Prepared by group 5

Group Project

Movie Data Analysis

Group Members

Vallary Banda
Leader

Catherine wangui
member

Michael Omondi
member

Abdirizak Jarso
member

Introduction

● ● ● ● ●

────────────────────
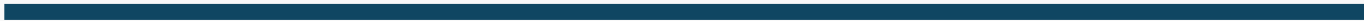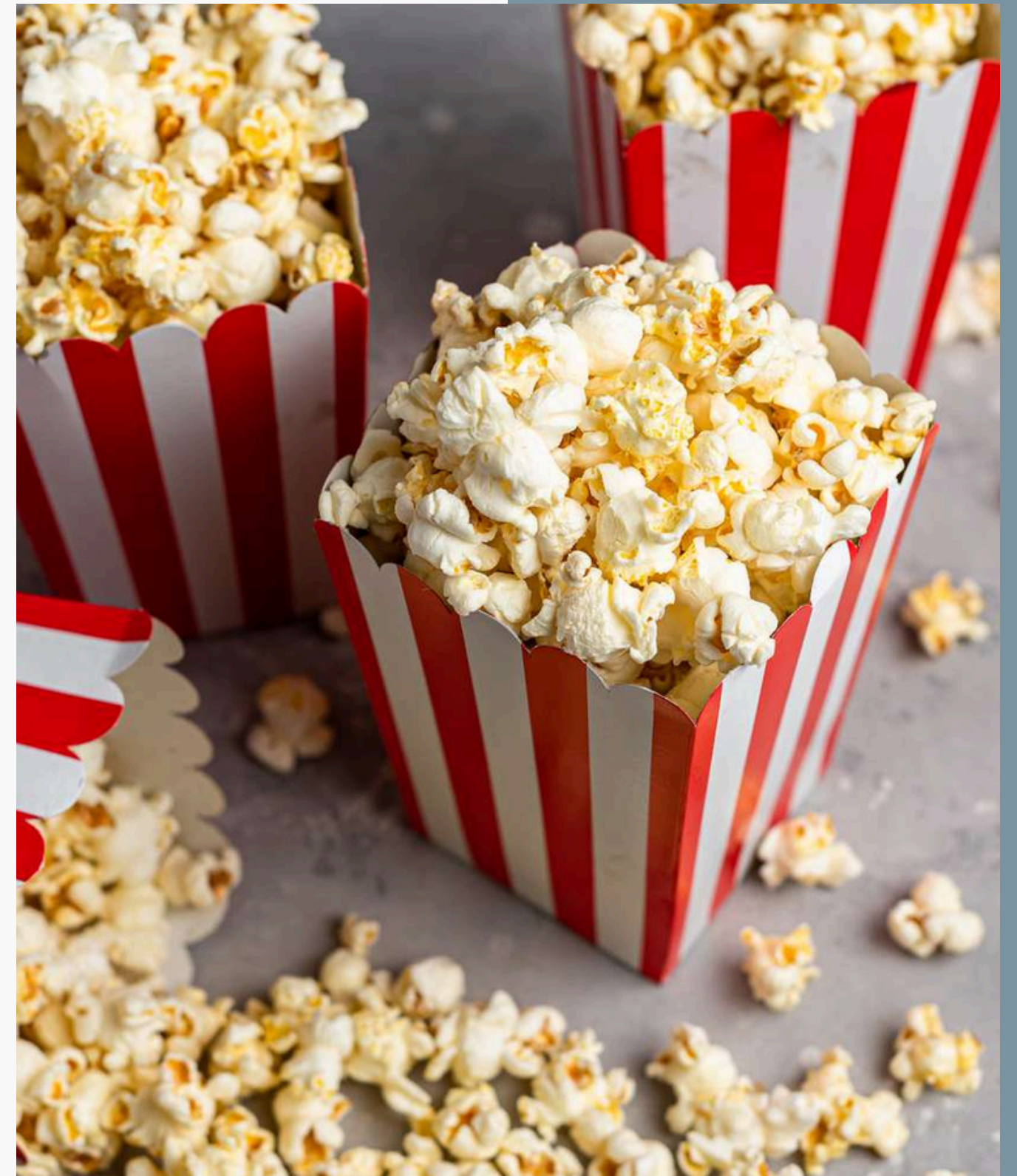
This project involves conducting an in-depth analysis aimed at generating actionable insights for stakeholders in the movie business. The analysis will explore various aspects of the movie industry, including financial performance, audience reception, and market trends. Our goal is to equip stakeholders with a comprehensive understanding of key metrics and trends, enabling them to make informed decisions that drive the company's growth and success in the competitive movie market.

────────────────────

● ● ● ● ●

Problem objectives

1. Identify types of films that are performing the best.

2. Explore and analyze current box office trends.

3. Identify factors that contribute to the success of a movie.

4. Translate findings into actionable insights.

DATA QUESTION TO BE ANSWERED

1. *Top-Performing Films:*
   - *Characteristics: Traits contributing to high box office gross.*
   - *Financial Success: Movies with the greatest financial success.*
2. *Popular and Successful Genres:*
   - *Trends: Popular and successful genres over time.*
   - *Audience Preferences: Content that resonates with target audiences based on ratings and reviews.*
3. *Impact of Budget on Revenue:*
   - *Budget vs. Revenue: How movie budgets affect box office revenue.*
   - *Profitability: Can smaller budget films be profitable? Identifying an optimal budget range.*
4. *Seasonal Trends in Performance:*
   - *Patterns: Seasonal trends in movie performance.*
   - *Optimal Release: Best times to release a movie to maximize revenue.*

DATA SOURCES

1. Box Office Mojo:
- Comprehensive resource for box office revenue data.
- Used to analyze the financial success of films.
2. IMDb:
- Information on TV shows, movies, and the entertainment industry.
- Includes cast and crew details, plot summaries, reviews, ratings, and user-generated content.
3. Rotten Tomatoes:
- Review aggregation site for movies and TV shows.
- Compiles professional critic and audience reviews to gauge reception.
4. TheMovieDB:
- User-driven database for movies, TV shows, and celebrities.
- Includes cast and crew details, plot summaries, user ratings, and reviews.
5. The Numbers:
- Detailed box office data and financial analyses of movies.
- Provides industry statistics, insights, home video sales, and movie budgets.

Data Exploration and Cleaning Process

Familiarizing with the Datasets:

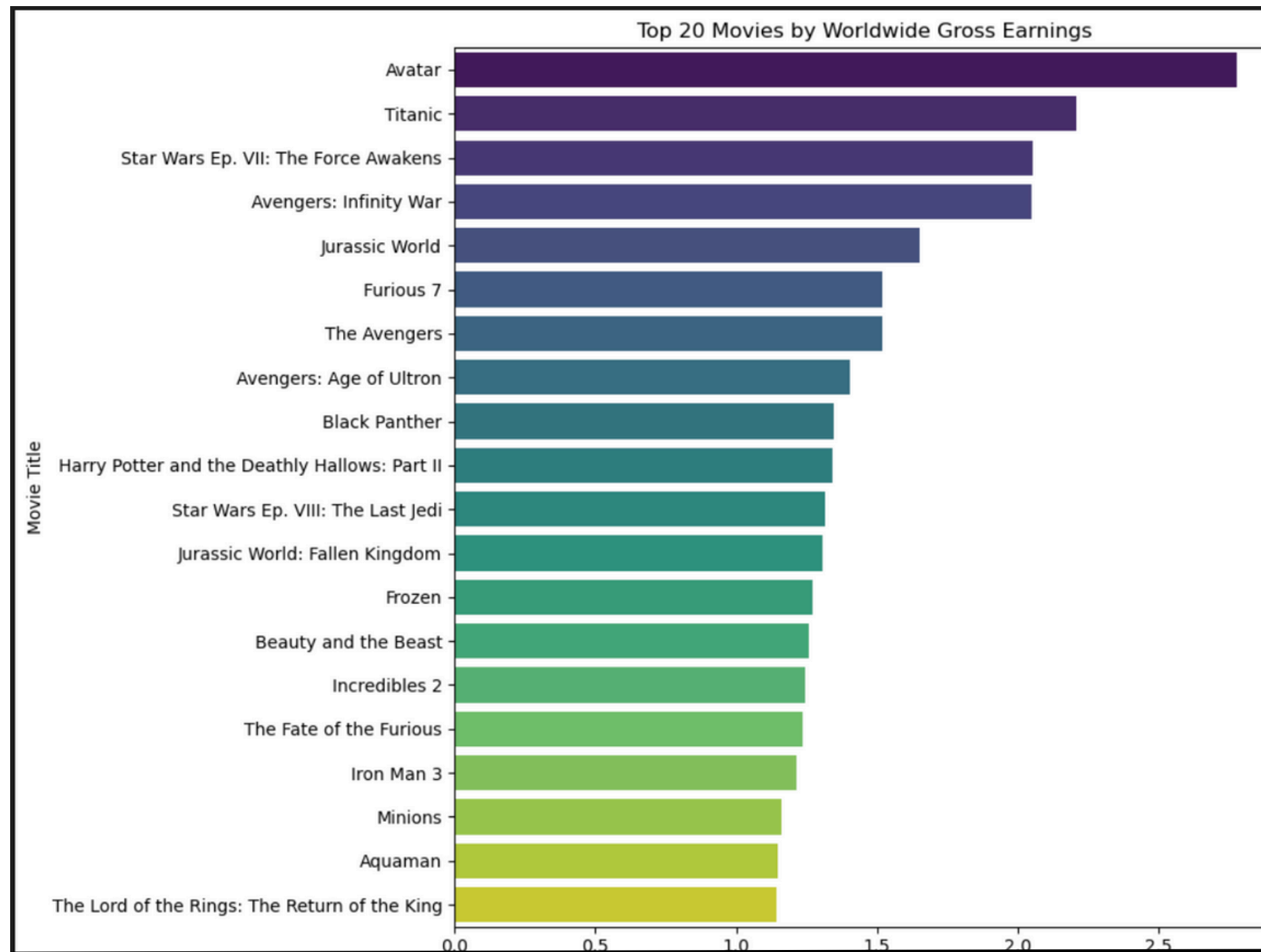- Explored various datasets and analyzed their columns.

Data Cleaning Methods

- Handling Missing Values:
    - Identified and managed missing data through imputation or removal.
- Data Transformation:
    - Converted data types and standardized formats for consistency.
- Removing Duplicates:
    - Checked for and removed duplicate records to ensure data uniqueness.
- Outlier Detection:
    - Identified and handled outliers using techniques such as z-score and IQR method.

Data Visualization Results
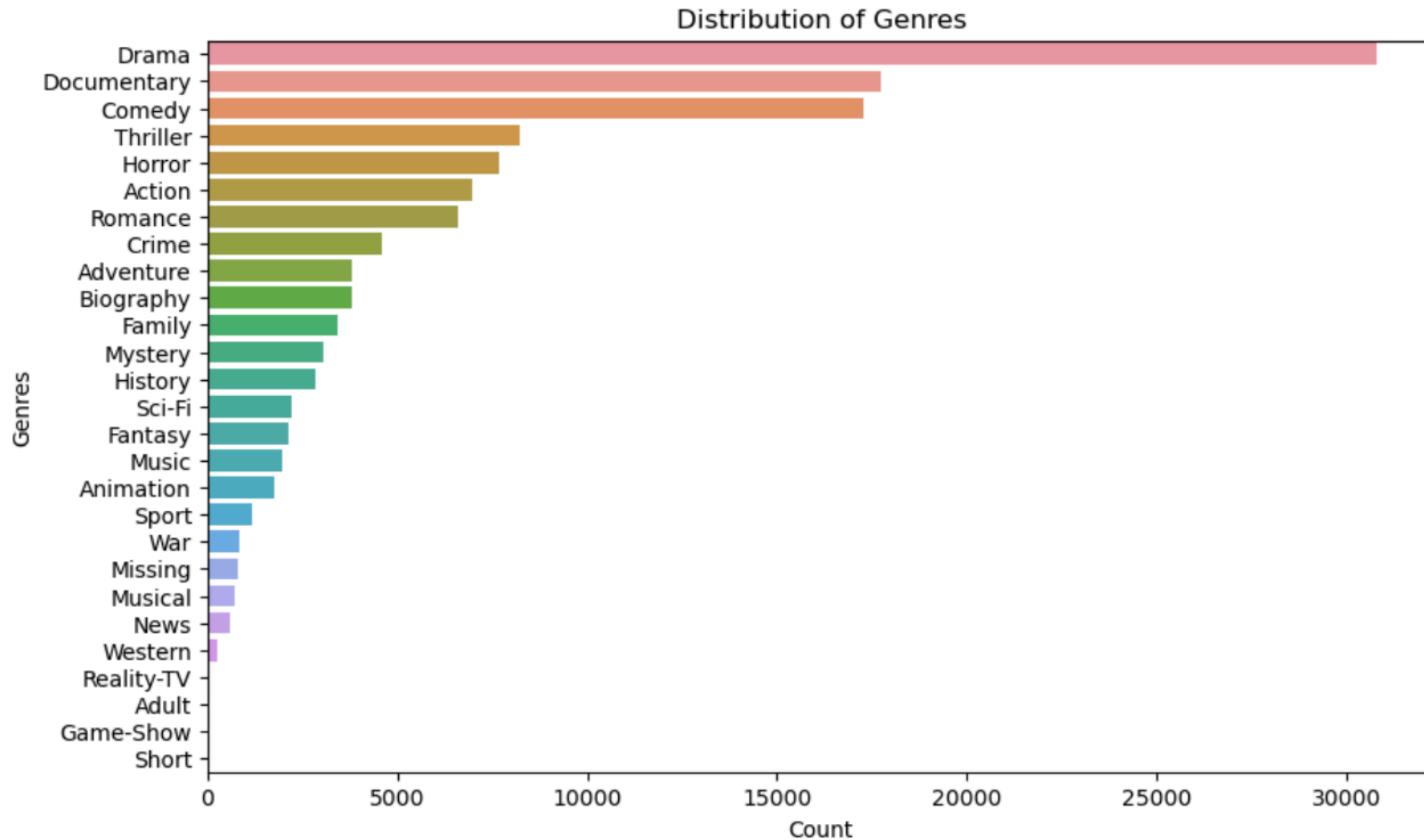# Top 20 movies by worldwide gross earnings

Top 20 Movies by Worldwide Gross Earnings

- The top three movies by revenue are Avatar with approximately $2.8 Billion followed by Titanic at $2.2 Billion and Star Wars with $2.0 Billion
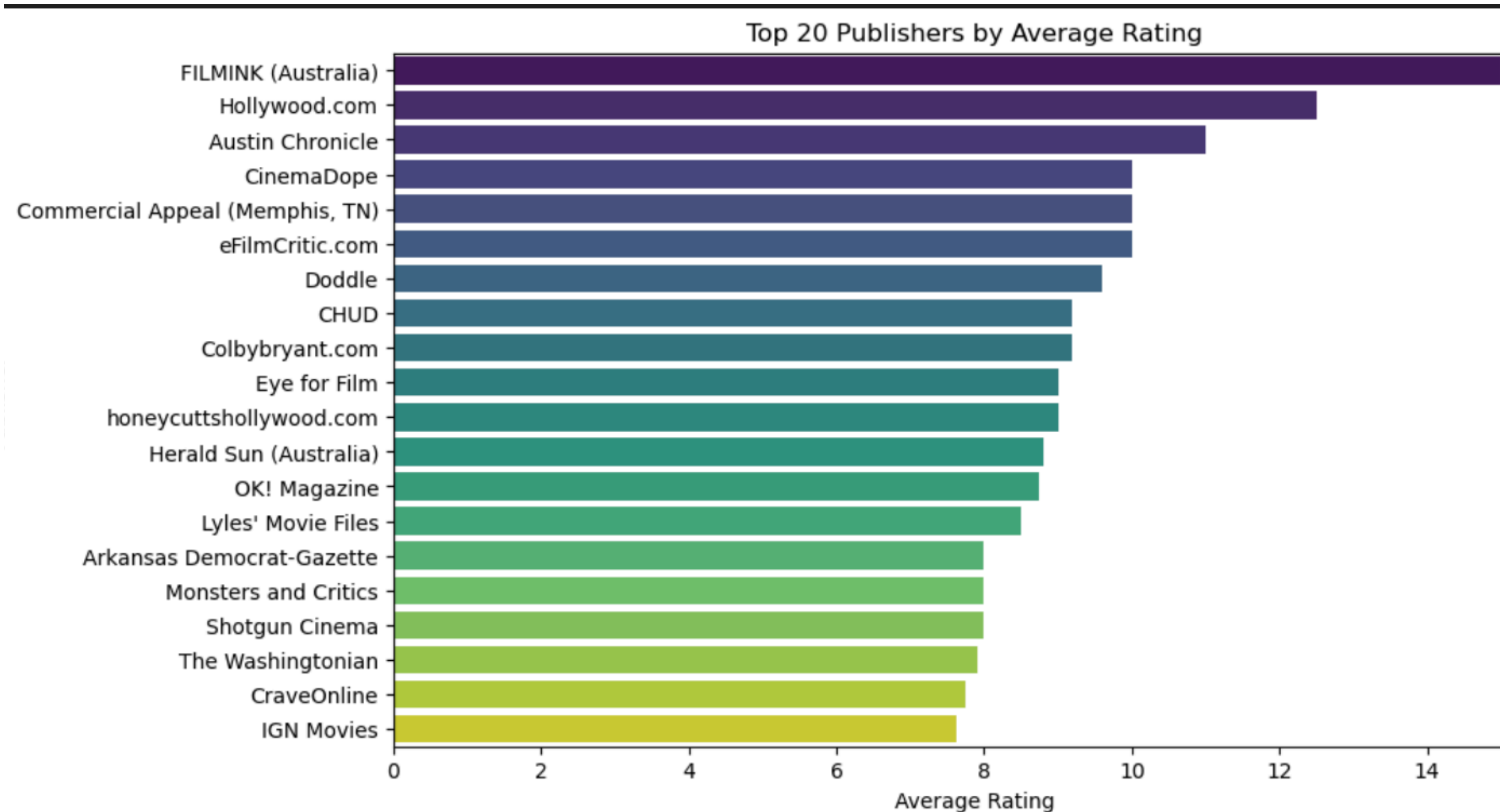
Data Visualization Results

Distribution of Genres

- The genres with the most distributed is Drama followed by Documentary and Comedy

Data Visualization Results
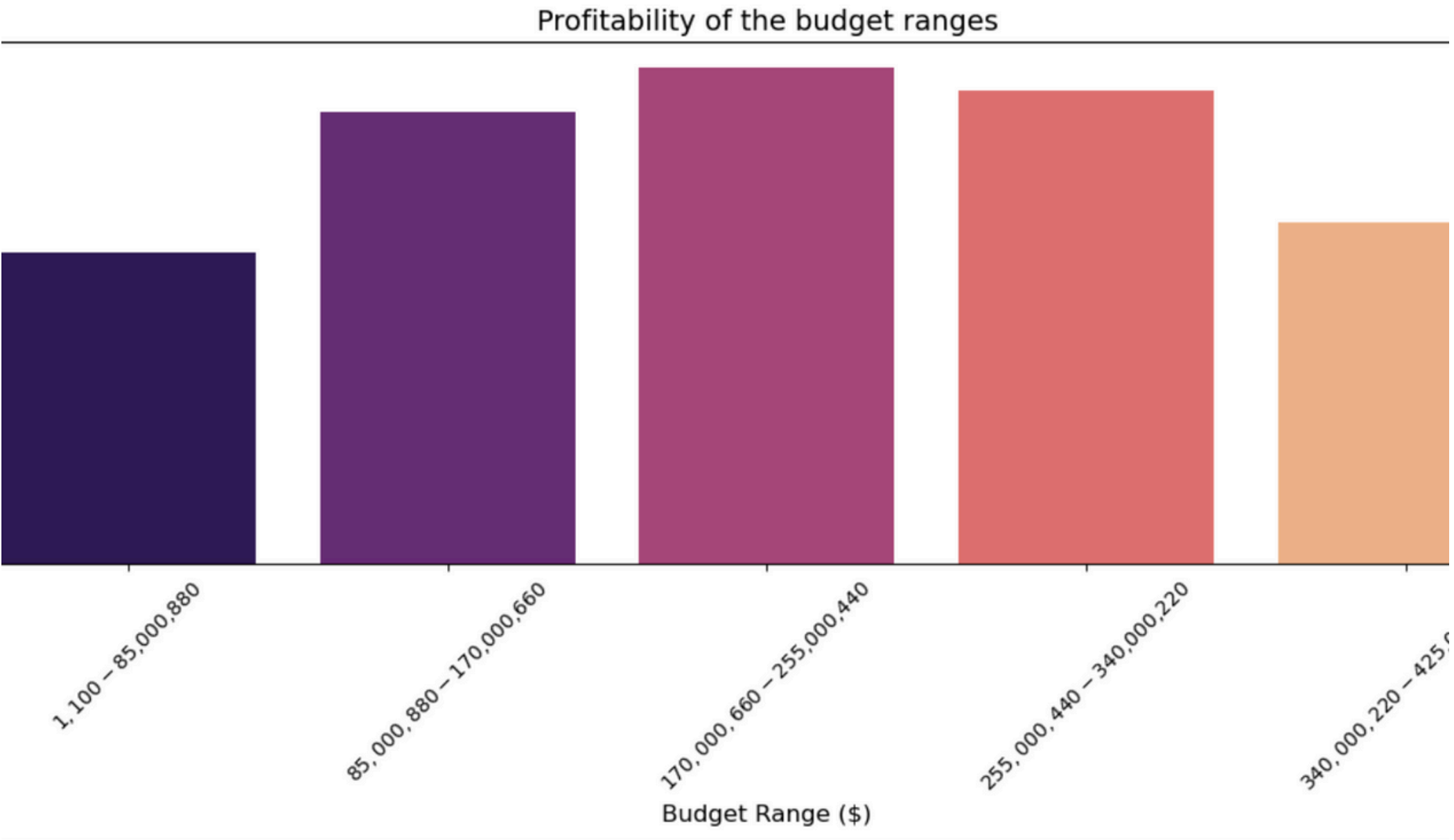
Top 20 publishers by Average Rating

Top 20 Publishers by Average Rating

From the bar graph the writer with the highest rating is FILMINK(Australia followed by hollywood and the publisher with list rating is IGN movies
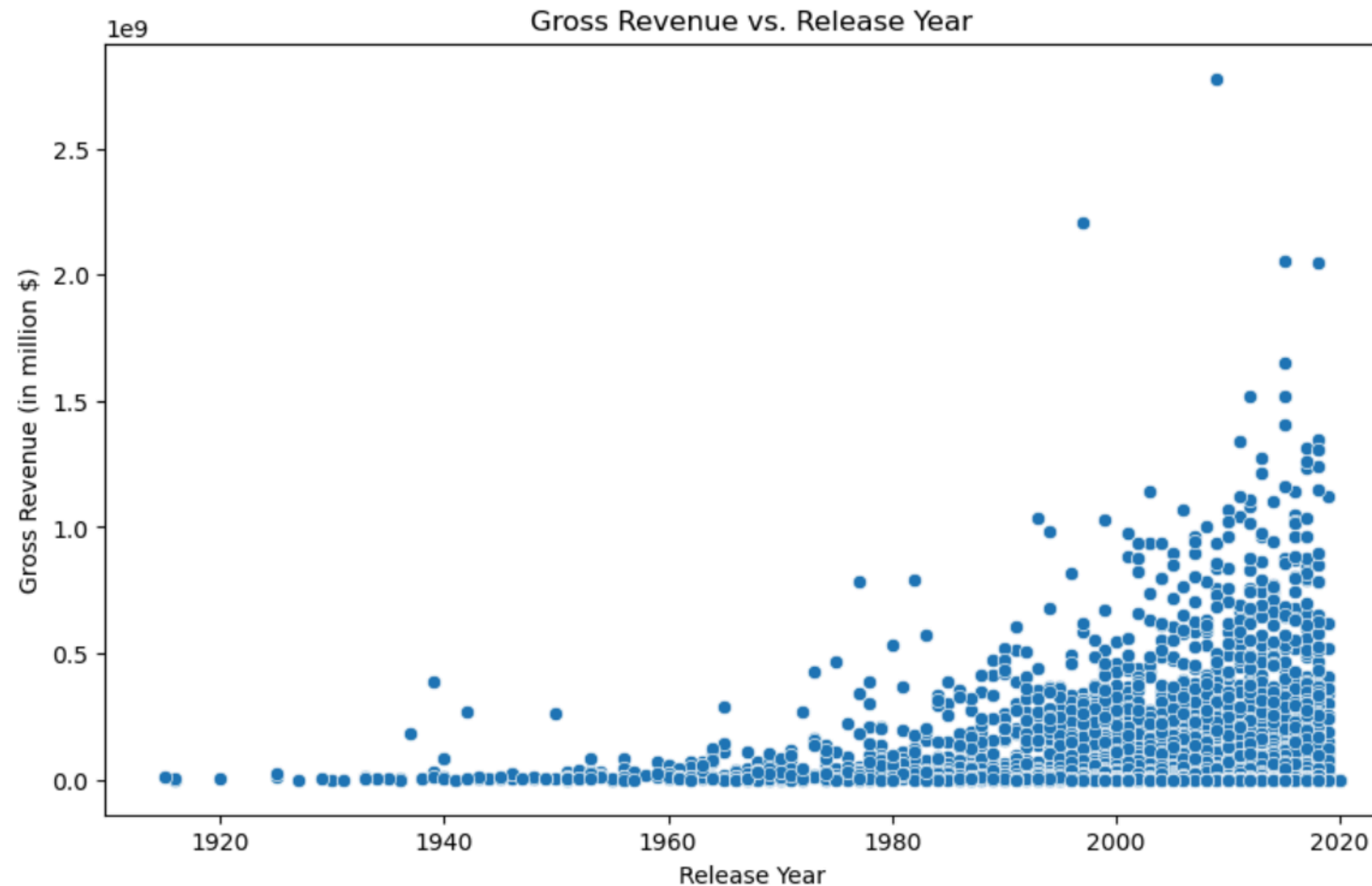
Data Visualization Results

Profitability of the budget ranges

Profitability of the budget ranges

The optimal budget range is 170,000,660 - 255,000,440 .With a profitability range of 95%
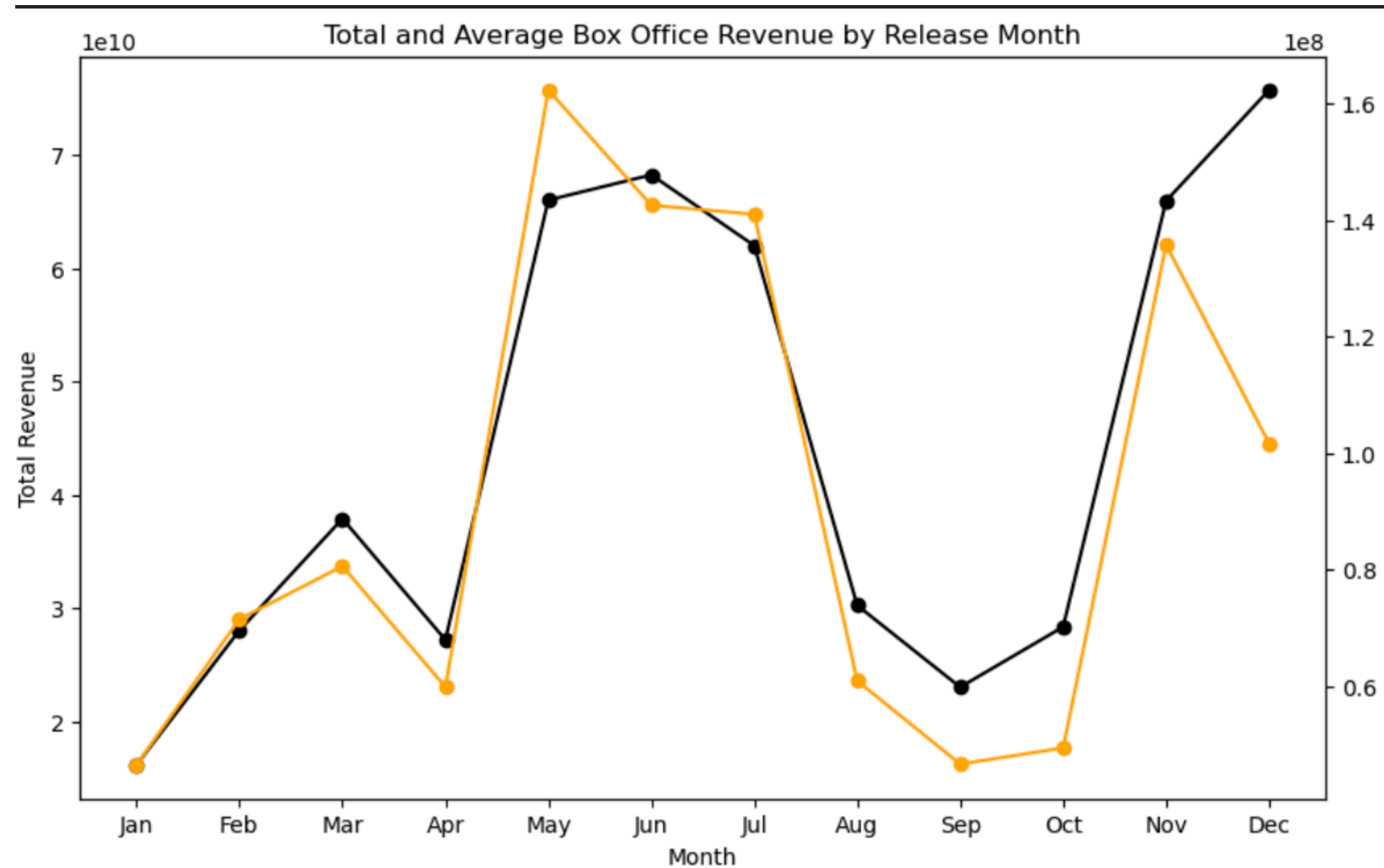
Data Visualization Results

Gross Revenue vs Release Year

The plot above shows an that the gross revenue from movies has generally increased from 1920 to 2020. This suggests that movies release in later years tend to earn more revenue.

Data Visualization Results

Total revenue release by Release month

Total and Average Box Office Revenue by Release Month

The ideal months to release movies are:

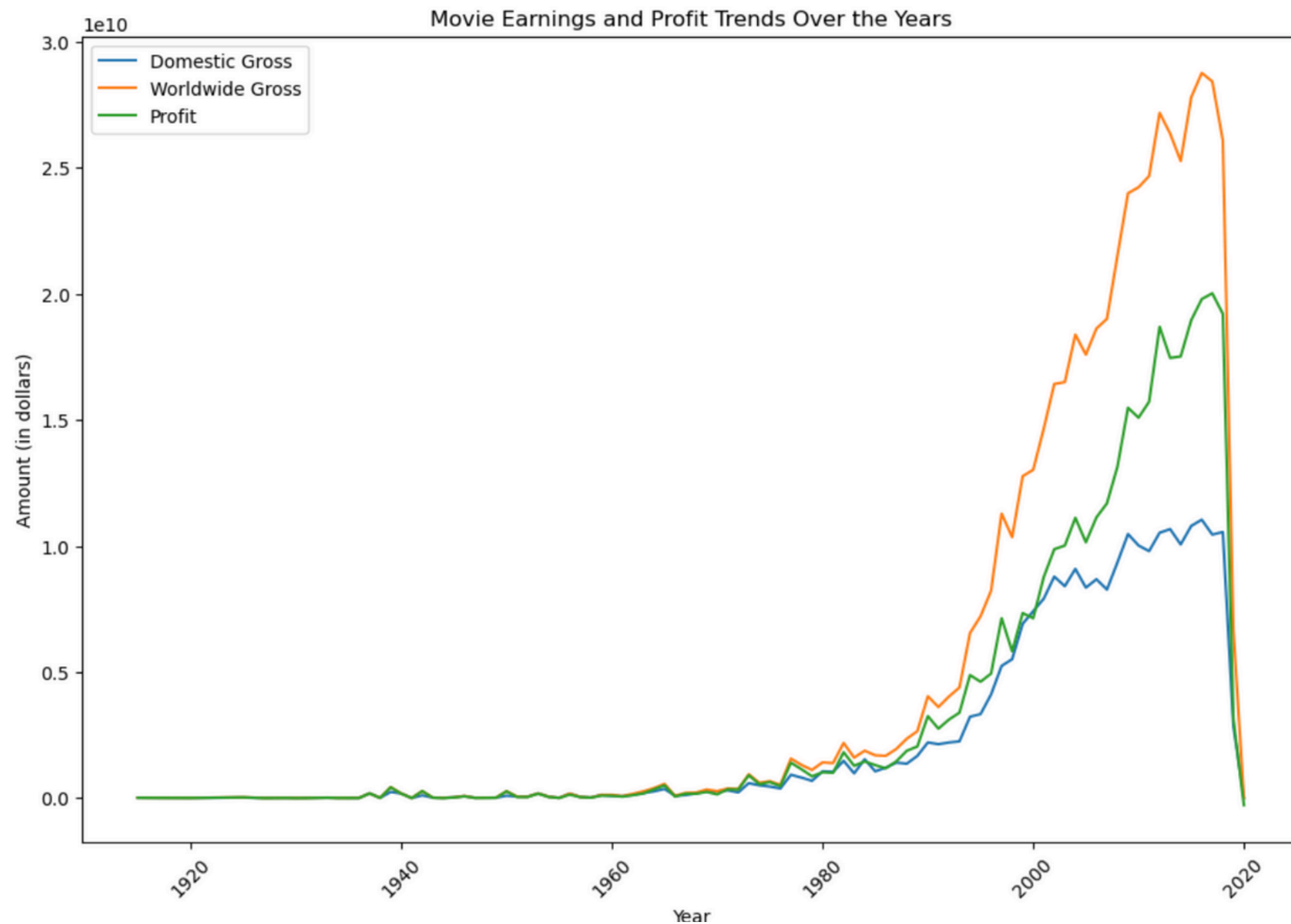**May** has the highest average worldwide gross earnings,
June and then **July**
* These months could be considered the best time to release a movie for maximum revenue as they generate the highest worldwide gross revenue on average.
December has the highest total revenue

Data Visualization Results

Movie earnings and profit Trends over the years

Over time, we notice `domestic_gross`, `worldwide_gross` and `profit` have been generally increasing.

Recommedation

1. The top most profitable movies are action, adventure, Animation, Comedy, Drama, 'Fantasy, Horror, Romance, Sci-Fi, and Thriller we would advise the stakeholders to focus on these genres

2. The top 5 most popular movies are all from the action genre so we would recommend the stakeholders focus more on action genres

3. Target production budgets to be within the optimal budget range to maximize on profitability.

4. Plan the release of movies to fall in the months of May, June, July, November, and December.

Thank you