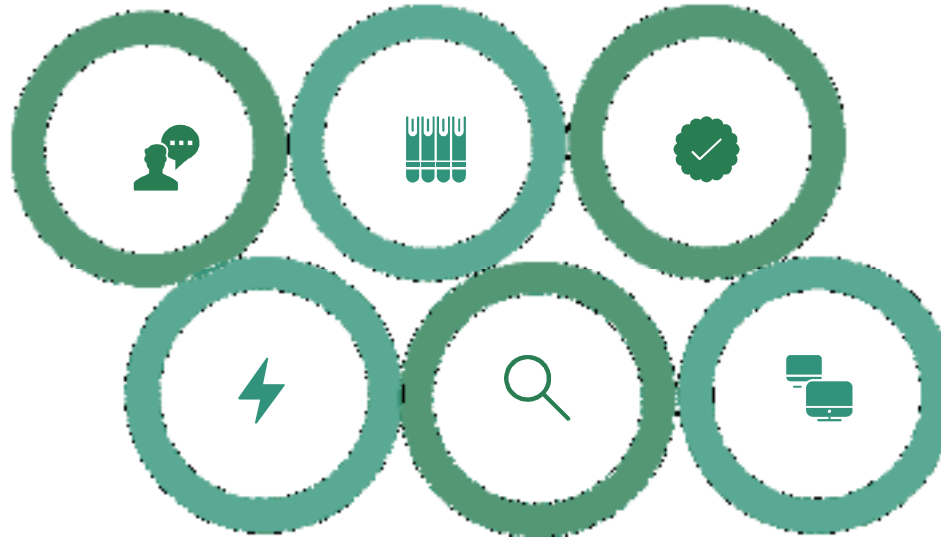


Microsoft Movie Analysis

Author: Brian Kariithi

PROJECT OVERVIEW



Microsoft has made the strategic decision to establish a new film studio and is seeking deeper insights into the most successful film genres at the box office. This initiative involves employing descriptive statistical analysis on datasets collected from the IMDb website to understand the combinations of genres that excel in various key areas. The analysis encompasses four distinct datasets, examining the genres that performed exceptionally well in terms of domestic gross sales, foreign gross sales, average ratings, number of votes, and production budgets relative to domestic and worldwide gross earnings. The findings highlight that the combination of Action, Adventure, and Sci-Fi consistently emerged as the top performer across the categories of Domestic Sales, Foreign Sales, and number of votes. Adventure played a pivotal role in the majority of the top 20 combinations across these three categories. Based on this analysis, the recommendation for Microsoft's movie production would be to focus on the Action, Adventure, and Sci-Fi genre, given its prevalence and success in the dataset. It's noteworthy that after data cleaning, there were 322 unique genre combinations considered. Additionally, Adventure and Action paired with either Animation or Fantasy proved to be successful combinations, making them a secondary recommendation for consideration. The combination of Adventure, Animation, and Comedy also demonstrated favorable performance in both Domestic and Foreign Sales, standing as a third recommendation. The analysis underscores Adventure as a robust genre, consistently associated with popular and successful films.

• Business Problem •

Microsoft aims to create profitable movies and is keen on identifying the most successful genres. To address this inquiry, an analysis was conducted on Domestic and Foreign Sales data, as well as the production budget in relation to Domestic gross and worldwide gross. The objective was to determine the financially lucrative genres. Additionally, the average rating and number of votes were examined for each genre, providing insights into the correlation between popularity and financial success.

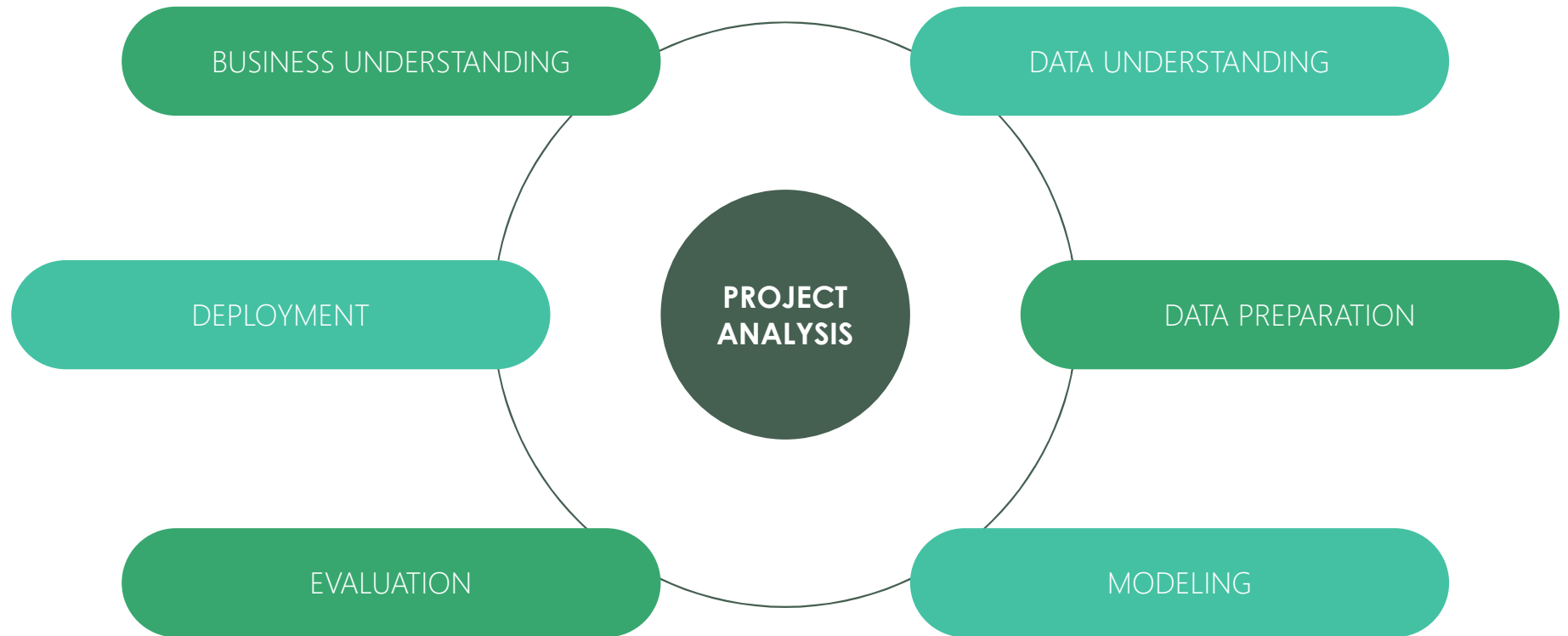
QUESTIONS FOR ANALYSIS

What considerations need to be taken into account when deciding on the optimal mix of genres for Microsoft's movie production endeavors?

which types of movies to produce in relation to audience preferences and industry trends

How can the insights gained from this analysis be leveraged to inform long-term strategic planning and decision-making within the Microsoft Movie Making industry?

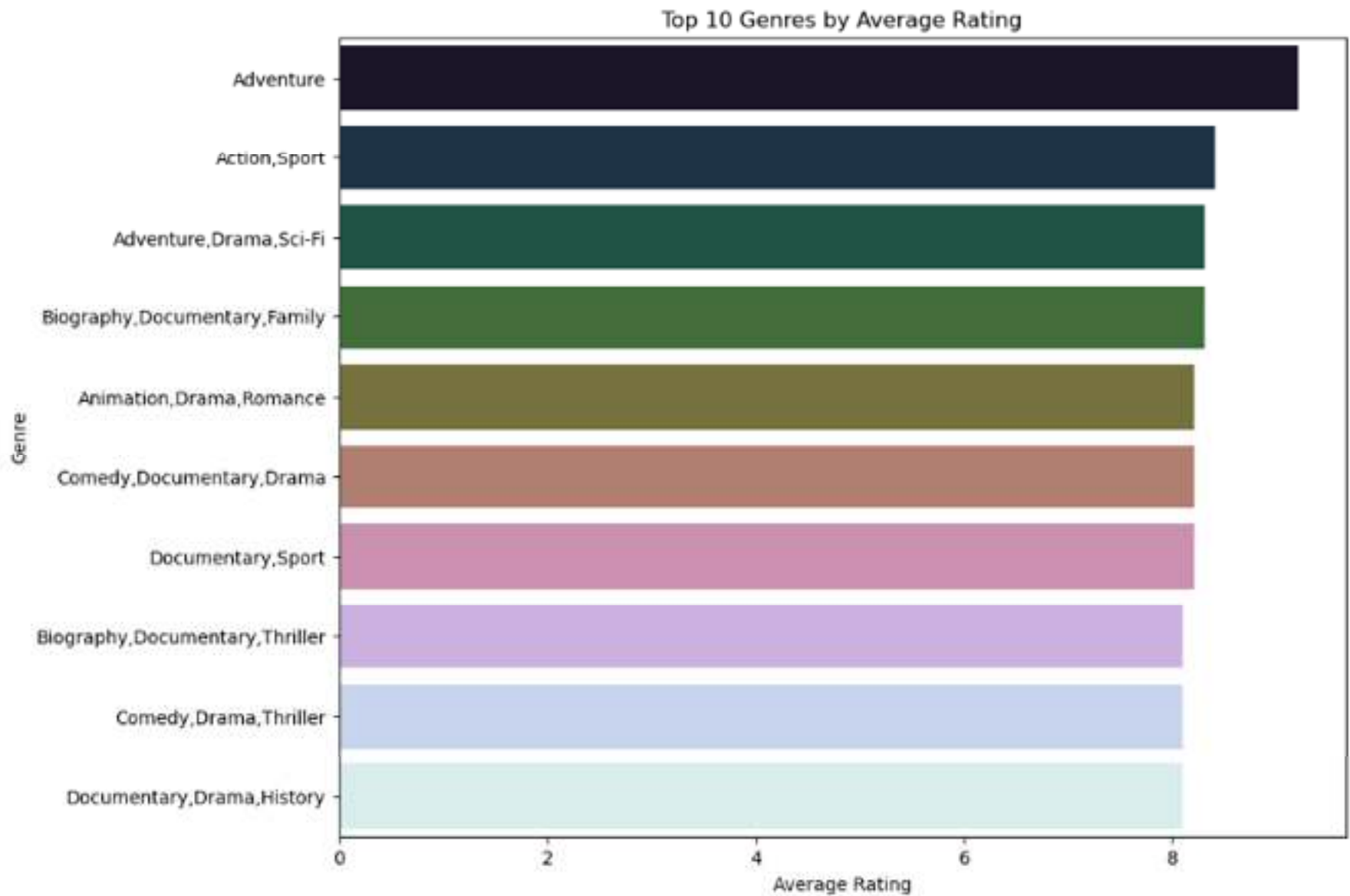
Project Analysis



DATASETS TO BE USED:

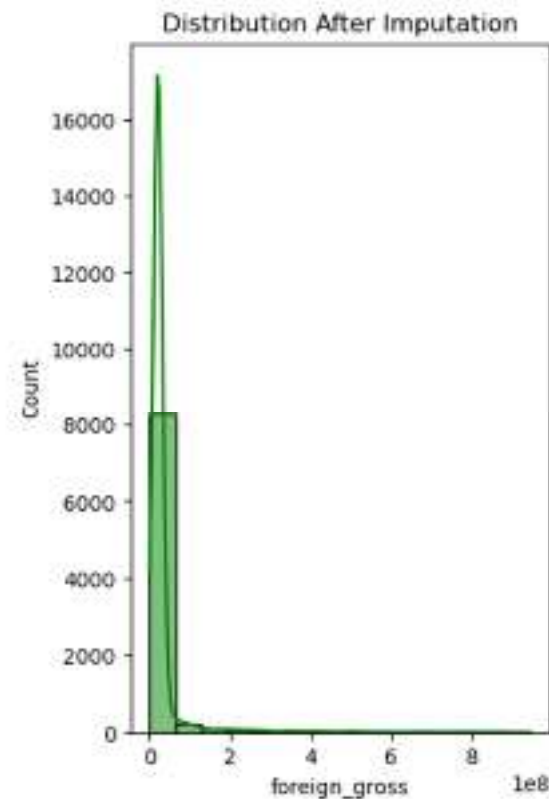
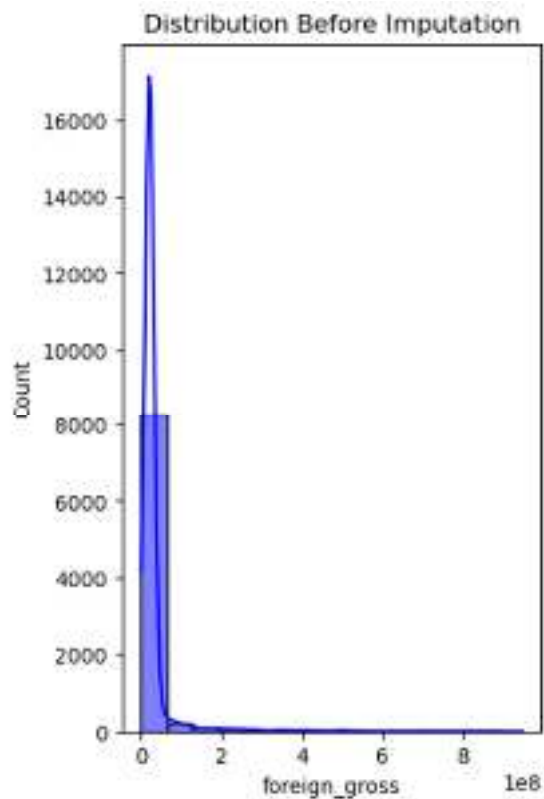
1. imdb.title.basics
2. imdb.title.ratings
3. bom.movie_gr
4. tn.movie_budgets.csvos

These are the datasets from which we will be extracting the data to analyze.



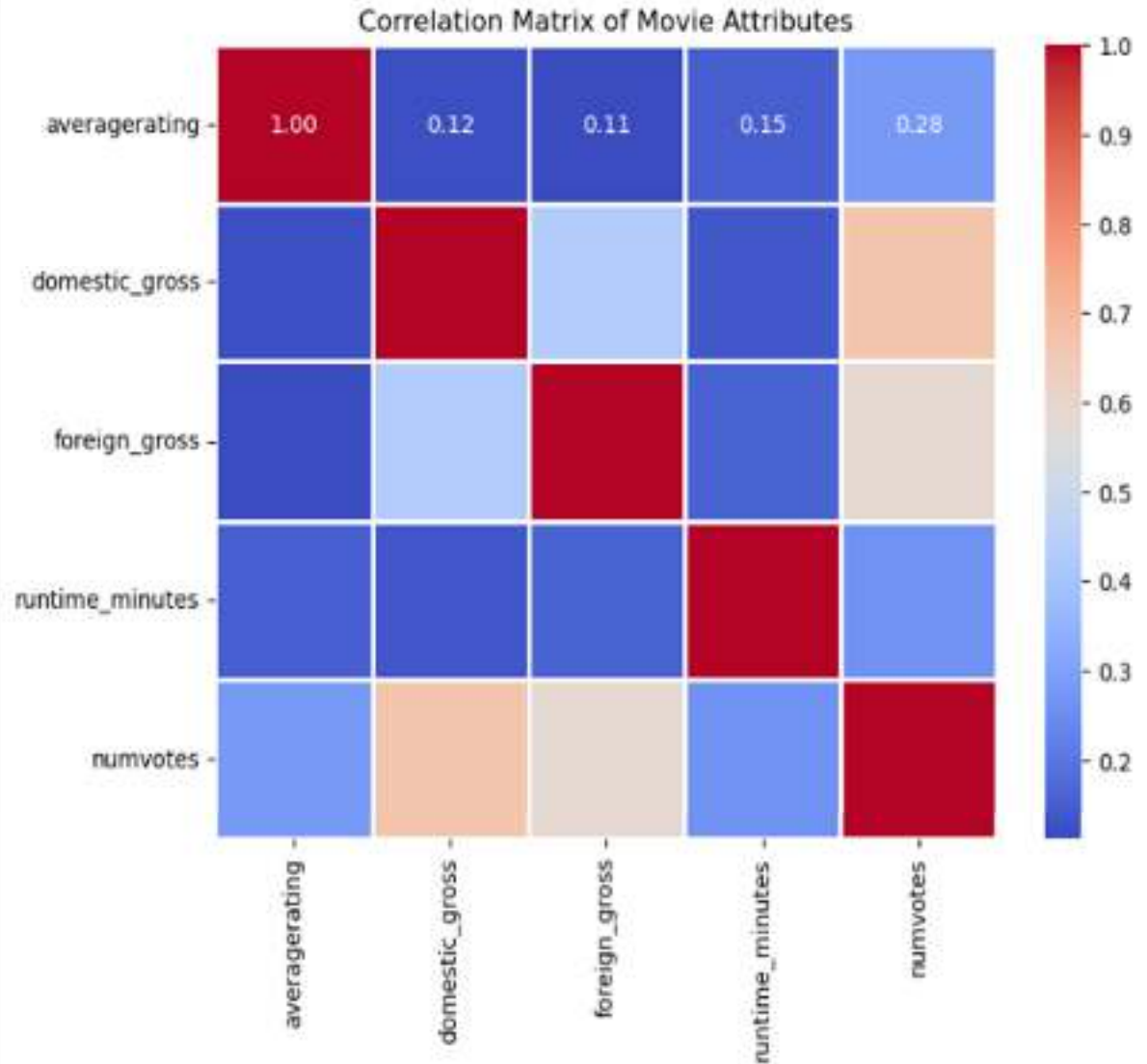
DATA IMPUTATION

- Data imputation is a statistical technique used to replace missing or incomplete values in a dataset with estimated or substituted values. The goal of imputation is to enhance the completeness and usability of the dataset, allowing for more accurate analysis and modeling. Missing data can occur for various reasons, such as measurement errors, data collection issues, or intentional non-responses.



CORRELATION ANALYSIS

Correlation analysis is a statistical technique used to evaluate the strength and direction of the linear relationship between two quantitative variables. The purpose of this analysis is to determine whether and to what extent changes in one variable are associated with changes in another variable. The result of a correlation analysis is a correlation coefficient, which quantifies the degree of correlation between the variables.



SUMMARY

THIS IS A SUMMARY OF THE FINDINGS OF THE ANALYSIS:

The Number of genres was 322

The Highest average rating genre: Adventure

The Lowest average rating genre: Comedy,Thriller

The Genre with highest domestic gross: Adventure,Drama,Sport

The Genre with lowest domestic gross: Comedy,

The Genre with highest foreign gross: Adventure,Drama,Sport

The Genre with lowest foreign gross: Biography,Documentary,Thriller

The Genre with longest runtime: Drama,History,

The Genre with shortest runtime: Action,Sport

The Genre with most numvotes: Adventure,Drama,Sci-Fi

The Genre with least numvotes: Documentary,Drama,Romance

The Best Performing Genre: Adventure,Drama,Sport

Thank You

E-MAIL: bryankinesh@gmail.com

GITHUB NAME: Bryanmunene

LINKED-IN: www.linkedin.com/in/brian-kariithi-0a5011192

