PHASE 2 PROJECT
FILM STUDIO RECOMMENDATONS

MIRRIAM MUMBUA

INTRODUCTION

Objective

The primary objective of this project is to provide actionable insights for a new movie studio aiming to enter the competitive film industry. By analyzing diverse datasets from sources like Box Office Mojo, IMDB, Rotten Tomatoes, TheMovieDB, and The Numbers, we aim to identify key trends and factors that influence box office success. Our analysis will focus on uncovering what types of films—considering genre, release timing, and other factors—are most likely to perform well and generate substantial profits.

Significance

The entertainment industry is highly competitive, with numerous variables affecting a film's success.

By leveraging data-driven insights, the studio can make strategic choices about which genres to focus on, optimal release timings, and how to allocate resources effectively.

This analysis will help mitigate risks, maximize returns, and position the studio to compete successfully in the market.

Business Problem

Understanding Film Performance

One of the primary hurdles is determining which types of films are most likely to succeed at the box office. With a plethora of genres, release strategies, and production budgets, making informed decisions is crucial for maximizing profitability and minimizing risk.

From the provided datasets, we can get information on the release dates , the movie titles, calculate percentage profit each movie genre, their studios,etc...

Giving us a foundation to start our understanding.

Goal:

Our goal is to provide actionable insights that will guide the new movie studio in making strategic film production decisions. By analyzing historical data on movie performance, we aim to uncover patterns and trends that can inform the studio's choices.

DATA UNDERSTANDING

DATA SOURCES

1. Box Office Mojo

-Compressed CSV (CSV.gz)

-Content: Movie box office performance data, including gross earnings, studio, title and release details.

2. TMDB

-SQLite Database

-Used only one table, Movie_basics:

    -Content: movie_id, primary_title, start_year, genres

3. The Number

-Compressed CSV

-Content: id, release_date, movie, production_budget, domestic and worldwide gross

DATA ANALYSIS

DATA CLEANING AND PREPROCESSING

1. Handling Missing Values:

- Used methods like .isna() and .sum() to find missing values.

- Dropped rows with missing critical values in columns related to box office performance.

2. Handling Duplicates

- Used drop_duplicates() methods to get rid of redundancy by using primary keys as the subset of elimination.

3. Data Integration:

- Combined datasets into a unified DataFrame for analysis, ensuring all relevant columns are included and properly aligned.

4. Adding needed columns in the dataframe.

- I had to perform some basic maths on specific columns and create new columns to store the results so I can get useful data e.g percentage profit.

5. Transforming Data Formats:

- Normalization: Converted financial figures from string formats with currency symbols to numeric values.

- Date Parsing: Transformed release dates into datetime format for time series analysis.

- Genre Handling: Split and normalized genre strings for consistent representation across datasets.

DATA ANALYSIS (EDA)PROCESS

- Created various plots (e.g., histograms, scatter plots) to identify patterns and trends.

- Computed summary statistics (mean, median, standard deviation) to understand data distribution.

- Examined relationships between variables, such as vote average and percentage profit.

RECOMMENDATIONS

RECOMMENDATION 1:

The bar chart reveals which genres have the highest average percentage profit, highlighting the genres that have historically performed well financially

Strategy: Invest in producing films within the genres that show the highest average percentage profit.

The top 5 being Horror, Mystery, Thriller, Sport, Romance all with over 600% average percentage profit.

RECOMMENDATION 2:

Average Percentage Profit by Release Month
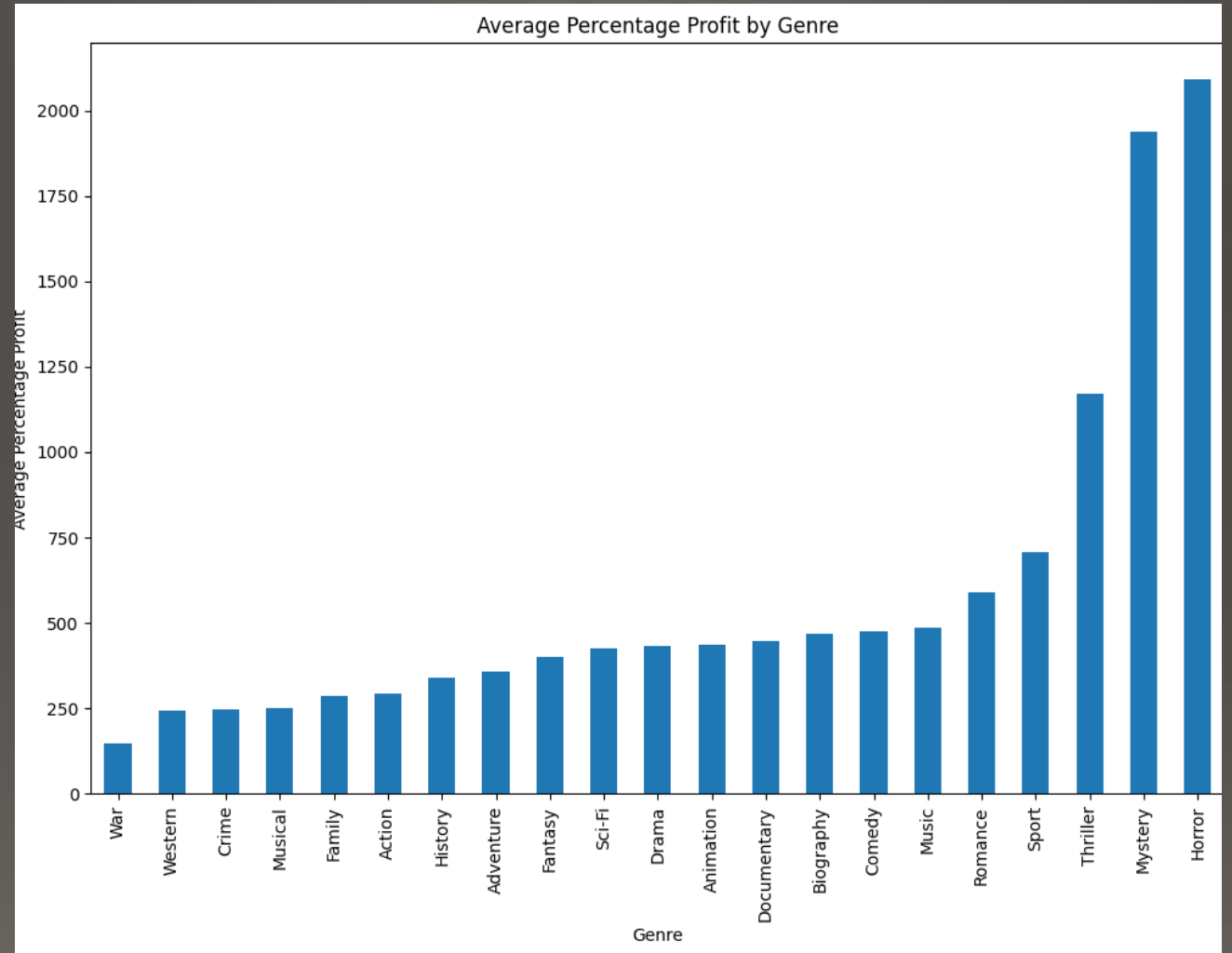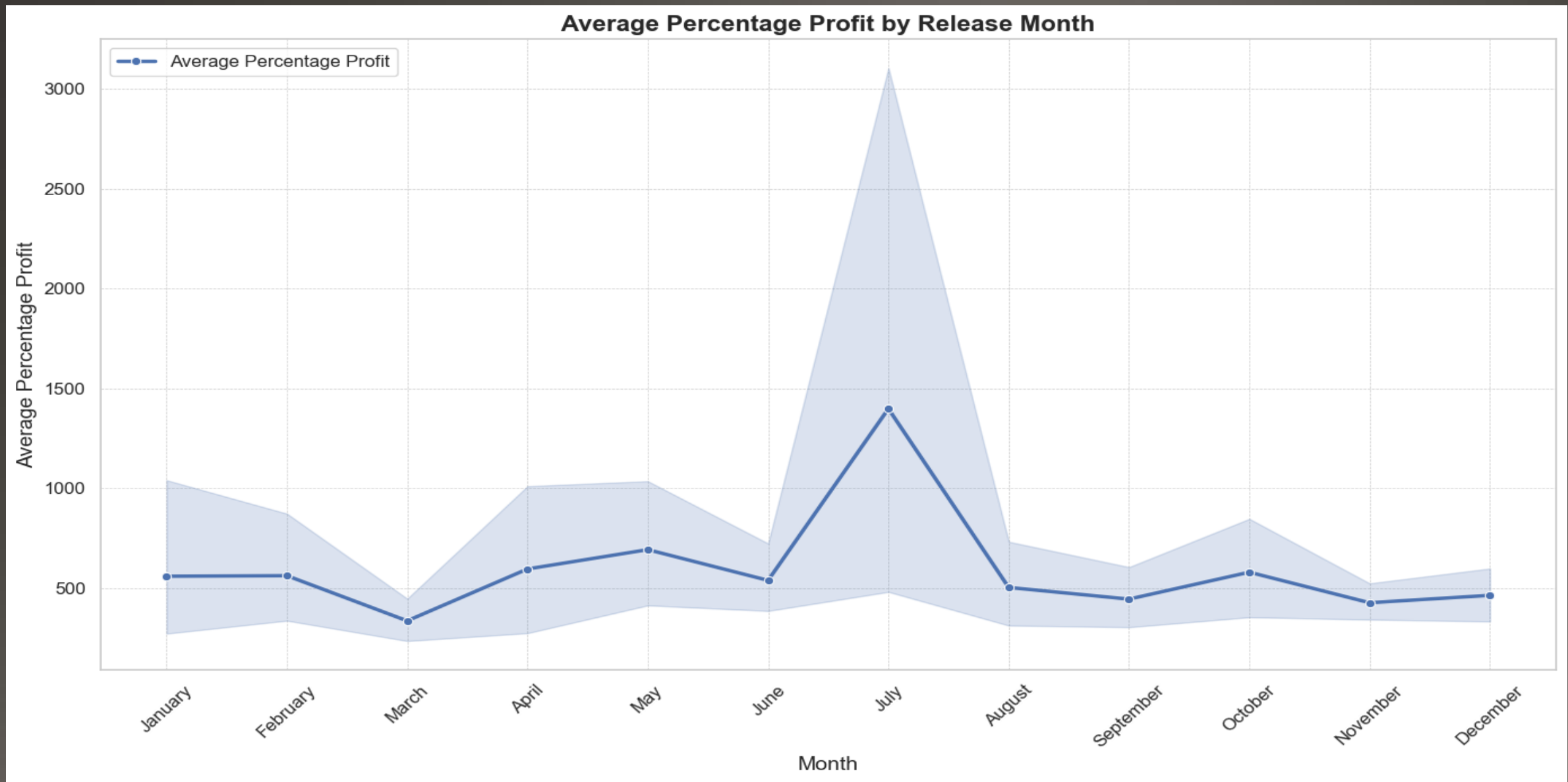
- The line plot reveals trends in average percentage profit by month, indicating which months tend to have higher profits for movie releases.

- **Strategy:** Plan new movie releases during the months that historically show higher average percentage profits.

- It is advisable to set to release the films during the month of july, since that's when the profit is maximized.

RECOMMENDATION 3:

- **Weak Correlation:** The scatter plot and Pearson correlation coefficient (-0.03) reveal a weak relationship between vote average and percentage profit.

- Most movies show varied profitability regardless of their critic scores.

- Critic ratings have little to no impact on a movie's financial success.

- **Strategy:** Invest in marketing, distribution strategies, and audience engagement as they may have a more substantial impact on profitability than critic reviews.

Correlation between Vote Average and Percentage Profit
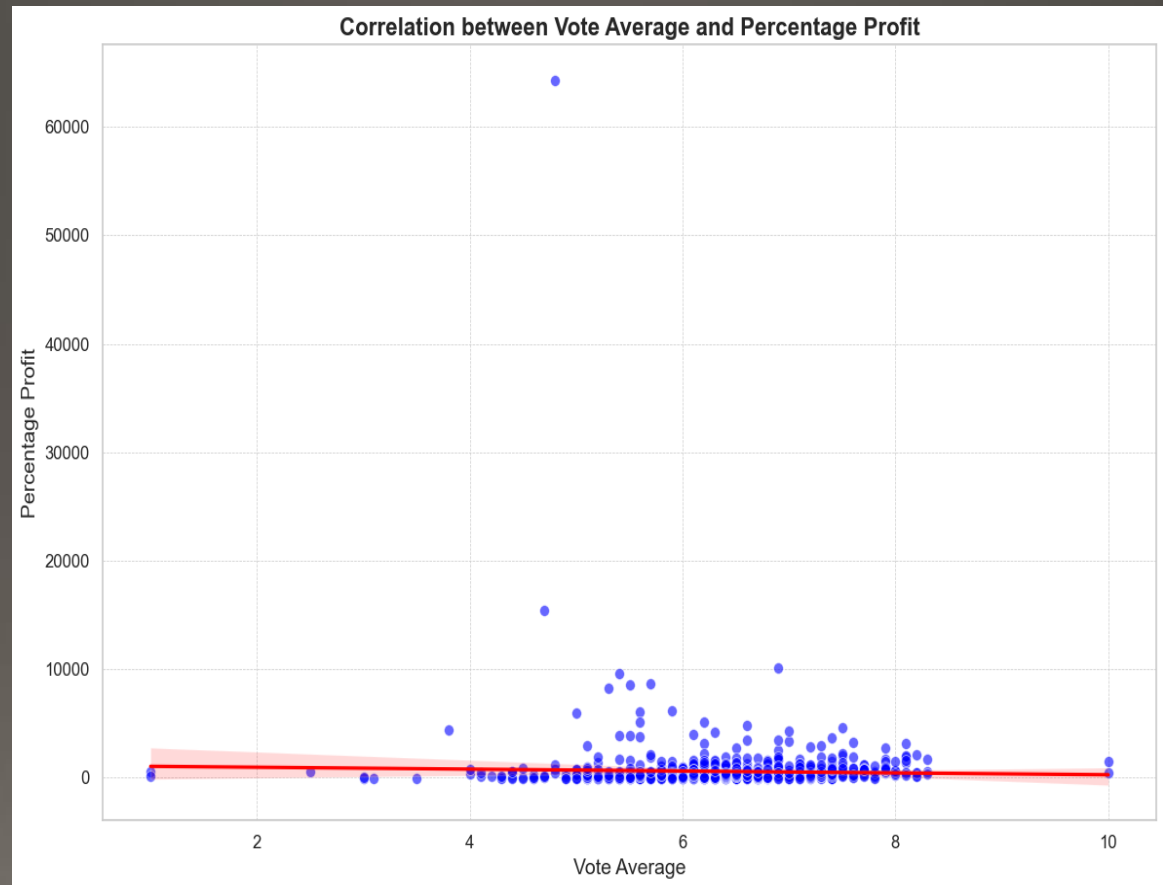
FUTURE WORK AND NEXT STEPS

1. Audience Demographics:

- Understand the characteristics and preferences of different audience segments.
- Collect and analyze data on age, gender, location, and viewing habits to tailor film content and marketing strategies.

2. Marketing Strategies:

- Evaluate the effectiveness of various marketing tactics.
- Gather data on promotional campaigns, social media engagement, and advertisement spends. Analyze the impact of these factors on box office performance.

3. Production Cost Optimization:

- Assess how production budgets influence profitability.
- Analyze the relationship between production costs and financial success. Explore cost-effective production methods and budget management practices.

CONCLUSION

- Analyzed movie data to provide actionable insights for a new movie studio.
- **Key Findings:**
- Genre Analysis.
- Seasonal Release Impact.
- Correlation Analysis.
- **Main Recommendations:**
- Prioritize producing films in genres with proven profitability.(Horror, Mystery, Thriller, Sport, Romance)
- Target release dates in months with historically higher average profits.(Around July mostly)
- Invest in marketing efforts and other factors rather than relying solely on critic reviews.

THANKYOU!

HAPPY FILMING! ☺

Please feel free to ask any questions or request further details about the analysis and recommendations.

Email: *mirriam.mumbua@student.moringaschool.com*