

MICROSOFT MOVIE STUDIO

By :Chris Mukiri



Overview

- In this project, the goal is to assist Microsoft in entering the movie industry by providing actionable insights based on exploratory data analysis (EDA).
- The business problem is to determine what types of films are currently performing well at the box office, ultimately guiding Microsoft's new movie studio in making informed decisions on film production.
- The project involves data analysis on various factors, including movie genres, production budgets, box office earnings, and release dates, to answer key data questions.

Key Objectives

1. Analyze the movie industry data to understand factors influencing box office performance.
2. Identify trends related to movie genres, production budgets, and audience ratings.
3. Make data-driven recommendations for Microsoft's movie production strategy.

Data Sources

- [Box Office Mojo](data-sources/box-office-mojo.csv)
- [The Numbers](data-sources/the-numbers.csv)
- [IMDb](data-sources/imdb.csv)

Data Preparation

- We performed the following data preparation steps:
 - Data cleaning to remove missing values and duplicates.
 - Merging data from various sources to create a unified dataset.

Data Analysis

- Top Performing Movies We analyzed the top 20 highest-grossing movies both domestically and worldwide, examining the relationship between production budget and earnings.
- Lowest Production Budget Movies We explored the 20 movies with the lowest production budgets to understand their performance.
- Top-Rated Shows We ranked the top 20 highest-rated shows and identified preferences for documentary content. Additionally, we examined the correlation between runtime and average ratings.

Recommendations

- Based on the data analysis, we make the following recommendations to Microsoft's movie studio:
 1. Consider investing in blockbuster movies, as there is a general trend of better income generation for such movies.
 2. Explore the documentary genre, as it is preferred among the top-rated shows.
 3. Take into account that the runtime of a movie does not strongly correlate with its average rating, and shorter movies can also perform well.

These recommendations can help Microsoft make informed decisions when entering the movie industry.

Data Visualizations

- We have created various data visualizations to support our analysis. These include:
 - Box plots for budget and gross earnings.
 - Scatter plots for budget vs. earnings.
 - Time series line charts for release dates.

Thank You!

- Email: chrisjnganga@gmail.com
- Github: <https://github.com/CMukiri/dsc-phase-1-project>