Data Science Project Phase 1

Joshua Ooko

Business understanding

Microsoft's new movie studio aims to establish itself in the competitive film industry by creating engaging and successful content. To achieve this, it is crucial for Microsoft to understand the current trends and genres that resonate with audiences. This report explores the types of films that have performed exceptionally well at the box office in recent years. By analyzing industry landscape and successful movies, actionable insights are presented to guide Microsoft's decision-making process. These insights will help Microsoft strategically position their movie studio, create content that appeals to audiences, and drive box office success. With limited knowledge and experience in film production, Microsoft can leverage these insights to make informed decisions and establish a strong presence in the evolving entertainment landscape. By understanding the pulse of the film industry, Microsoft can enhance their chances of producing commercially successful and captivating content that engages audiences worldwide.

Problem Statement

Exploring what types of films are currently doing the best at the box office. You must then translate those findings into actionable insights that the head of Microsoft's new movie studio can use to help decide what type of films to create.

objective

The objective of this report is to analyze the current landscape of the film industry and identify the types of films that have been successful at the box office. By examining trends, genre preferences, and factors contributing to the success of recent movies, the report aims to provide actionable insights that Microsoft's new movie studio can use to make informed decisions about the types of films they should create. The ultimate goal is to guide Microsoft in strategically positioning their movie studio to produce compelling and commercially successful content that resonates with audiences and drives box office revenue

Specific objective

1. Analyze the current landscape of the film industry, including market trends, audience preferences, and box office performance.
2. Identify the genres that have been most successful at the box office in recent years.
3. Examine the factors contributing to the success of these films, such as average ratings, number of votes, and domestic and worldwide gross.
4. Determine the correlation between production budget and box office performance to guide decision-making regarding investment in film production.
5. Provide insights on the average production budgets for different genres and their impact on box office success.
6. Investigate the relationship between average ratings and production budget to understand the trade-offs between artistic quality and financial returns.
7. Analyze the performance of different budget ranges to identify potential opportunities for Microsoft's movie studio.
8. Present actionable recommendations on the types of films Microsoft should consider creating based on the analysis of successful genres, budget considerations, and audience preferences.
9. Guide Microsoft in effectively positioning their movie studio to create content that captivates audiences, drives box office success, and establishes their presence in the competitive film market.
10. Enable Microsoft to make informed decisions that enhance the chances of creating compelling and commercially successful films, resulting in long-term profitability and audience engagement.

Notebook Structure

Reading Data
Data Cleaning
Data Analysis
Conclusion
Recomendation

Data understanding

Data understanding is a critical step in analyzing the film industry. It involves exploring and acquiring relevant datasets such as movie budgets, gross earnings, genre information, ratings, and other factors that contribute to box office success. By examining these datasets, we gain insights into the current state of the film industry, including trends, genres, and factors influencing movie performance. This understanding lays the foundation for further analysis and provides valuable information for decision-making in Microsoft's new movie studio venture.

Methodology

- Reading the Data
- Data cleaning

Before diving into the analysis, it is essential to clean the data by removing irrelevant columns, handling missing values, standardizing formats, and addressing inconsistencies. This ensures data integrity and prepares it for further exploration.

- Data Analysis:

Exploring the data through descriptive statistics, visualizations, and statistical analysis helps uncover patterns, relationships, and outliers. It provides valuable insights into the data, identifies trends, and validates assumptions. The findings guide decision-making and provide a foundation for further analysis or modeling.
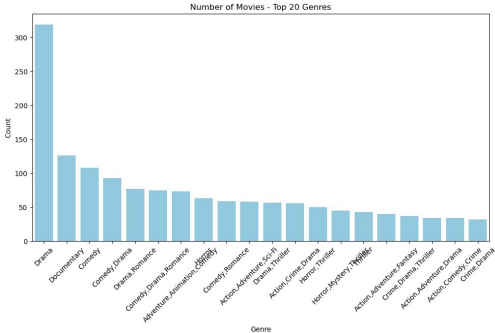
- Conclusions

In conclusion, our analysis of the film industry landscape and successful movies provides valuable insights for Microsoft's new movie studio. We have identified the genres that have been performing exceptionally well at the box office, highlighted trends, and explored factors contributing to box office success. This understanding can guide Microsoft in strategically positioning their movie studio to create compelling content that resonates with audiences and drives box office revenue.

- Recommendations
- Focus on popular genres: Prioritize the production of films in genres that have consistently performed well at the box office, such as action, adventure, comedy, and animation. These genres have a proven track record of attracting audiences and generating revenue.
- Embrace current trends: Stay updated with the latest industry trends and audience preferences. Consider the growing popularity of superhero films, remakes, and franchises, and incorporate elements that resonate with audiences into the studio's film concepts and storytelling.
- Invest in production quality: Allocate resources to ensure high-quality production values. Emphasize visually stunning and technically impressive films that captivate audiences and set the studio's content apart from the competition. Attention to detail and production value can significantly enhance the viewing experience.
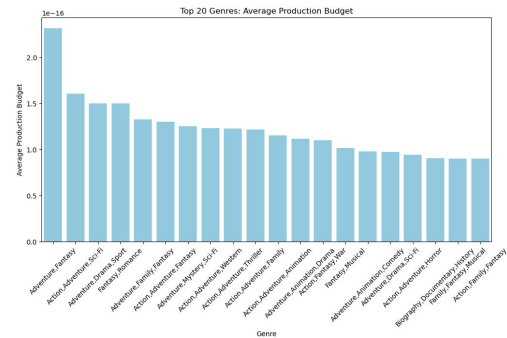
Data Analysis

This bar graph is displaying the count of movies for the top 20 genres. The x-axis represents the genre names, and the y-axis represents the movie count. Each bar corresponds to a genre, allowing you to compare the number of movies across different genres.

Number of Movies - Top 20 Genres

Data Analysis

This bar graph shows calculatd the average production budget for each genre and select the top 20 genres with the highest average production budget. The resulting bar graph will display the average production budget on the y-axis and the genre names on the x-axis. Each bar represents a genre, allowing you to compare the average production budgets across different genres.

Top 20 Genres: Average Production Budget

Data Analysis

 This graph selecst the top 20 genres with the highest total domestic gross. The resulting bar graph will display the total domestic gross on the y-axis and the genre names on the x-axis. Each bar represents a genre, allowing you to compare the box office performance across different genres.

Top 20 Genres: Box Office Performance