MICROSOFT MOVIE STUDIO FILMS ANALYSIS.

Author: Brian Kipruto Ngeny.

Overview.

Microsoft Software Company is creating a movie studio and thus, they want an analysis of the films that are currently doing the best on box office so as to gain insight into the types of films that are currently marketable. This project analyzes the types of films currently doing the best on Box Office to find actionable insights that the head of Microsoft's new movie studio can use to help decide what type of films to create. Exploratory Data Analysis and Evaluation shows that some of the highest rated genres are; Adventure, Drama, Sci-Fi, Comedy, Action and Romance which are also among the most profitable genres, indicating that they are the best performing film types currently.

Business Understanding

I)   Business Problem

Microsoft sees all the big companies creating original video content and they want to get in on the fun. They have decided to create a new movie studio, but they don't know anything about creating movies. Like any new business venture, market analysis is an important step in the establishment of the business. It provides crucial information on market trends, peak seasons and recommendable product or services that can give a business a niche over the competition. To address this issue, I plan to use exploratory data analysis (EDA) to provide the head of Microsoft movie studio with actionable recommendations and insights into the type of films the studio should focus on producing so as to take the movie industry by storm and avoid making too many loses that may lead to the shutting down of the movie studio.

II)   Objectives.

The goal of this project is to:

1. Analyze the findings of the exploratory data analysis model (EDA) and provide the company with actionable insights that answer the following specific business questions;

    i.   What is the mean, minimum, maximum and position of outlier amounts for movie gross return values?
   ii.   Does the popularity and original language of a movie affect its rating?
 iii.   How does the production budget amount translate to the gross returns?
  iv.   Do higher movie ratings translate to higher gross returns?
   v.   Which are the highest rated and most profitable genres of movies currently?

Data Understanding.

The data used was sourced from:

- [Box Office Mojo](#)
- [IMDB](#)
- [Rotten Tomatoes](#)
- [The MovieDB](#)
- [The Numbers](#)

These data has been read from their respective files and loaded into data frames which go by the followings names within the project:
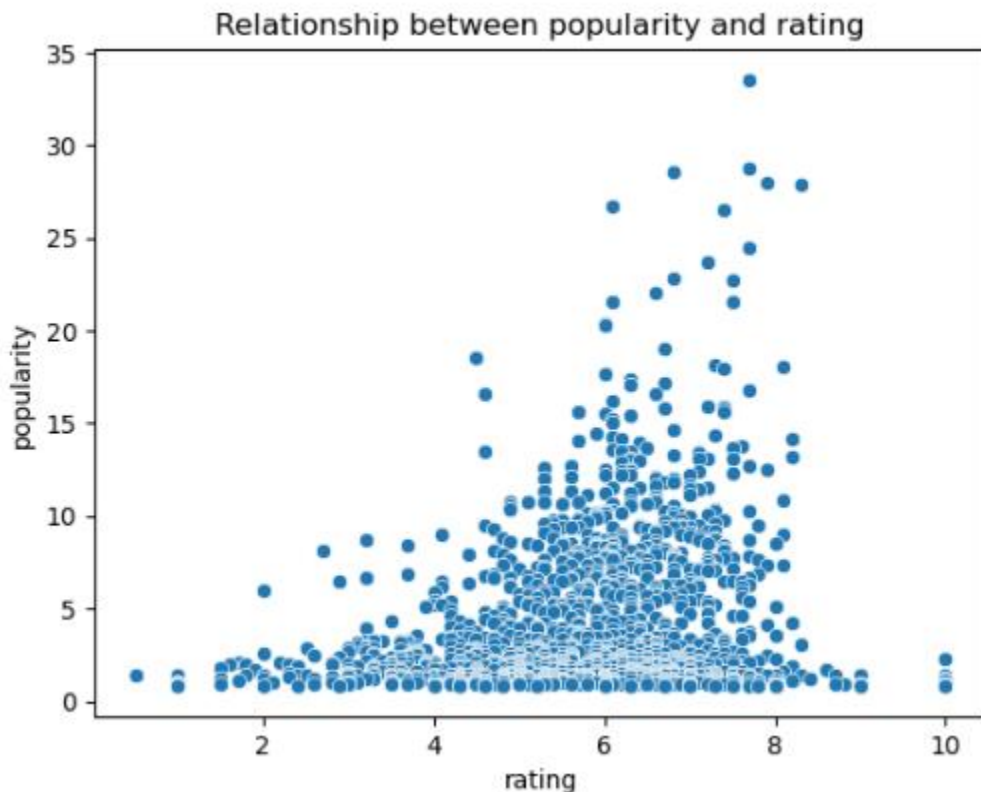
- **bom_data** sourced from Box office mojo.
- **tmdb_data** sourced from The MovieDB.
- **tn_movie_budgets** data sourced from The Numbers.
- **movie_ratings** table sourced from IMDB database.
- **movie_basics** table sourced from IMDB database.
- **imdb_data** table. This dataset is created by merging the two tables from the IMDB database into one table using the movie_id as the unique identifier.
- **merged_data** table. This dataset is created by merging the imdb_data table with the bom_data table into one table using the title as the unique identifier.
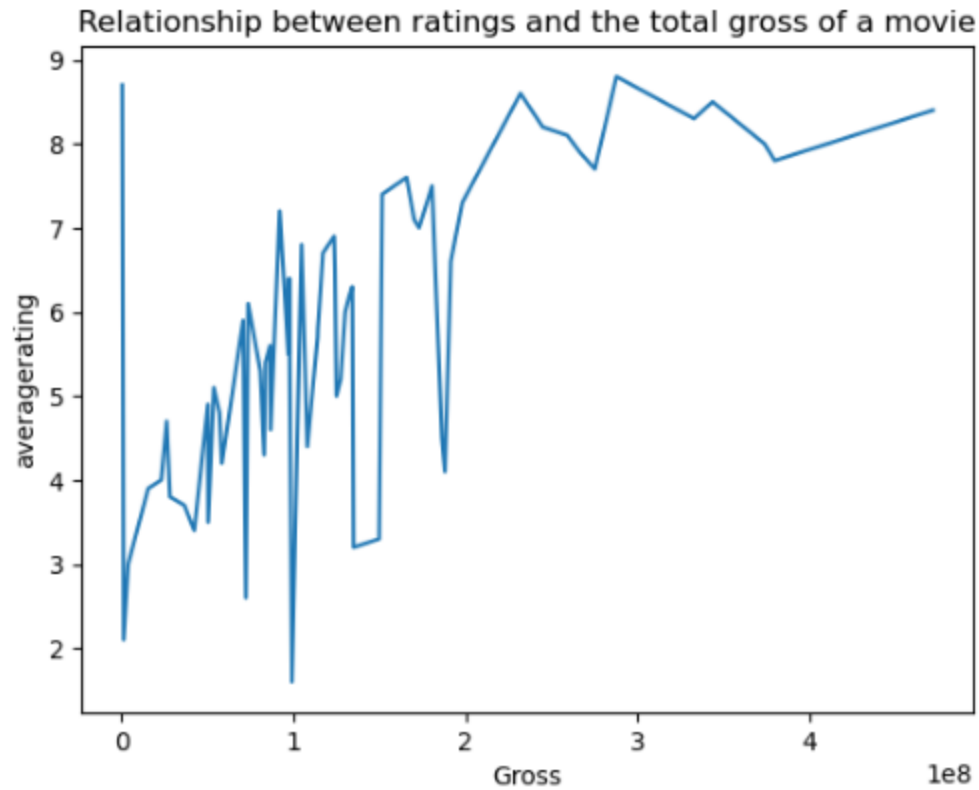
Data Analysis results.

1. The statistical values for the data provided on the production budgets, domestic gross and worldwide gross are as follows:

    a) The average amount for the production budget of a movie is 31,587,757 dollars. The lowest amount on record is 1,100 dollars and the highest amount on record is 425,000,000 dollars.

    b) The average amount for the domestic gross of a movie is 41,873,326 dollars. The lowest amount on record is 0 dollars and the highest amount on record is 936,662,225 dollars.

    c) The average amount for the worldwide gross of a movie is 91,487,460 dollars. The lowest amount on record is 0 dollars and the highest amount on record is 2,776,345,279 dollars.
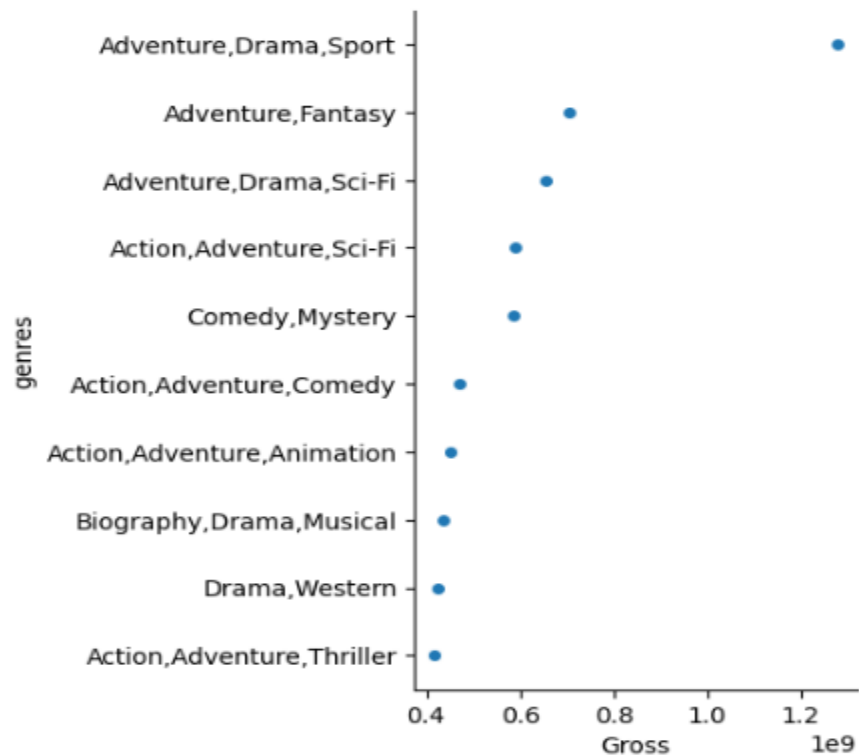
2. There is no correlation between popularity and rating as evidenced by the scatter plot displayed below:

Relationship between popularity and rating

3. There is however a close relationship between the rating of a movie and the gross revenue it generates as shown below:

Relationship between ratings and the total gross of a movie

4. The top most profitable genres are; Adventure, Drama, Sci-Fi, Comedy and Action movies as evidenced below:

Conclusions.

Based on the analysis done, this are the conclusions that can be drawn up:

I. A movie with an average production budget of close to 30,000,000 dollars is expected to return an approximate total gross amount of 100,000,000 dollars. The good news is that, the box plot shows that most of the outlier values are found above this approximate value, meaning that a movie is more likely to surpass its approximated gross return value.

II. The original language of a movie also has a bearing on the rating and popularity of a movie. English movies are by far the most produced movies in the dataset analyzed but are very lowly rated and not as popular compared to Japanese, French and Spanish movies which are the highest rated and most popular movies in terms of the original languages.

III. Among the top ten most produced genres of movies, Drama, Comedy and Romance are the most recurrent genres. These three genres also appear more times than any other genre in the top five highest rated list of most produced genre of movies.

IV. After merging the dataset on ratings and the dataset on gross returns, I concluded that some of the genres that fetch the highest amount of gross returns are, Adventure, Drama, Sci-Fi, Comedy, Action, Fantasy and Sport.

Recommendations.

The Exploratory Data Analysis and conclusions drawn from the analysis lead to these recommendations for the Microsoft movie studio:

- **The production budget allocated to an individual movie, should on average be between 25,000,000 dollars and 35,000,000 dollars.** This will help in mitigating the risk factor involved in any business venture such as this one. If the movie is a success, it will raise upwards of 100,000,000 dollars gross revenue but if it fails it should break-even and generate approximately the same amount as the production budget.

- **The Studio should focus on producing content with highly rated features e.g. highly rated original languages and highly rated genres.** From the analysis done, there is a close correlation between the rating of a movie and the gross revenue it generates. Combining highly rated features in movie production should result in a good, quality movie that is bound to be highly rated and thereafter high revenue returns.

- **The studio should prioritize these genres; Adventure, Drama, Sci-Fi, Comedy, Action and Romance.** These genres have been determined from the analysis as the highest rated

genres and the most profitable ones. Creating content found within these genres and even combinations of these genre types is an assured tactic of making the studio a success.

Thank you very much for your consideration and for your time.