel's The Avengers," "Avengers: Age of Ultron," "Black Panther," "Harry Potter and the Deathly Hallows Part 2," and "Star Wars: The Last Jedi."



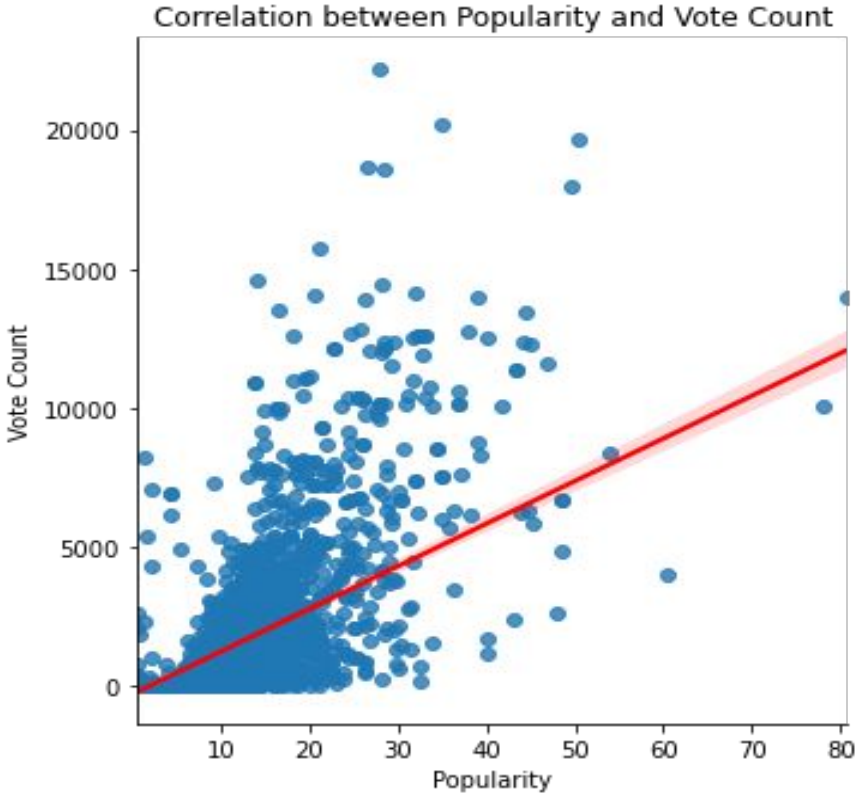
Top Highest Voted Movies

This dataset features a collection of highly-rated and popular films, each with its unique blend of genres and characteristics. Notable films include "Inception," "The Dark Knight Rises," "Interstellar," "Django Unchained," and "The Avengers."



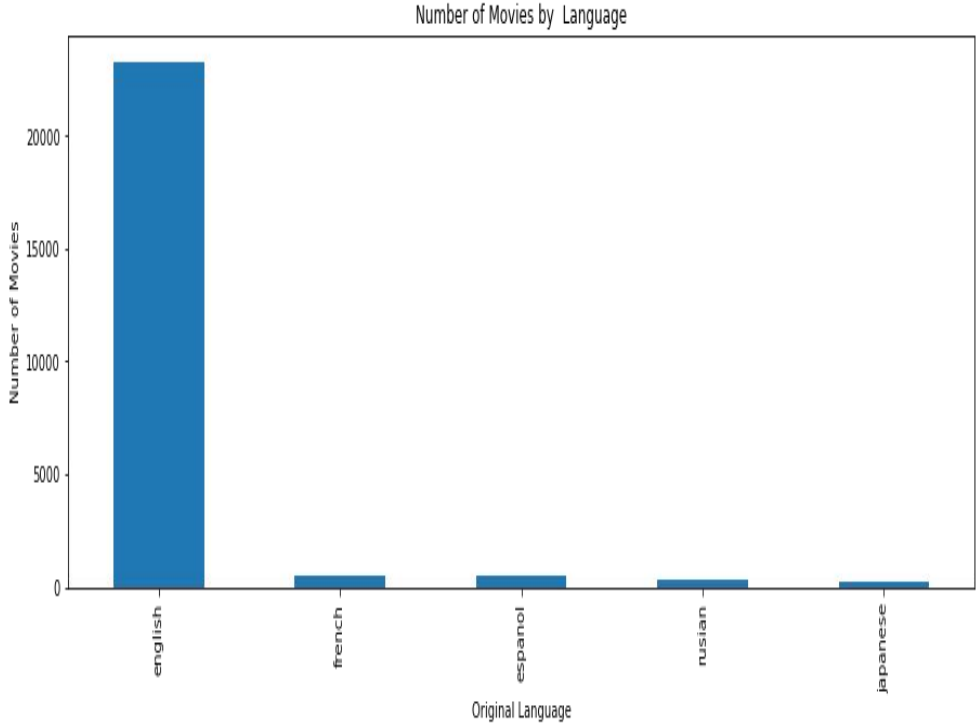
Correlation between Ratings and number of votes.

Our analysis revealed a substantial positive correlation between movie ratings and the number of votes received, indicating that highly-rated movies tend to attract more viewer engagement and votes.



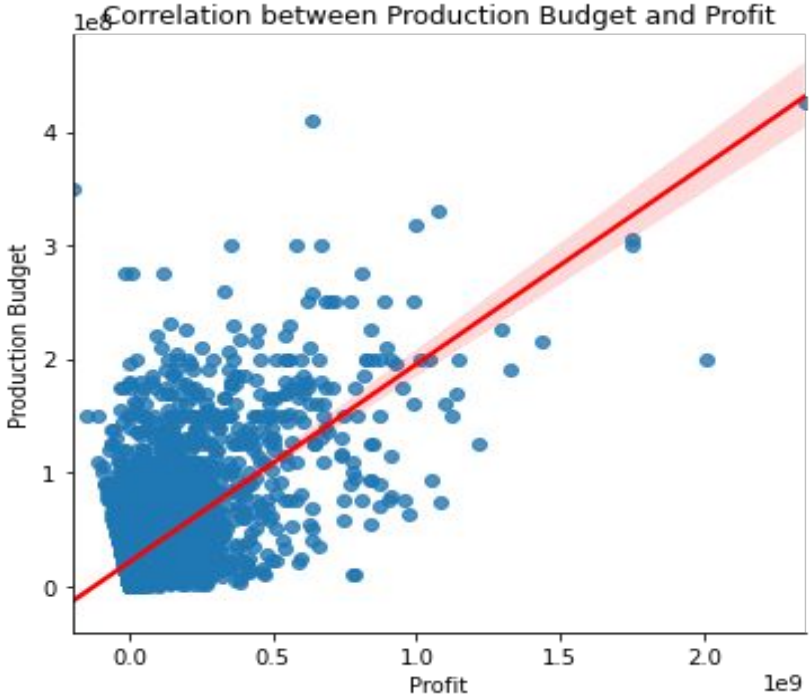
Number of Movies by Language.

The dataset exhibits linguistic diversity, with English (en) being the predominant language choice, highlighting the importance of understanding the prominence of English in the dataset.



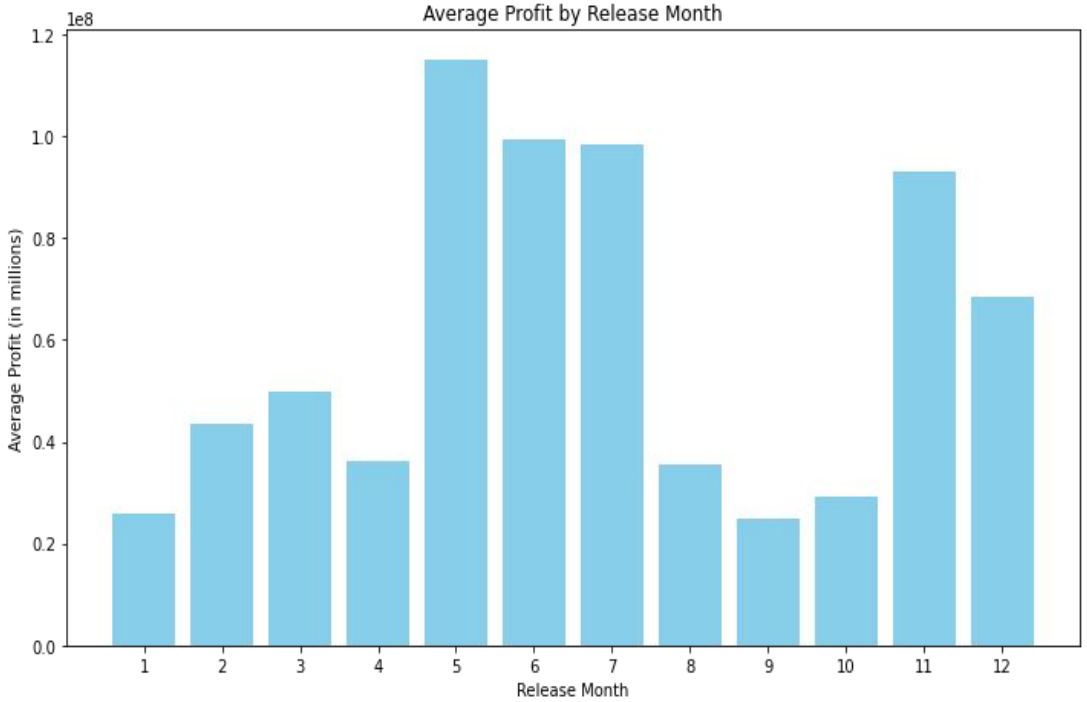
Correlation between Production Budget and Profit.

The analysis revealed a substantial positive correlation between Production Budget and the Profit.The average production budget for these movies is approximately $31,587,757.10 and a return of 59.9 Million so there is a return on investment.



Best Months to Release a Movie.

Our analysis identified the best months for movie releases in terms of achieving the highest average profit. May, June, and July, representing the summer season, lead the way, followed by November and December, which align with the holiday season.



Conclusions

In conclusion,the top genres for viewing pleasure were Action,Adventure,Fantasy,sci-fi etc. The top studios were Disney’s buena vista(BV),Fox etc.The preferred language was English. There was a good correlation between ratings and number votes. The sweet spot for production budget was $31.587 which gives a sufficient amount of return of $59.9 million.

Recommendations

- The average production budget for these movies is approximately $31,587,757.10 and average profit is around $59.9 million.
- Our analysis identified the best months for movie releases in terms of achieving the highest average profit. May, June, and July, representing the summer season, lead the way, followed by November and December, which align with the holiday season.
- These accomplished directors have left an indelible mark on the world of cinema, each bringing their own unique style and storytelling prowess to the screen.E.g Christopher Nolan,Quentinn Tarantino”s.
- Focus on all the following genres to produce the best competitive movies.This are action,adventure, fantasy, sci-fi,drama

Thank You!

Email: youremail@email.com
GitHub: @username
LinkedIn: linkedin.com/in/username/