

IMDB MOVIE ANALYSIS

AFEEFA SIDDIQUA



PROJECT DESCRIPTION

- The goal of this project is to investigate the factors that influence the success of movies on IMDB, where success is defined by high IMDB ratings.
- Understanding these factors is crucial for movie producers, directors, and investors who aim to make informed decisions to enhance the quality and success of their future projects.
- **Tech-stack Used:** MS Excel

Approach:

The data set contained 5043 rows and 28 columns.

Data Cleaning:

- **Duplicate Records:**

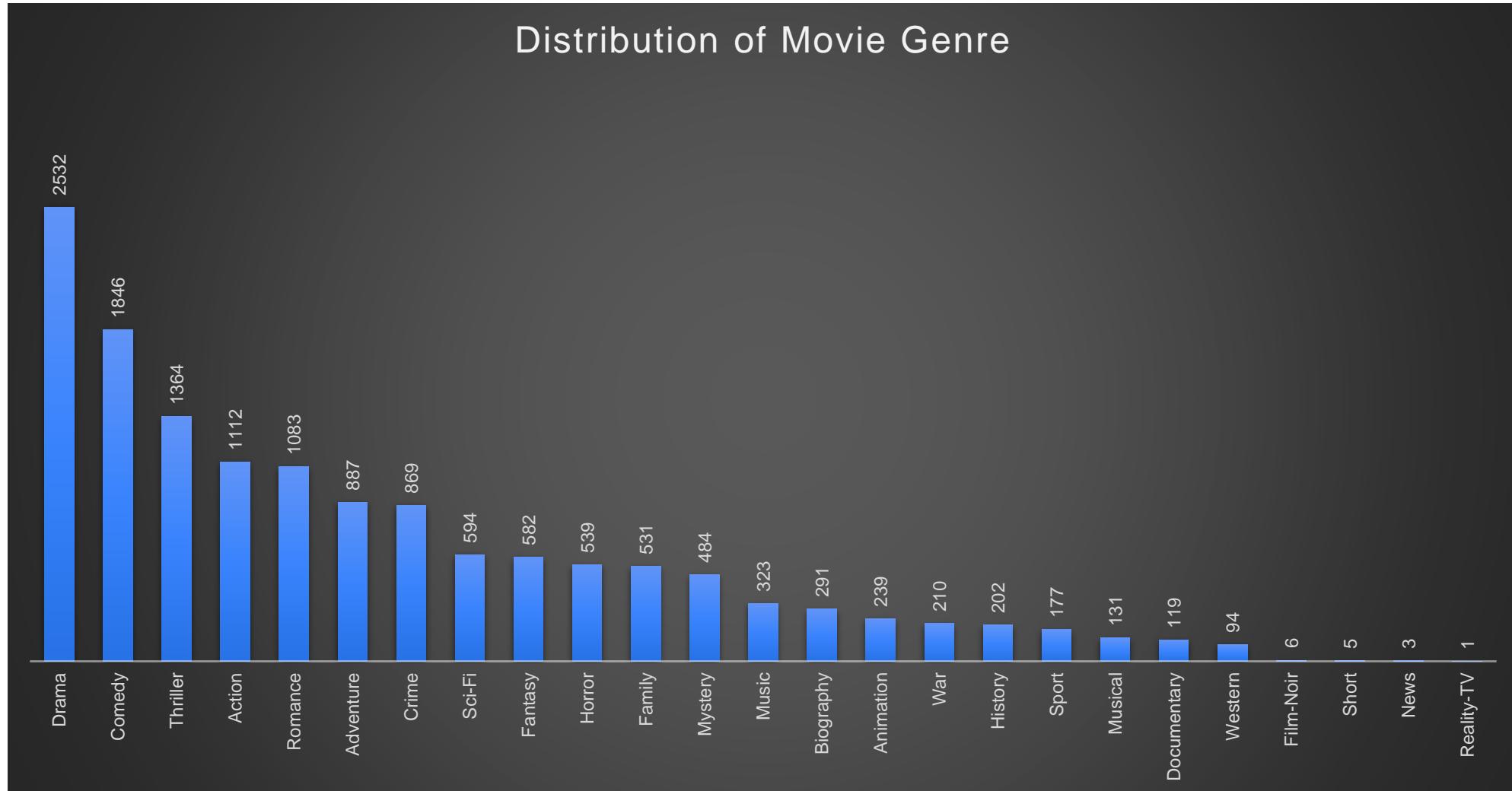
Identified and removed duplicate values from the record to ensure each observation is unique.

- **Handling Missing Values:**

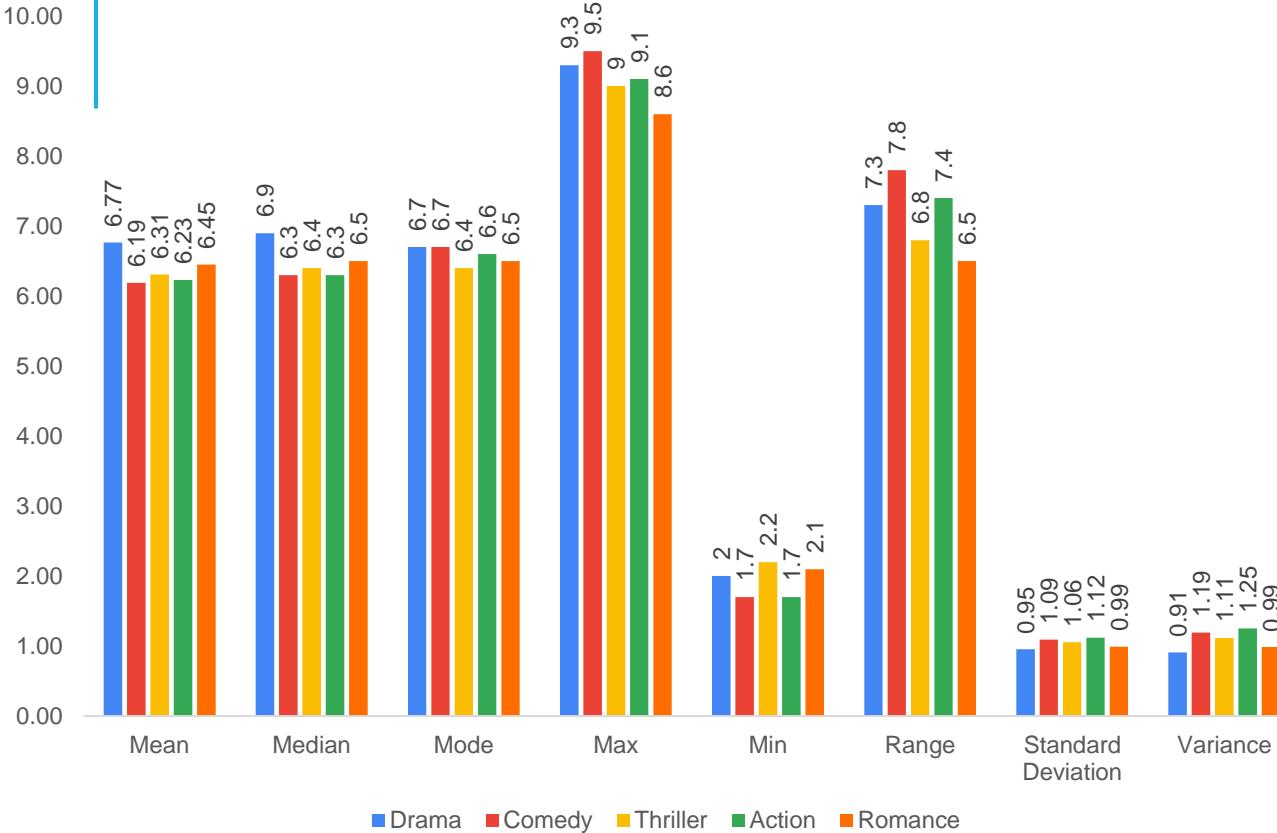
1. Eliminated records having missing values greater than 9 by using count blank(range)
2. For null values in columns director name, language, country, duration scrapped it from the movie page.
3. For null values in the gross and budget column implemented a liner regression technique based on the imdb score for gross and budget and then predicted the null values for gross and budget respectively.

Movie Genre Analysis: Analyze the distribution of movie genres and their impact on the IMDB score.

The top 5 genres are Drama, Comedy, Thriller, Action and Romance



Descriptive Statistics of Top 5 Genre



➤ Insights

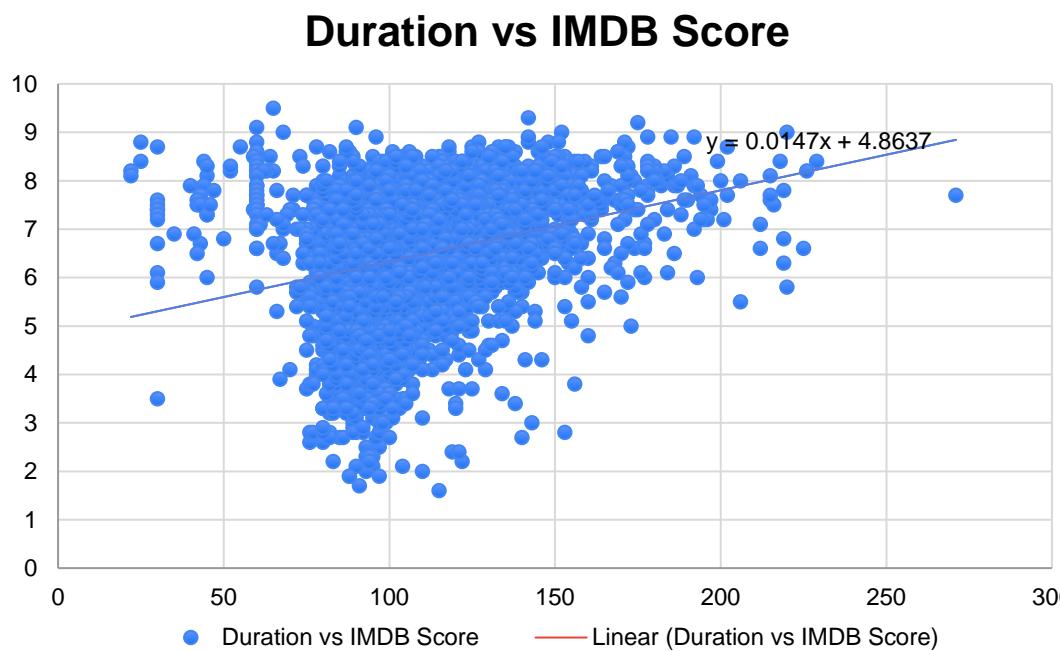
Drama tends to be the most consistently high-rated genre with lower variability in ratings.

Comedy and **Action** are generally rated lower on average and show more variability in their ratings.

Adventure movies have the most variability in ratings, indicating diverse audience reactions.

Movie Duration Analysis: Analyze the distribution of movie durations and its impact on the IMDB score.

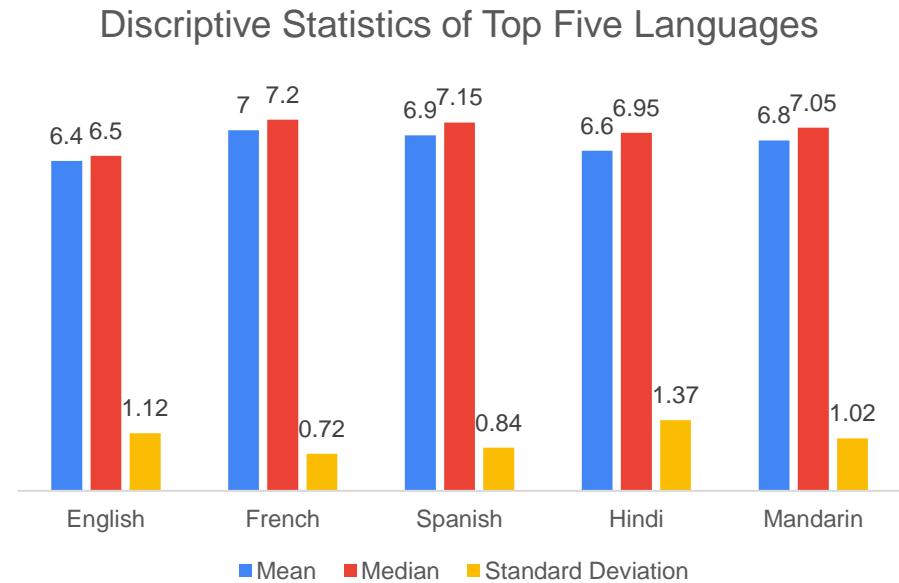
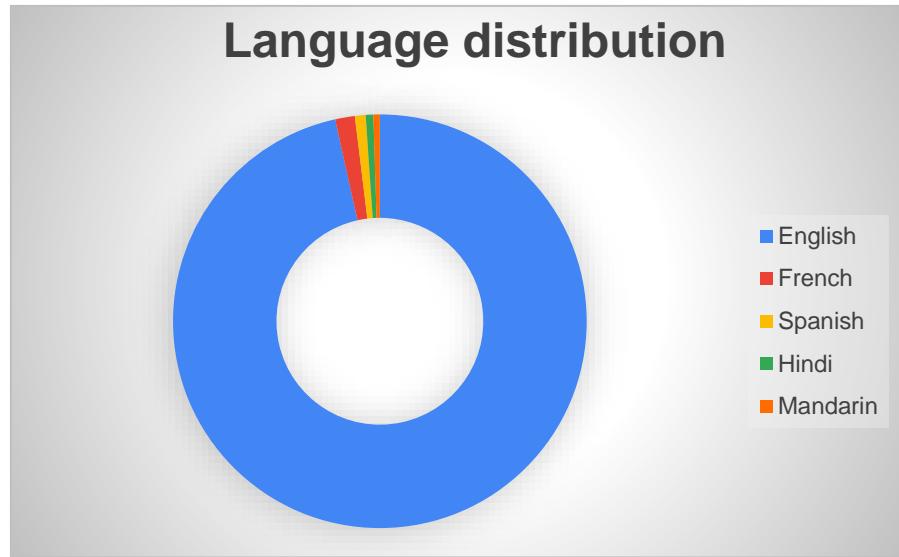
Descriptive Statistics	Values
Mean	107.13
Median	103.00
Mode	90.00
Standard Deviation	22.16



➤ Insights

- ❖ The average movie duration is around 107 minutes.
- ❖ The most common movie duration is 90 minutes.
- ❖ There's a moderate spread of movie durations around the mean (22.16 minutes)
- ❖ A positive correlation between longer movies and higher ratings.

Language Analysis: Examine the distribution of movies based on their language.

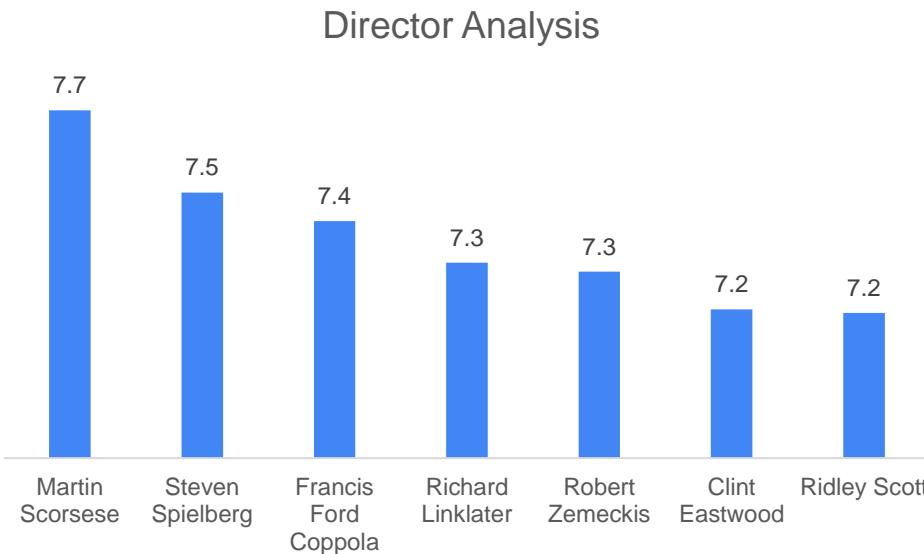


➤ Insights

- **High-Rated Languages:** French, and Mandarin movies generally receive higher and more consistent IMDB scores compared to English, Spanish, and Hindi movies.
- **English Movies:** While English movies dominate in quantity, they exhibit a wide range of ratings, reflecting diverse movie qualities.
- **Consistency:** French and German movies show more consistency in their ratings, as indicated by lower standard deviations.
- **Implications:** For stakeholders looking to produce or invest in high-quality movies, focusing on languages with historically higher and more consistent ratings might be beneficial. However, the larger volume of English movies provides more opportunities for hits despite the broader range of quality.

Director Analysis: Influence of directors on movie ratings.

Row Labels	Count of movie_title	Average of imdb_score	PERCENTILE
Martin Scorsese	20	7.7	7.2
Steven Spielberg	26	7.5	7.2
Francis Ford Coppola	11	7.4	7.2
Richard Linklater	11	7.3	7.2
Robert Zemeckis	13	7.3	7.2
Clint Eastwood	20	7.2	7.2
Ridley Scott	17	7.2	7.2



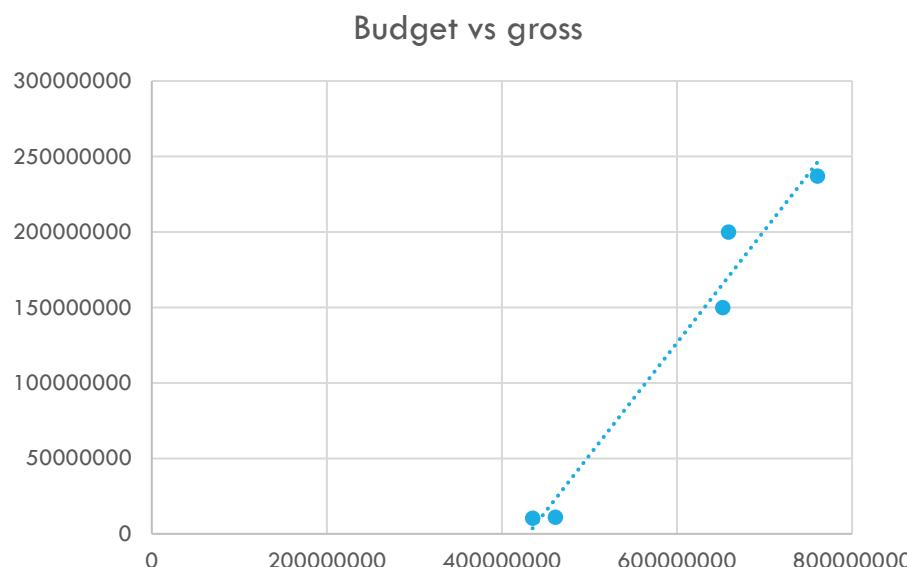
➤ Insight

- Directors with average scores above 7.0 are generally well-regarded and produce movies that are well-received by both critics and audiences. This indicates that their involvement in a movie project can be a key factor in its success.

Budget Analysis: Explore the relationship between movie budgets and their financial success.

correlation between budget and Gross 0.984311735

Top 5 movie_title	Profit Margin
Avatar	523505847
Jurassic World	502177271
Titanic	458672302
Star Wars: Episode IV - A New Hope	449935665
E.T. the Extra-Terrestrial	424449459



➤ Insight

- As the budget increases, the gross earnings also tend to increase.
- The strong positive correlation suggests that investing more in movie production typically results in higher gross earnings.
- This could be due to higher budgets allowing for better production quality, more effective marketing, and attracting larger audiences.

□ CONCLUSION:

- By conducting a thorough analysis of the IMDB Movies dataset, this project aims to provide actionable insights that can drive decision-making and improve the likelihood of producing successful movies.



THANK YOU