

IMDB MOVIE ANALYSIS REPORT

July 2024

CONTENT

01

Overview

02

Objectives

03

Approach

04

Additional Changes

05

Insights and Knowledge Gained

06

Conclusion

OBJECTIVES

Objective

The objective of the project is to analyze IMDb movie data to uncover insights related to various aspects such as genres, directors, budgets, gross earnings, and their impact on IMDb scores. The project aims to understand patterns, correlations, and trends within the dataset to derive meaningful conclusions about factors influencing movie success.

APPROACH

Q1 . Movie Genre Analysis: Analyze the distribution of movie genres and their impact on the IMDB score.

I have determined the most common genres of movies and calculated Descriptive statistics Such as (Mean,Median,Standard Deviation, Variance) based on IMDB Score.

 [Click Here To View Excel Sheet](#)

I have achieved it using excel formulas. I have also calculated Descriptive statics for each specific genera

According to my analysis the most common genera is **Drama** in dataset as most of movies specified in this genres

I have achieved this by Calculating total number of movies for Each genres and calculated Statistics.

APPROACH

Q2. Movie Duration Analysis: Analyze the distribution of movie durations and its impact on the IMDB score.

I have Calculated Average, Median and Standard Deviation of Duration as given in the task and additionally i have added a new column that contain duration in hours (basically i have converted minutes into hours). And based on duration and IMDB score I have plotted scatter plot

 [Click Here To View Excel Sheet](#)

The plot shows that most of movies have a good imdb score as their duration increases meaning movies having duration more than 100 minutes have a good decent rating and the rating increasing as movie duration increases. Movies within a moderate range tended to have more consistent rating.

APPROACH

Q3 . 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.

 [Click Here To View Excel Sheet](#)

I have created a table Consisting of columns Language, Count, Average, Median, Std. i have calculated the average of language based on imdb score. The most Common languages are **English, Sanish, Mandarin, German, Japanese** maximum number of movies contains this languages.

According to my analysis after analyzing movies descriptive statics with respect to average of imdb score the common languages movies have a good imdb rating greater than 6.

APPROACH

Q4. 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.



[Click Here To View Excel Sheet](#)

I have created a pivot table consisting of director name and imdb score and identified top 10 directors based on imdb score and calculated contribution to success of movies using percentile

I have used large function to find highest imdb score and then calculated percent rank for that imdb score and then calculated percentile. So it tells the directors contribution to the success of movies is 100%

APPROACH

Q5. 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.

 [Click Here To View Excel Sheet](#)

I have Calculated the Correlation between the movie budget and gross earnings and also created Profit column that contain profit of movies (gross-budget).

I Have also Shown the Highest Profit margin of movie which is 523505847

ADDITIONAL CHANGES

1

Added Profit Column

This Shows the profit of Movies .It is created by using gross and budget column and subtracting them.

2

Added Profit and Loss Column

This Column tell where there is profit of loss it is simply created by using conditional column where if $\text{profit} > 0$ then Profit. else Loss

3

Added Duration Column

Created a duration column that consist of movies duration in hours.

INSIGHTS AND KNOWLEDGE GAINED

1 Genre Impact on IMDb Scores:

- Finding: Certain genres consistently received higher IMDb scores than others.
- Insight: History, War ,Drama tended to have higher average IMDb scores compared to horror and Adventure genres, indicating viewer preferences for specific types of content.

2 Director Influence on Ratings:

- Finding: Top directors, identified by their average IMDb scores, significantly influenced movie ratings.
- Insight: Directors known for their consistent high-quality productions often correlated with higher IMDb ratings, underscoring the impact of directorial talent on movie success.

3 Language and IMDb Ratings:

- Finding: Movies in specific languages showed varying IMDb score distributions.
- Insight: English-language movies generally exhibited higher average IMDb scores, reflecting broader international appeal and accessibility.

INSIGHTS AND KNOWLEDGE GAINED

4

Financial Success and Budgets:

- Finding: There exists a positive correlation between movie budgets and gross earnings.
- Insight: Movies with higher budgets tended to achieve higher gross earnings, although outliers with lower budgets also achieved notable success, suggesting diverse strategies for financial viability in filmmaking.

5

Duration and Viewer Perception:

- Finding: Movie duration impacts IMDb ratings to a certain extent.
- Insight: While very short or very long movies may polarize ratings, movies within a moderate duration range tended to receive more consistent IMDb scores, indicating viewer preference for balanced storytelling lengths.

DESCRIBE WHAT YOU ACHIEVED THROUGH THE PROJECT

Data Exploration:

- Successfully imported, cleaned, and manipulated IMDb movie data in Excel.
- Split genres and languages to analyze their impact on IMDb scores.
- Used Excel functions and pivot tables for effective data summarization and analysis.

Statistical Analysis:

- Calculated descriptive statistics (mean, median, mode, range, variance, standard deviation) to understand distribution and trends across genres, directors, budgets, and more.
- Utilized correlation analysis to explore relationships between movie budgets, gross earnings, and IMDb scores.

Insights into Movie Genres and Directors:

- Identified genres and directors associated with higher IMDb scores.
- Analyzed how these factors influence viewer perceptions and ratings.

Financial Analysis:

- Explored the financial aspects of movies by correlating budgets with gross earnings.
- Identified movies with the highest profit margins, providing insights into financial success strategies in filmmaking.

CONCLUSION

The project not only provided quantitative insights into factors influencing movie ratings but also highlighted qualitative aspects such as directorial influence and genre preferences. These findings are valuable for stakeholders in the movie industry seeking to optimize production strategies, enhance viewer engagement, and maximize financial returns based on IMDb ratings.