Academy Xi

Project Phase 1

Data Analysis for Microsoft

Summary

Microsoft aims to venture into the world of original video content creation by establishing a new movie studio. Despite lacking experience in the film industry, Microsoft is eager to capitalize on the success seen by major companies in the movie sector. To make informed decisions and establish a successful entry strategy, there is a need to analyze the current trends and preferences in the film industry.

The task involves researching and identifying the types of films that are currently performing well at the box office. This research will be crucial in providing actionable insights to guide Microsoft's new movie studio in selecting the right genres or themes for their upcoming productions.

Outline

BUSINESS PROBLEM

DATA

METHODS

RESULTS

CONCLUSIONS

Business Problem

Microsoft recognizes the success of major companies in the creation of original video content and aims to venture into this domain by establishing its own movie studio. Despite **lacking expertise in the film industry**, Microsoft is enthusiastic about entering this competitive market. The challenge lies in understanding the current trends in the film industry and identifying the types of films that are garnering success at the box office.

The task at hand is to conduct research on top-performing films, analyze their characteristics, and translate these findings into actionable insights. These insights will guide the head of Microsoft's new movie studio in making informed decisions on the types of films to produce, ensuring a competitive edge in the market.

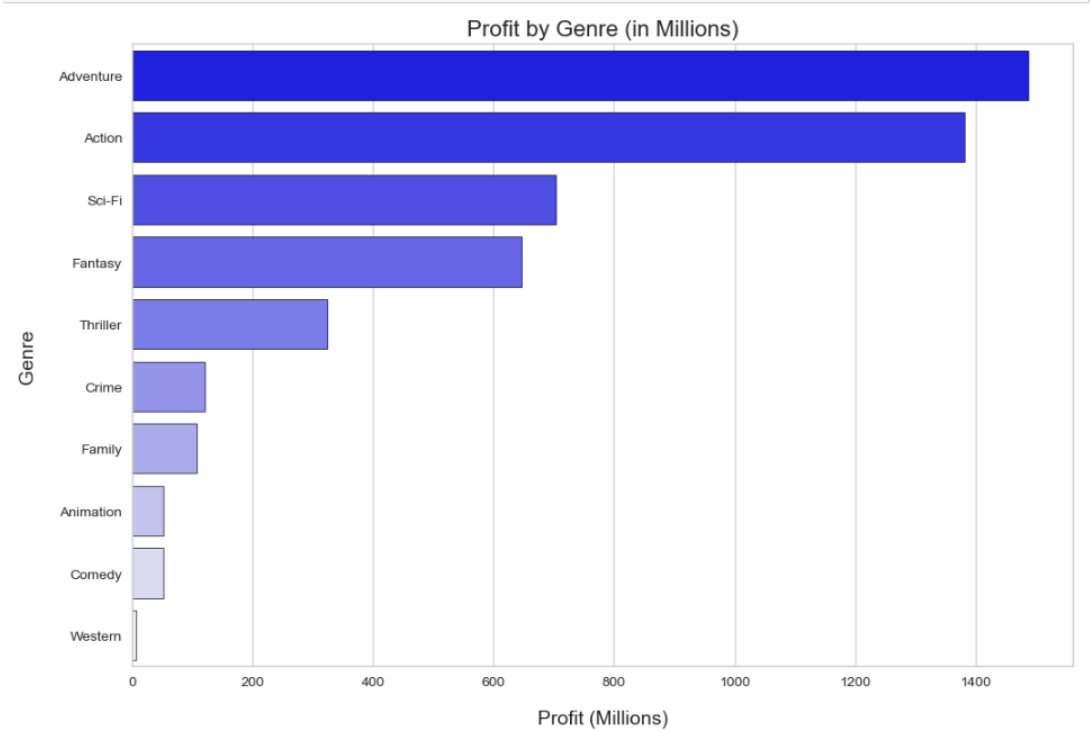
Data

The project utilizes three CSV files for data analysis, which need to be merged to form a comprehensive data frame. Data sets 1 and 2 share a common primary key, facilitating their straightforward merging. However, Data set 3 necessitates the creation of a new primary key to enable its integration with the other data sets for a complete analysis

Methods

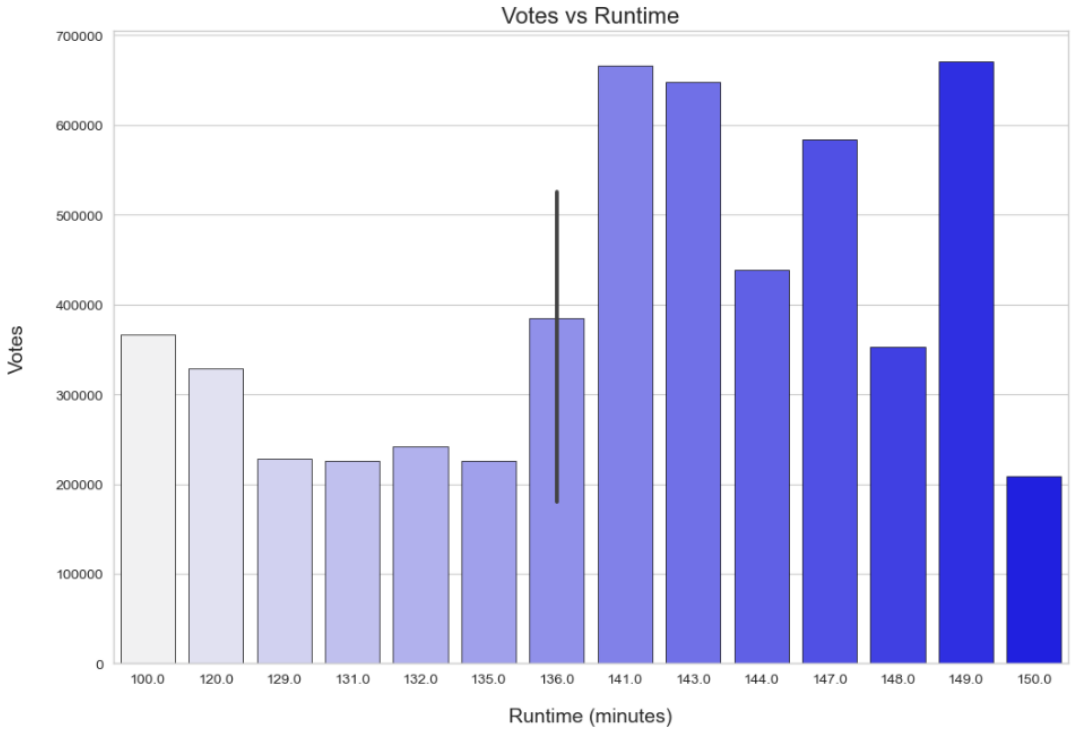
- The project begins by importing essential libraries like Pandas, NumPy, Matplotlib, and Seaborn in a Jupyter notebook and uploading the CSV files.
- The next phase involves merging data sets using their respective primary keys.
- Following the merging process, the data undergoes a thorough inspection for null values, duplicates, and statistical analysis.
- Finally, the findings are visually presented using graphs and statistics to offer clear insights from the merged data sets.

Visual Presentation



Results

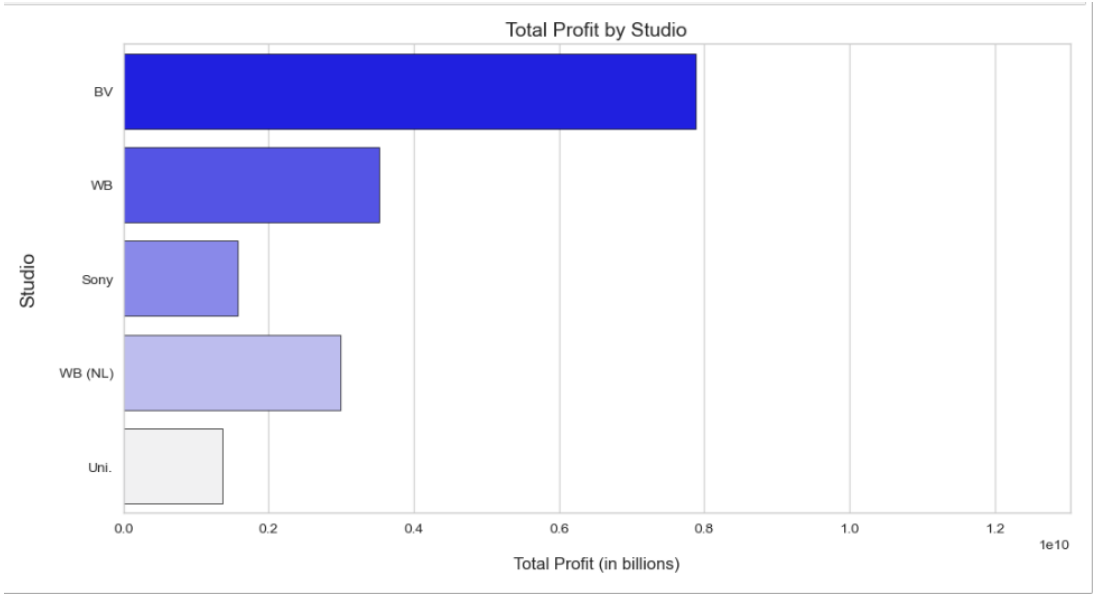
The graph depicts the relationship between genres and their profit in millions. Adventure, action, and Sci-Fi are the top three genres in terms of revenue.



Results

The graph displays the distribution of votes according to the movie's runtime in minutes. The highest number of votes falls within the 141 to 149-minute range.

Results



The graph illustrates the profit (measured in billions) of five movie producers, with Buena Vista ranking highest and Universal Studios at the bottom of the list.

Conclusions

In conclusion, Adventure, action, and Sci-Fi are the most profitable movie genres, with the highest revenue generated. Additionally, movies with a runtime between 141 to 149 minutes receive the most votes from viewers. Lastly, among the five movie producers depicted, Buena Vista leads in profit, while Universal Studios trails at the bottom.

Thank You!

Email: ben_valencia12@yahoo.com.au
GitHub: @ben valencia
LinkedIn: https://www.linkedin.com/in/benvalencia/