

## Microsoft Movie Analysis

Author : Mark Bundi Kimathi

## Overview

The project's goal is to assist Microsoft's new movie studio in making informed decisions on the types of films to create in order to maximize box office success. The process involves an exploratory analysis of box office trends to identify profitable genres and translate those findings into actionable insights for the head of the movie studio.

## Business Problem

The business problem identified in the project is that Microsoft's new movie studio lacks the knowledge and expertise in creating movies, which puts them at a disadvantage in the competitive film industry. They need to make strategic decisions on the types of films to produce in order to maximize box office success and establish themselves as a significant player in the entertainment market.

## Data Understanding

Describe the data being used for this project.

---

The data analysed came from IMDb website. IMDb (an acronym for Internet Movie Database) is a popular worldwide online database of information relating to all movies, television programs, video games and streaming content online. I used 2 files from IMDb to answer the question of which genres were most successful got the most ratings.

The data from Box Office Mojo was added to the IMDB data to determine which genres are most financially successful.

After checking the information on each table to see column names and null values, I joined the two datasets, `df_titles` and `df_movie_ratings` together using the 'tconst' column as it was a unique identifier creating a new dataframe called `df_joined_data`. I then joined the dataset `df_movie_gross` with the new dataframe using the title as the unique identifier, creating a combined new dataset called `df_joined_data2`.

The data used for rating is `Ratings_data` from `df_joined_data`. The data used for monetar statistics is `Gross_data`.

## . Methods

I used the Ratings\_data to determine which genres were highest rated. The data was grouped by genre on the mean of the averarating column. This calculates the mean rating per genre.

The Gross\_data , domestic and foreign columns were made into type integer. Then the data was grouped by genre and domestic and foreign gross independently. Each had its own bar chart showing how successful they are in each market.

## Result

The data showed the highest rated genres are:

1. Comedy,Documentary,Fantasy
2. Documentary,Family,Musical
3. History,Sport
4. Game-Show
5. Music,Mystery

While the most economicaly successful genres in the foriegn and domestic markets are:

1. Action,Adventure,Sci-Fi
2. Adventure,Animation,Comedy
3. Action,Adventure,Fantasy
4. Action,Adventure,Comedy
5. Action,Adventure,Thriller
6. Comedy

## Conclusion

---

This analysis leads to three recommendations regarding types of movies that are successful:-

Movies with the genre combination Action, Adventure & Sci-Fi topped the leaderboard in both Domestic and Foreign Gross Sales, this combination is obviously a hit at the box office worldwide, make this the first type of movie to produce for success.

Movies with Adventure, Animation and Comedy were the next most successful in Foreign Gross Sales and Domestic Sales, use this combination as the next or alternative type of movie to produce.

The audiendce rating shows Comedy,Documentary,Fantasy got the highest ratings.

From the finding it is concluded that Adventure genre seems constant in all the above graphs so must be a key element of any movie to be produced.

