IMDB Movie Analysis Using Power BI

PROJECT DESCRIPTION:-

The IMDb Movie Analysis project aims to explore and analyze a comprehensive dataset of movies available on the IMDB platform. This dataset contains essential information about movies, including director names, movie titles, duration, genre, budget, gross earnings, IMDb ratings, and more. Through in-depth data analysis using Power BI this project seeks to extract valuable insights and trends that contribute to a movie's success.

In this project, I was required to provide a detailed report for the below data record mentioning the answers of the questions that follows:

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.

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.

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.

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.

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.

MY APPROACH :-

I have gone through the dataset and understood all the given columns.. This dataset consists of unwanted columns, Null values and Blank rows. So, I have decided to Clean this dataset thoroughly.
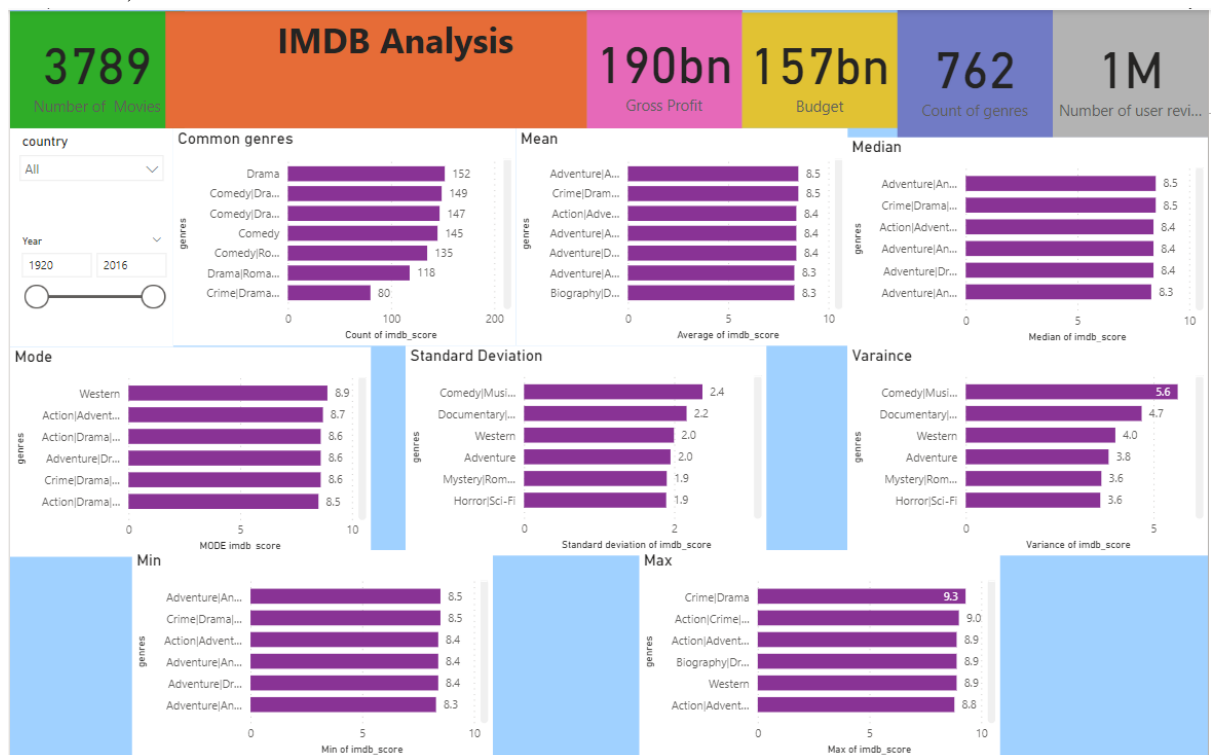
TECH STACK :-

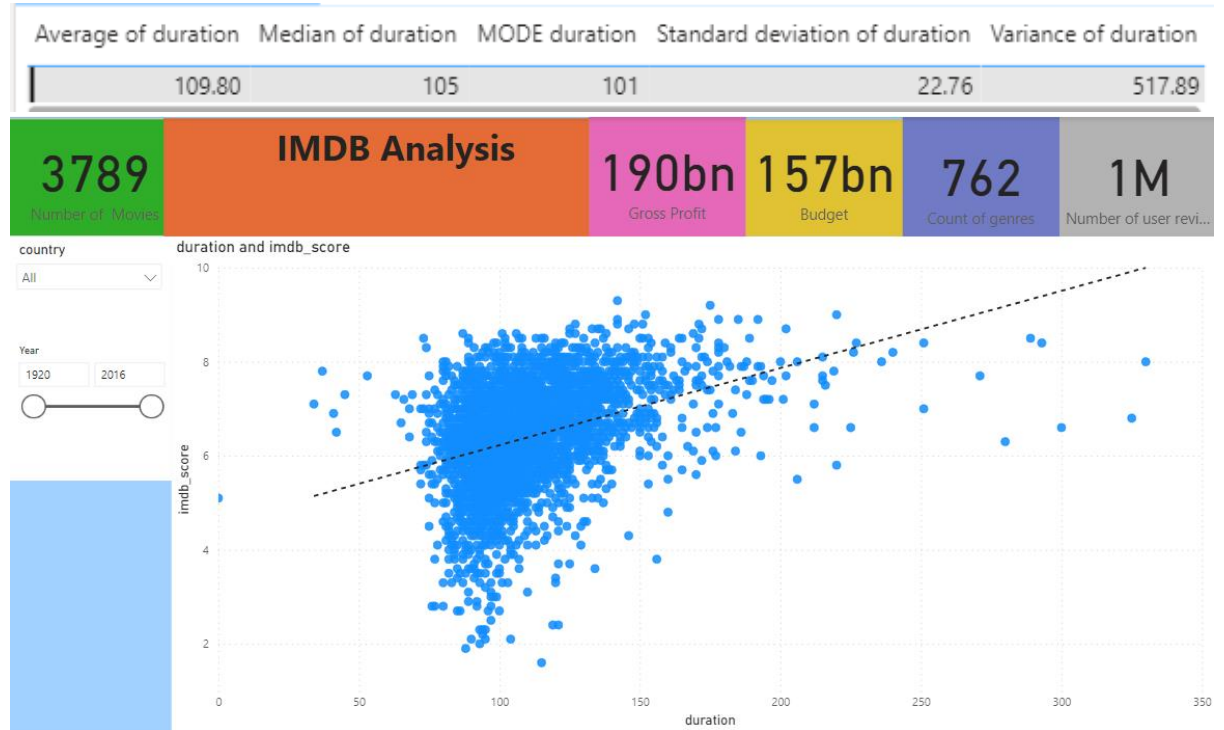For this project, I have used Microsoft Power BI to gain insights.

INSIGHTS :-

1) **Movie Genre Analysis:**

   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.
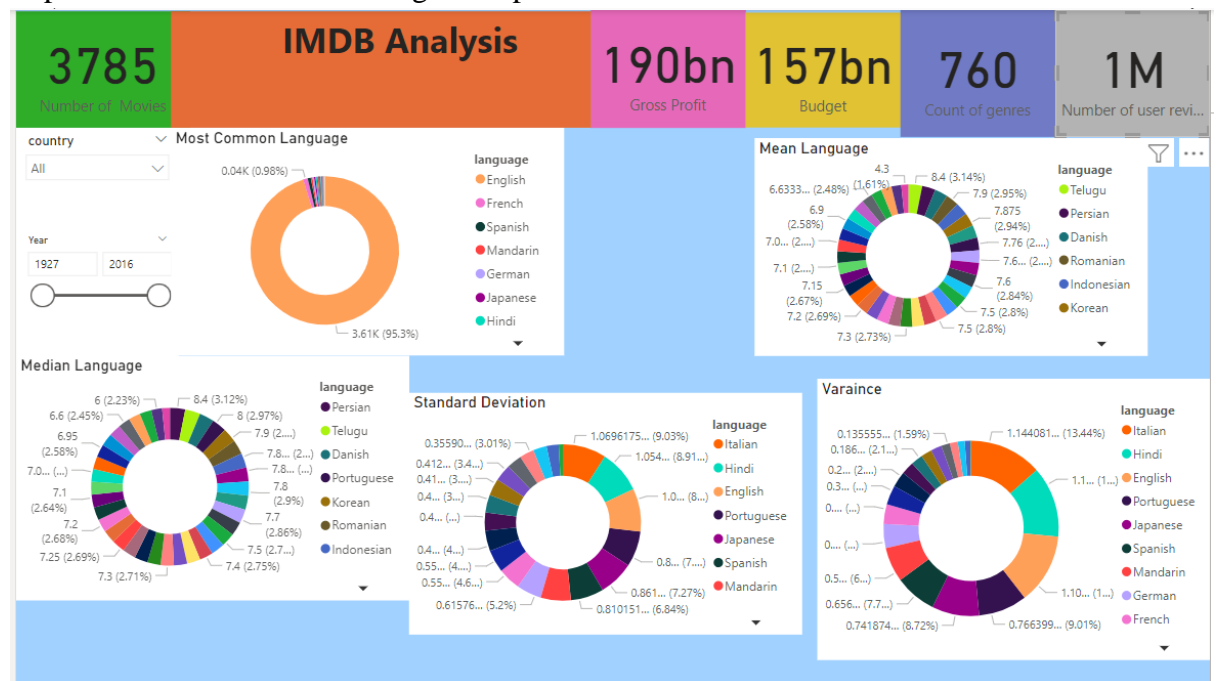
2) **Movie Duration Analysis:**

Task: Analyze the distribution of movie durations and identify the relationship between movie duration and IMDB score.
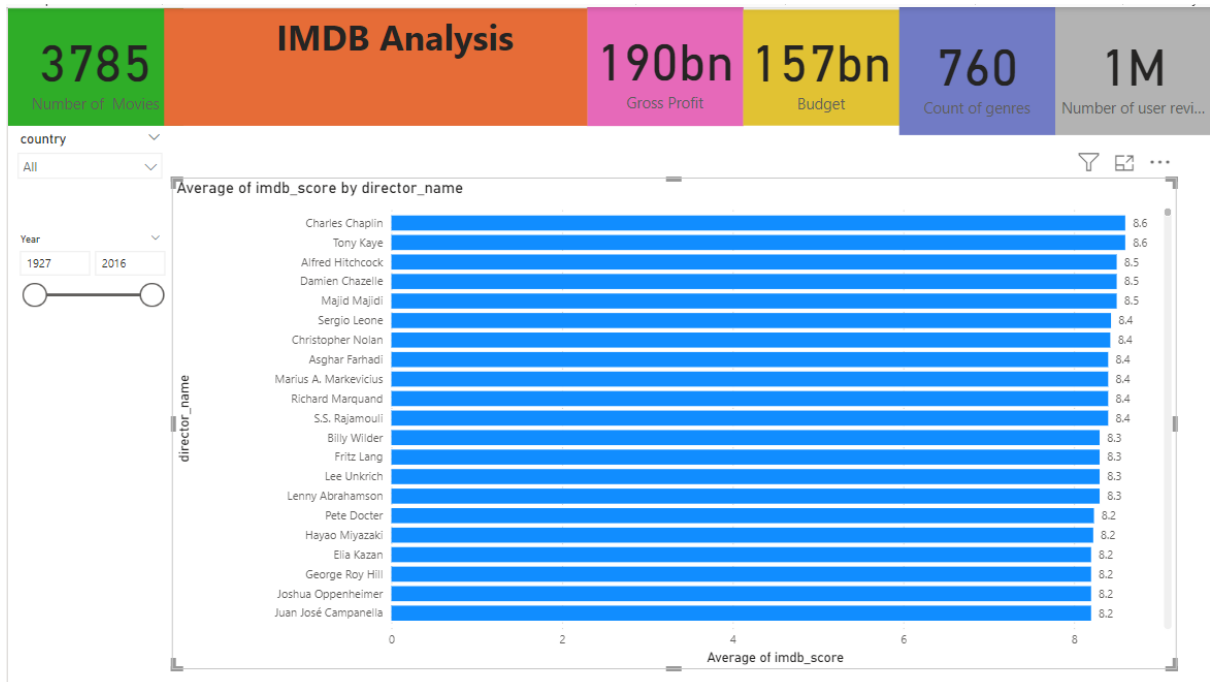
Average of duration	Median of duration	MODE duration	Standard deviation of duration	Variance of duration
109.80	105	101	22.76	517.89

3) **Movie Language Analysis:**

Task: Determine the most common languages used in movies and analyze their impact on the IMDB score using descriptive statistics.
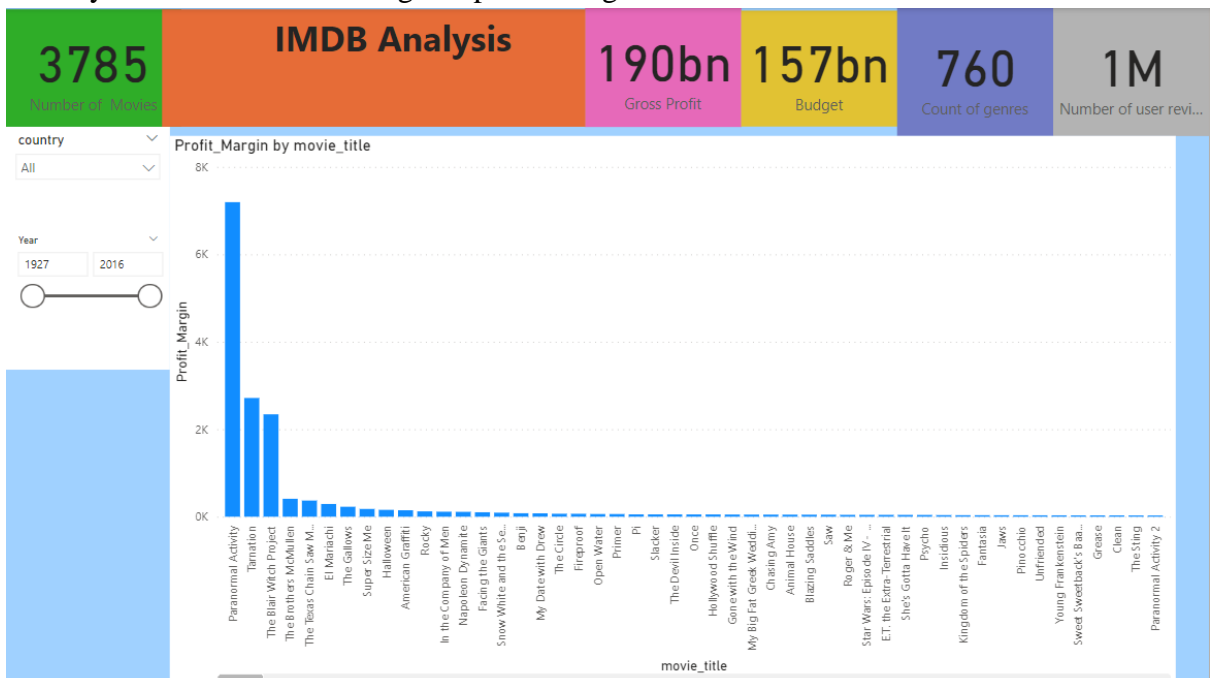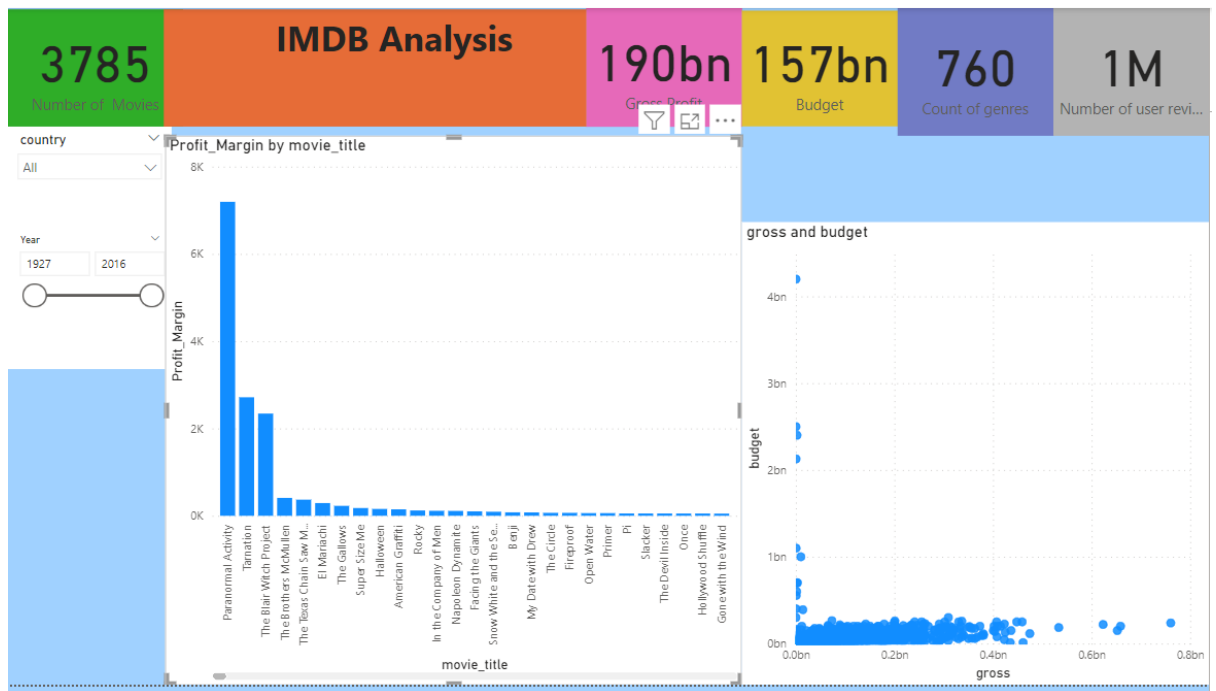
4) Movie Director Analysis:

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

5) Movie Budget Analysis:

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

I have noticed that,

1) The Most common movie genres from the dataset are Drama, Comedy, Thriller and Action.
2) The Average duration of a Movie is 109 minutes. The trendline between the duration vs imdb score is elevated upward with $R^2 = 0.131$
3) The Most common languages used in the movies are English, French, Spanish, Mandarin and German. I have also Observed that the languages Telugu and Persian have the highest average imdb score.
4) I have identified that Tony Kaye, Charles Chaplin, Alfred Hitchcock, Ron Fricke, Damien Chazelle, Majid Majidi, Sergio Leone, Christopher Nolan, SS Rajamouli and Richard Marquand are the top 10 directors with average imdb score >=8.4
5) The Top-5 with highest profit margins are Paranormal Activity, Tarnation, The Blair Witch Project, The Brothers McMullen, The Texas Chainsawman.

RESULTS :-

With the help of this project, I have gained valuable experience for data analysis using statistical knowledge and Power BI's data visualization. Through this, I have learnt to apply my data analysis skills in solving real life problems.