IMDB Movie Analysis

Project Description:

This project explores the factors that contribute to a movie's success on IMDB, with success defined by achieving high ratings. By analyzing various aspects of movies such as genre, duration, language, director, and budget, the project aims to uncover patterns and relationships that influence audience ratings. The insights gained from this analysis can help movie producers, directors, and investors understand which elements are most likely to lead to successful films, allowing them to make more informed decisions in future projects. The project combines data cleaning, statistical analysis, and visualization techniques to provide a comprehensive understanding of the factors driving movie success on IMDB.

Approach:

I started by cleaning the IMDB Movies dataset, handling missing values, removing duplicates, and ensuring data consistency. For the analysis, I used Excel to calculate descriptive statistics (mean, median, variance) and performed correlation analysis between key variables like genre, duration, language, director, and budget with IMDB ratings.

I utilized `COUNTIF` to analyze genre distribution and `CORREL` to explore relationships between budget and gross earnings. Scatter plots and trendlines were created to visualize these relationships. The findings were compiled into a report with visualizations, providing actionable insights for stakeholders in the film industry.

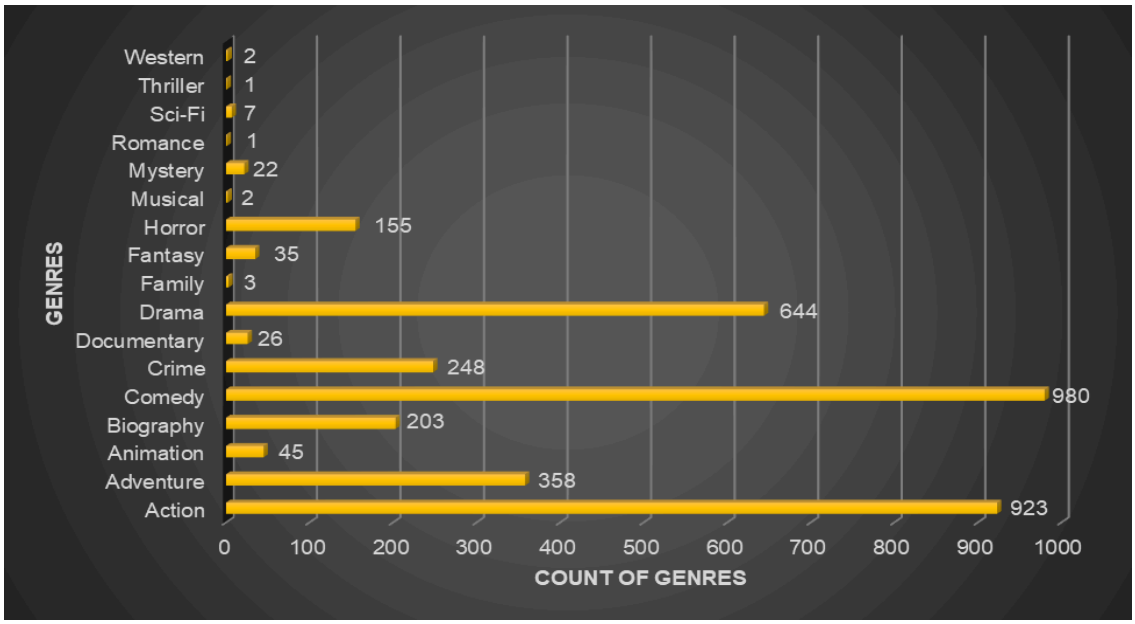
Tech-Stack Used:

Microsoft Excel 2021: Utilized for data cleaning, statistical analysis, and visualization. Excel was used to calculate descriptive statistics, perform correlation analysis, create various charts (e.g., bar charts, scatter plots), and compile the final report with visual aids. Excel’s powerful functions and tools made it the ideal software for analyzing and presenting the IMDB Movies dataset effectively.

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

- **Task:** Determine the most common genres of movies in the dataset. Then, for each genre, calculate descriptive statistics (mean, median, mode, range, variance, standard deviation) of the IMDB scores.

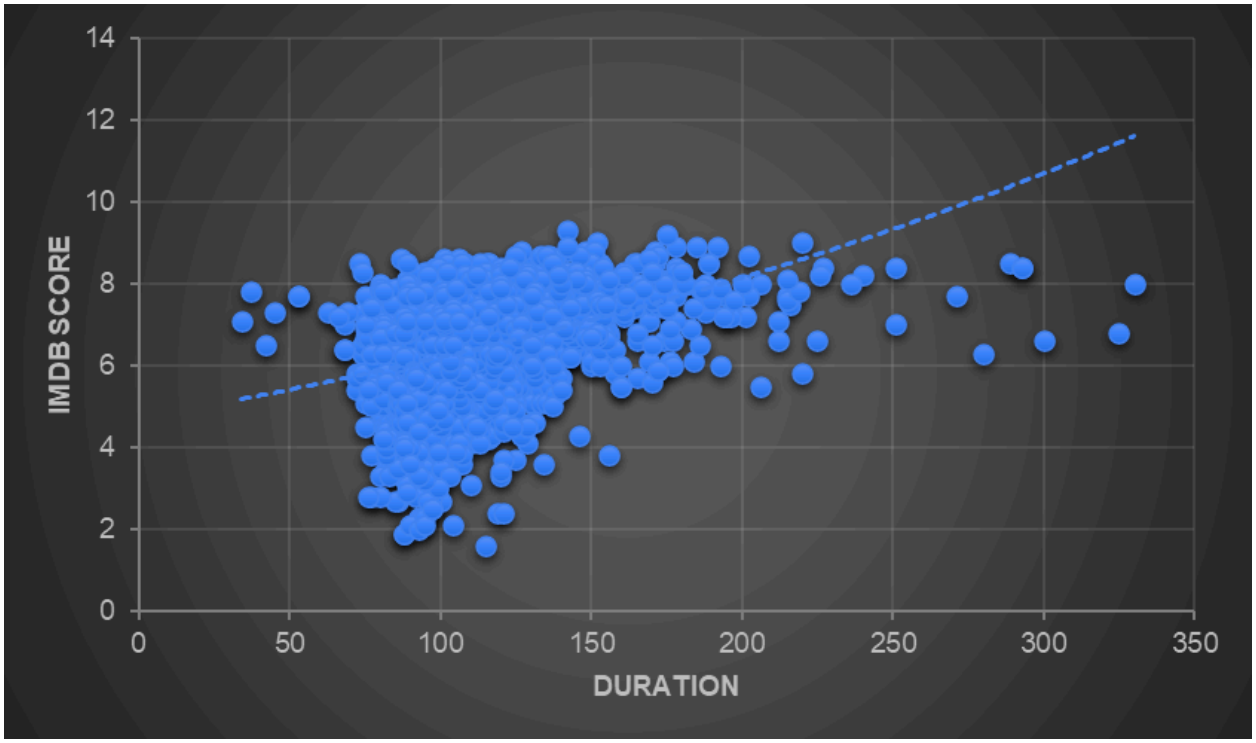
According to the data, ‘Comedy ’genre is most common among other movies in the dataset. Followed by ‘Action’ then ‘Drama’.



Count	17
Average	215
Median	35
Mode	2
Max	980
Min	1
Variance	106288.75
SD	326.01955

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

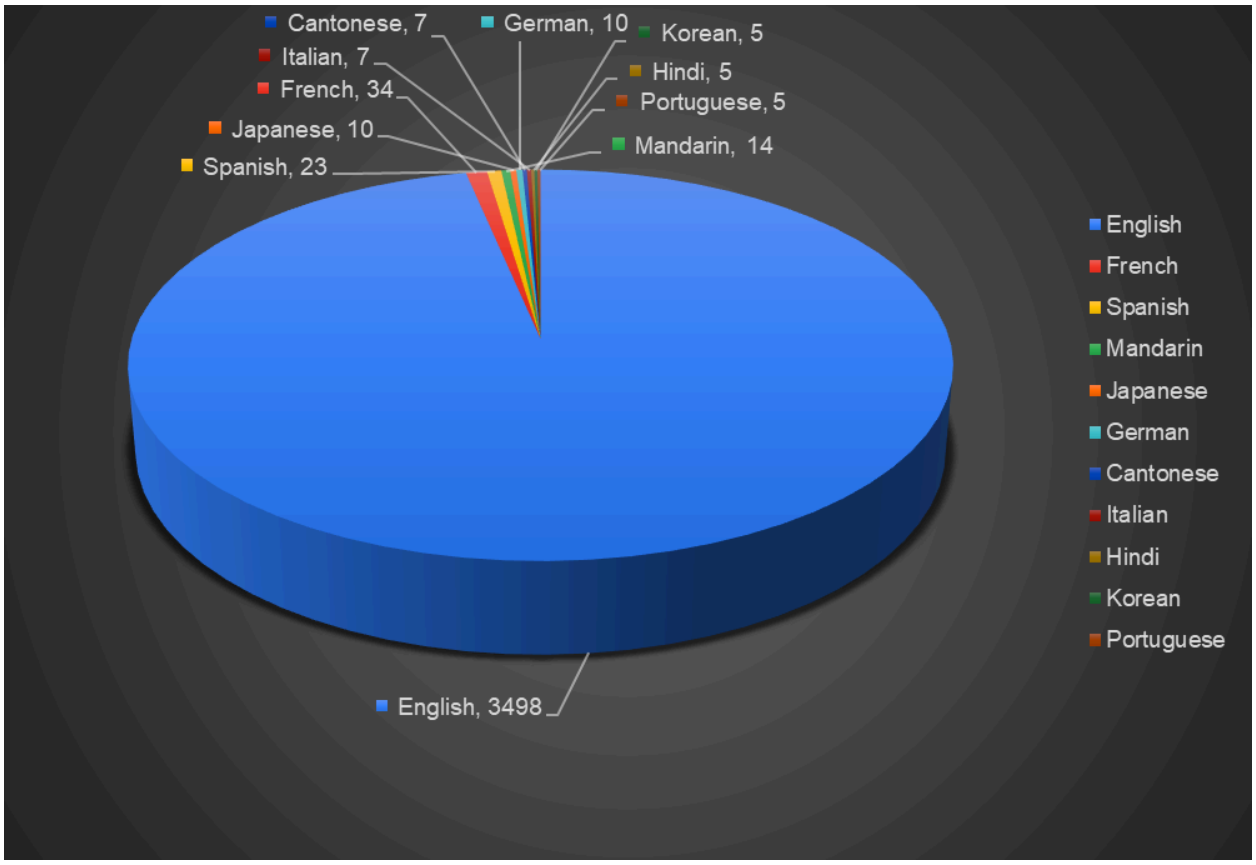
- Task: Analyze the distribution of movie durations and identify the relationship between movie duration and IMDB score.



Average	110.2318
Median	106
SD	23.00157

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

- **Task:** Determine the most common languages used in movies and analyze their impact on the IMDB score using descriptive statistics.



English is the most common language across all movies with 3498 movies.

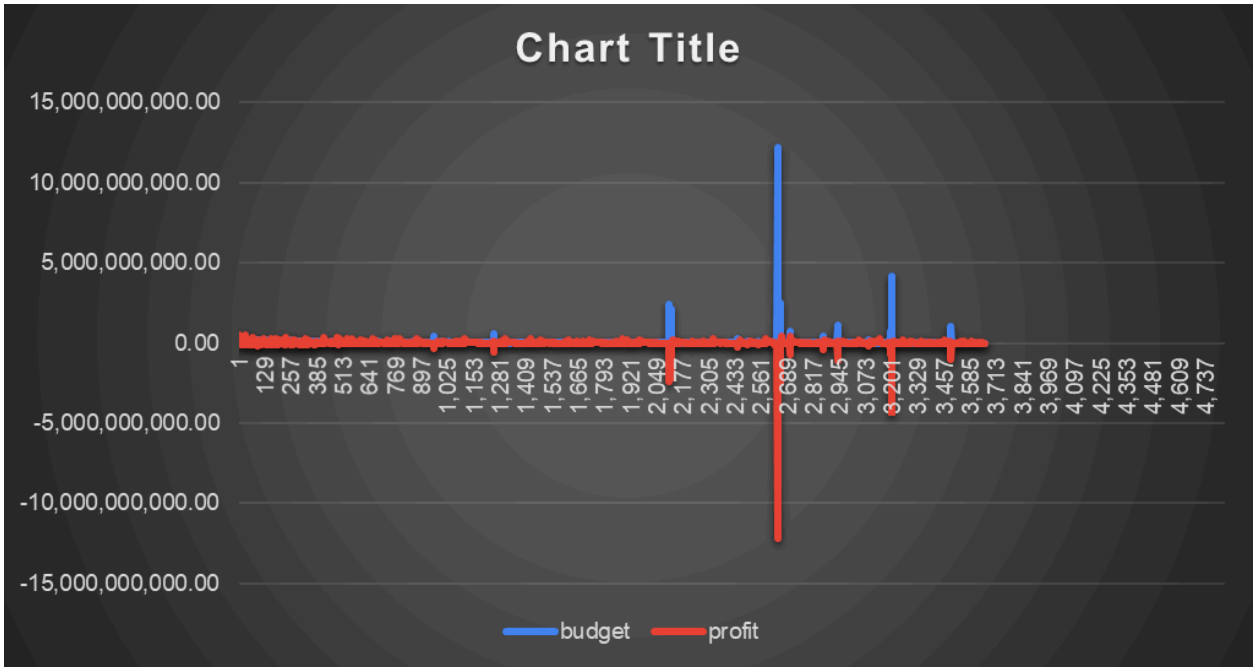
D. Director Analysis: Influence of directors on movie ratings.

- Task: Identify the top directors based on their average IMDB score and analyze their contribution to the success of movies using percentile calculations.

Directors like Akira Kurosawa, Tony Kayne, Charles Chaplin, etc have more scores on imdb.

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

- Task: Analyze the correlation between movie budgets and gross earnings, and identify the movies with the highest profit margin.



The movie with the highest profit margin is ‘Avatar’.

Result:
Through this project, I gained a deep understanding of what influences IMDB movie ratings. By analyzing genre, duration, language, director, and budget, I identified key factors that correlate with higher ratings and financial success. The insights revealed which genres and directors are linked to better ratings, how movie length impacts success, and the role of language and budget in financial performance. This analysis provides actionable guidance for movie industry professionals and enhances my skills in data analysis, allowing for more informed decision-making and strategic planning.

Excel File Link: project_5_excel.xlsx