

Title: "Unveiling Success: A Data-Driven Approach to Microsoft's Cinematic Ventures"

Introduction

In an era characterized by boundless technological innovation and ever-evolving consumer preferences, Microsoft embarks on an ambitious journey into the realm of filmmaking. Armed with a commitment to excellence and a data-driven ethos, Microsoft seeks to revolutionize the cinematic landscape by leveraging the power of information. This project represents a pivotal step in Microsoft's endeavor to establish its footprint in the entertainment industry, employing exploratory data analysis as the cornerstone of its strategic decision-making process. By delving into the depths of data sourced from prominent platforms such as IMDB and Box Office Mojo, we aim to unravel the intricacies of cinematic success, shedding light on the genres that hold the key to profitability and audience acclaim. Through meticulous analysis and insightful interpretation, this endeavor seeks to chart a course for Microsoft's cinematic aspirations, guiding the company towards a future characterized by innovation, prosperity, and cinematic excellence.

Overview

The objective of this project is to conduct exploratory data analysis aimed at extracting valuable insights for Microsoft's venture into establishing a new movie studio. By leveraging data sourced from IMDB and Box Office Mojo, the project endeavors to identify the most successful movie genres based on two key determinants: revenue generation and audience reception. Through comprehensive analysis, we aim to discern which genres exhibit the highest profitability and garner the most favorable audience feedback. Preliminary findings suggest that genres such as Sci-Fi, adventure, animation, and action emerge as lucrative options for Microsoft to prioritize, given their propensity for generating substantial revenue. By focusing on these genres, Microsoft can strategically position itself in the competitive landscape of the film industry, maximizing its chances of success in its cinematic endeavors.

Business Understanding

Microsoft's foray into video content production marks an exciting venture, yet it faces the challenge of navigating an unfamiliar terrain. With aspirations to establish its own movie studio,

Microsoft seeks to capitalize on industry trends to inform strategic decisions. This analysis is geared towards unraveling the intricacies of successful movies at the box office, offering insights that will empower the new movie studio head to make informed choices regarding genres and themes to prioritize. Key questions guiding this exploration include identifying genres with the highest gross income, determining genres with higher average ratings, and elucidating the relationship between average ratings and gross income. Additionally, we aim to dissect the domestic and foreign gross income trends by the start year of movies, providing Microsoft with a comprehensive understanding of the dynamics shaping the film industry landscape. Through rigorous data analysis, we endeavor to equip Microsoft's movie studio with the necessary intelligence to navigate the complexities of the entertainment industry and pave the way for

Data Cleaning

Data cleaning played a pivotal role in refining our datasets to ensure high-quality data for analysis, ultimately leading to more accurate and reliable findings. The process involved several steps to address issues such as missing values, duplicate entries, incorrect data types, and error values.

1. **Check for Missing Values:** The initial step in data cleaning was to identify and handle missing values. Through a systematic approach, we inspected each dataset for missing values and devised strategies to address them. For instance, in the `Movie_gross` dataset, entries with missing values for domestic and foreign gross were replaced with zeros, indicating no revenue for those movies. Additionally, missing values in the 'studio' column were replaced with 'None' to maintain data integrity. In the `imdb_title_df` dataset, missing values in the 'original_title' column were replaced with corresponding values from the 'primary_title' column, ensuring consistency. Furthermore, missing values in the 'runtime_minutes' column were imputed with the median value of the column to preserve the integrity of the runtime data, which is crucial for analysis. Finally, missing values in the 'genres' column were dropped, as genre information is essential for our analysis, and imputation would not be appropriate.
2. **Check for Duplicate Values:** Another aspect of data cleaning involved identifying and removing duplicate entries to prevent skewing of results. By carefully examining each dataset, we eliminated duplicate values where necessary, ensuring that each observation represented a unique movie entry.
3. **Amend Data Types:** Data types were adjusted as needed to ensure consistency and compatibility with analysis techniques. This included converting categorical variables to appropriate data types and ensuring numerical variables were represented accurately.

By systematically addressing these data cleaning tasks, we were able to enhance the quality and reliability of our datasets, enabling more robust analysis and yielding more insightful findings. This meticulous process of data refinement was instrumental in uncovering key trends and patterns in the movie industry, providing valuable insights to guide Microsoft's strategic decisions in establishing its own movie studio.

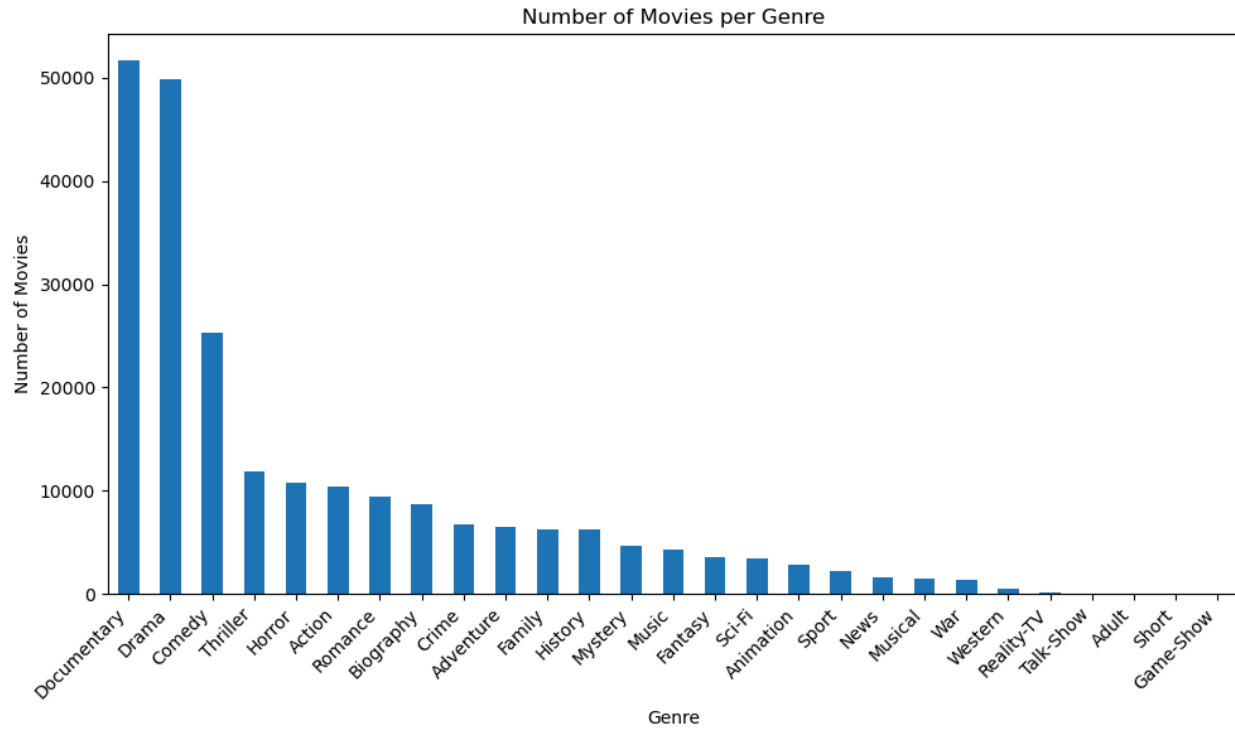
Data Visualization

Data visualization is the graphical representation of information and data. Through visual elements such as charts, graphs, and maps, complex datasets will be presented in a more accessible and understandable format. Data visualization is a crucial tool in data analysis because these graphs below will illustrate complex datasets into visual forms which will be easy to communicate findings.

In this chapter we shall use different graphs to illustrate information and make comparisons.

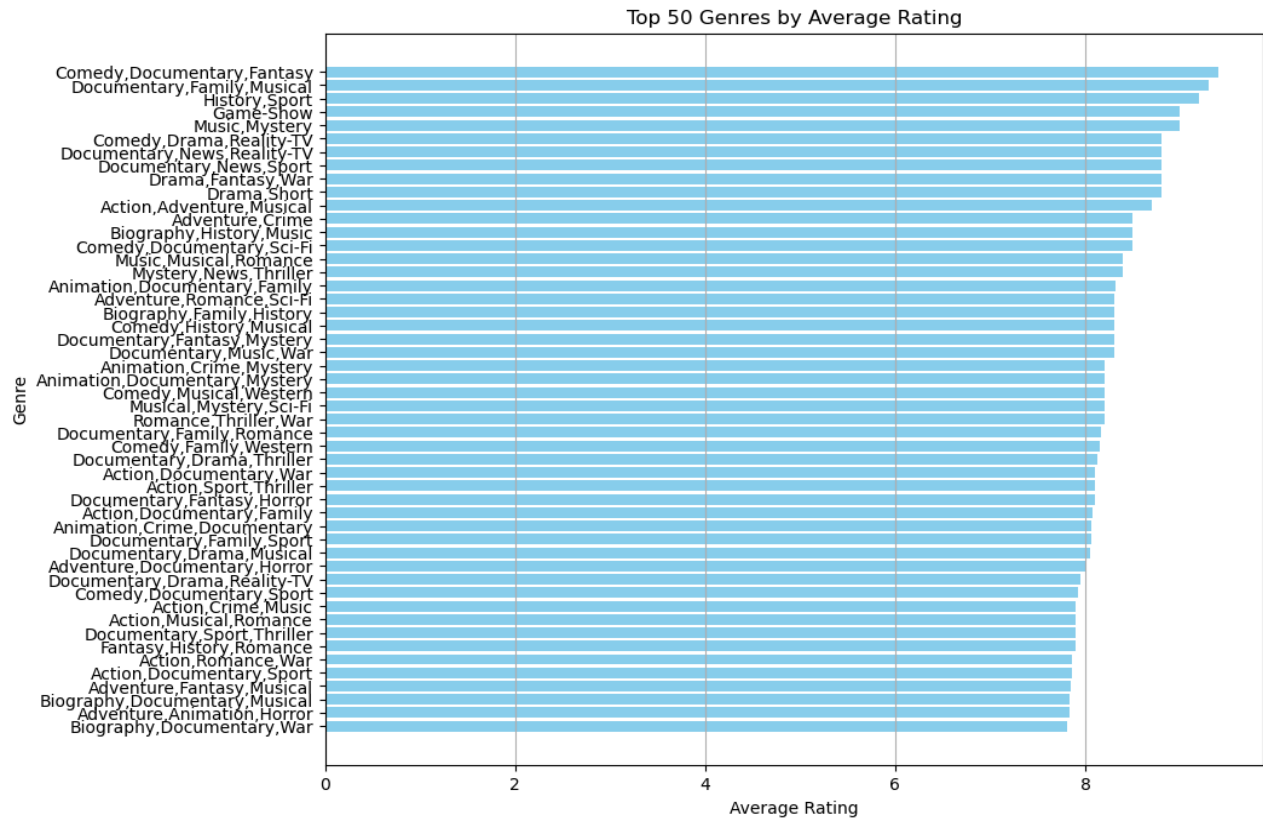
In the bar graph below, the dataset shows that Documentary, Drama, Comedy, and Thriller are the most prevalent genres among the released movies. Documentary and Drama, in particular, stand out as the largest film genres due to their broad and inclusive nature, often leading to movies being classified under multiple genres.

On the other hand, the genres with the fewest movies are Game-show, Short, Adult, Talkshow and Reality-TV. These genres are more specialized, focusing on specific themes or storylines in the movies.



Below, the bars representing the average ratings show minimal variation in height.

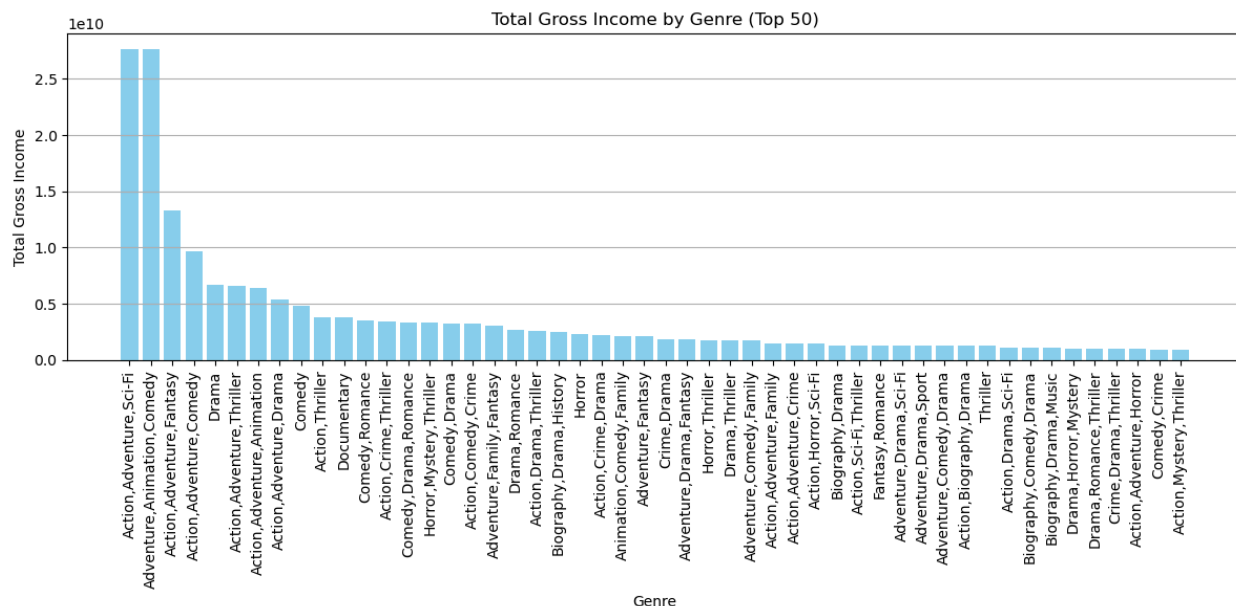
Comedy, Documentary, and Fantasy have the highest average ratings, while Biography, Documentary, War genres have the lowest average ratings.



Below the bar graph represents Box office mojo theatrical box office earnings, excluding other revenue sources like home entertainment sales, rentals, television rights, and product placement fees.

The genres that generate the highest income are Sci-Fi, adventure, animation, and action, whereas the genres with the lowest income are Action, Mystery and Thriller.

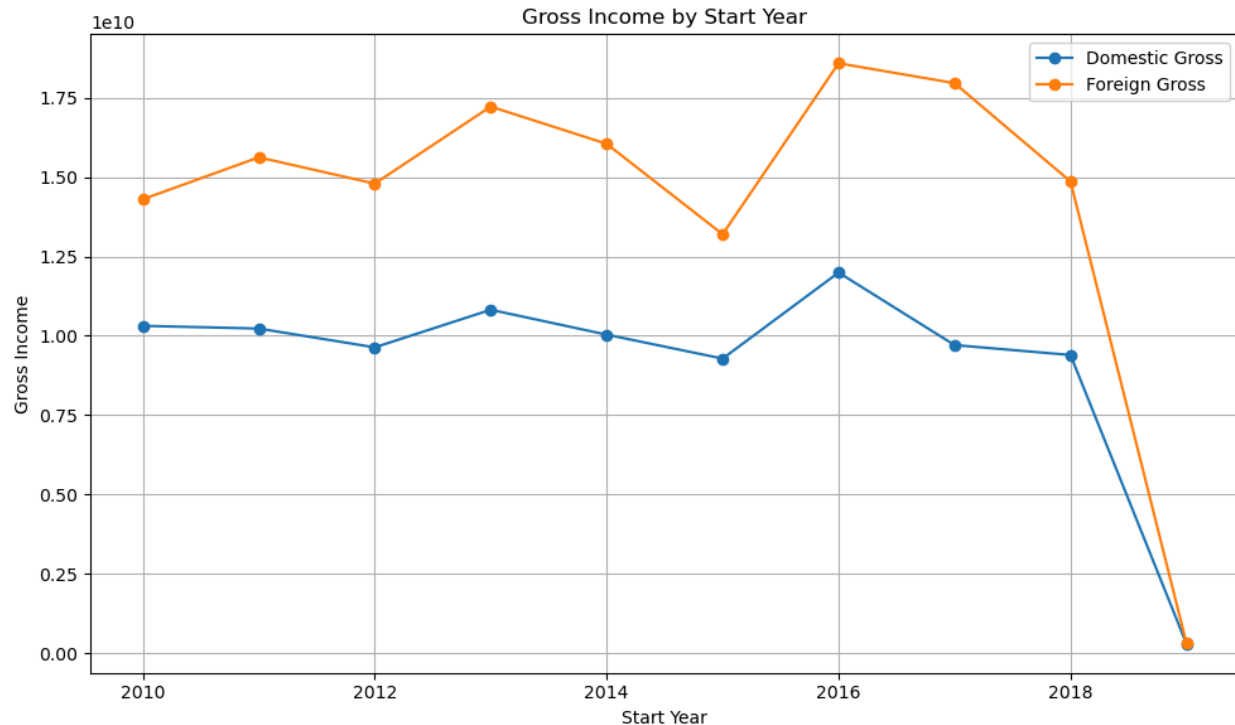
To explore the potential relationship between the average rating and gross income, various factors will be examined as they can influence a movie's overall earnings.



In the graph below we notice a trend: it is common for movies to achieve higher foreign gross income compared to their domestic gross income upon release.

In 2016, movies experienced their peak foreign gross and domestic gross income. However, these incomes significantly declined for movies released in 2019. This decline could be attributed to the increasing popularity of streaming services, with many people opting to use platforms like Netflix instead of going to theaters.

For example, Netflix witnessed remarkable global growth from 2019 onwards. In just the first three months of that year, the streaming video company added an impressive 9.6 million paid subscribers.



Conclusions and recommendations

In closing, this analysis offers valuable insights for Microsoft's endeavor to establish a new movie studio. I observed that while genres like Comedy, Documentary, and Fantasy boast high average ratings, it's the genres of Sci-Fi, Adventure, Animation, and Action that demonstrate the greatest income-generating potential. Additionally, the decline in theatrical gross income, potentially influenced by the rise of streaming services like Netflix, underscores the importance of adapting to evolving industry trends. As we move forward, we recommend a strategy that balances audience satisfaction with commercial success. This entails diversifying the genre portfolio to include lucrative genres while prioritizing quality content creation. Furthermore, leveraging strategic partnerships with streaming platforms and remaining adaptable to changing consumer preferences will be key in maximizing revenue potential and establishing a strong foothold in the dynamic entertainment landscape. With these insights in mind, Microsoft is poised to embark on an exciting journey in the world of filmmaking, capturing the hearts and minds of audiences worldwide.

Recommendations Microsoft should consider:

- Diversify Genre Portfolio: While high ratings are important for audience satisfaction, prioritize genres such as Sci-Fi, Adventure, Animation, and Action that have proven to be lucrative in terms of income generation.
- Leverage Streaming Platforms: Recognize the growing trend of streaming services and consider strategic partnerships

or distribution deals with platforms like Netflix to reach a wider audience and maximize revenue potential. -Focus on Quality Content: While considering commercial success, prioritize creating high-quality content that resonates with the target audience, as reflected in genres with high average ratings like Comedy and Fantasy. -Adapt to Changing Trends: Stay agile and adaptable to evolving consumer preferences and industry trends, including exploring opportunities in streaming, digital distribution, and innovative marketing strategies.

Thank You

Presented by Prossy Nansubuga Kamau

www.linkedin.com/in/prossy-nansubuga-kamau-mba-bbit-a66b5045