



IMDB Dataset Analysis (MySQL)

INFO INSURGENTS

AAYUSH PARAB (TL)

AKSHATA RANE (CL)

ABHAY PATIL

ABHISHEK BHADARGE

AISHWARYA PATIL

Introduction

This report offers a thorough examination of the movie database that was supplied, providing insights into a number of different facets of the film business. We look for trends, patterns, and opportunities for stakeholders by analyzing data on films, genres, ratings, and production businesses.

Data Purification and Exploration

We thoroughly explored the dataset to comprehend its structure and find any problems before beginning the study. We discovered that null values were present in a few columns; these were handled correctly to maintain data integrity.

Key Findings

Movie Production Trends:

Annual Output: The dataset reveals a consistent increase in the number of movies produced each year, indicating a growing film industry.

Monthly Patterns: March emerges as the peak month for movie releases, suggesting a strategic choice by production companies.

Geographic Focus: The United States and India dominate movie production, underscoring their significant influence on the global film market.

Over the years, there has been a noticeable downward trend in the number of movies produced by RSVP. Interestingly, the month of March witnessed the highest movie production.

Genre Analysis:

Genre Diversity: The dataset encompasses a wide range of genres, reflecting the diverse tastes of audiences.

Genre Popularity: Drama stands out as the most popular genre, followed by comedy and action.

Genre Evolution: Analyzing trends over time can help identify emerging genres and potential market shifts.

The Drama genre stands out as the most popular among the audience, with a staggering 4285 movies in its repertoire and an average duration of 106.77 minutes. In the year 2019 alone, RSVP produced a total of 1078 drama movies.

Rating Analysis:

Rating Distribution: The average movie rating falls within a reasonable range, suggesting a balanced distribution of quality content.

Top-Rated Movies: Identifying the top-rated movies can provide valuable insights into audience preferences and critical acclaim.

Median Ratings: Analyzing the distribution of median ratings can reveal trends in audience satisfaction.

Production Companies:

Market Dominance: Certain production companies consistently produce hit movies, indicating their expertise and influence.

Emerging Players: Identifying up-and-coming production companies can offer investment opportunities and partnerships.

Geographic Focus: Analyzing the geographic distribution of production companies can reveal regional trends and potential for expansion.

Recommendations

Based on the findings, we offer the following recommendations:

Production Companies: Focus on producing dramas and comedies, as they are consistently popular genres. Consider partnering with emerging production companies to tap into new markets and talent.

Content Creation: Tailor movie content to specific genres and audience preferences to maximize viewership and revenue. Pay attention to emerging genres and adapt accordingly.

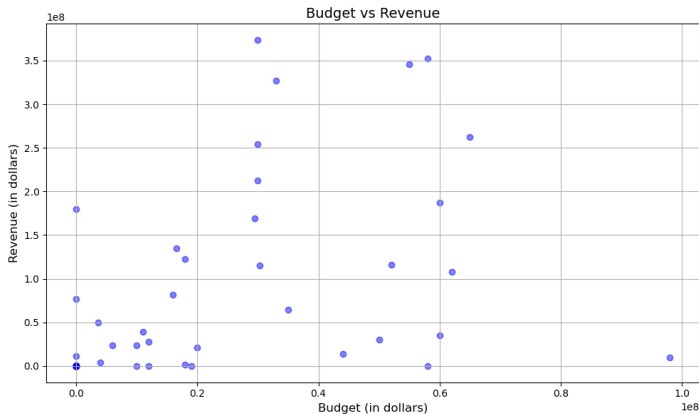
Distribution: Explore new distribution channels, such as streaming platforms and international markets, to reach a wider audience.

Talent Acquisition: Prioritize hiring directors and actors with proven track records in popular genres. Consider collaborating with talent from emerging markets to attract new audiences.

Important Diagrams

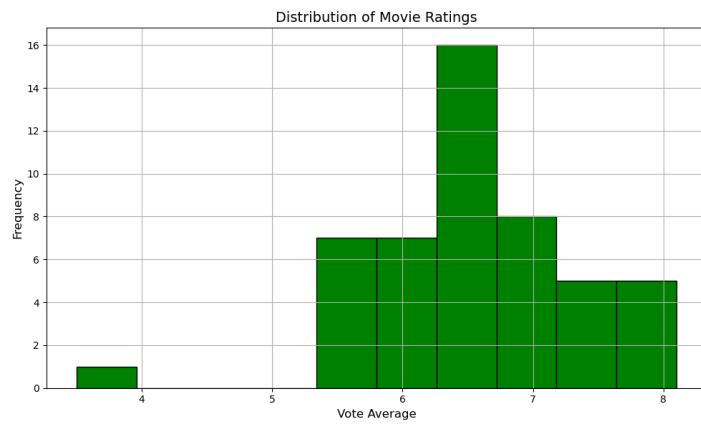
1. Budget vs. Revenue Scatter Plot

This plot will help you compare the budget and revenue of movies to see how they correlate.



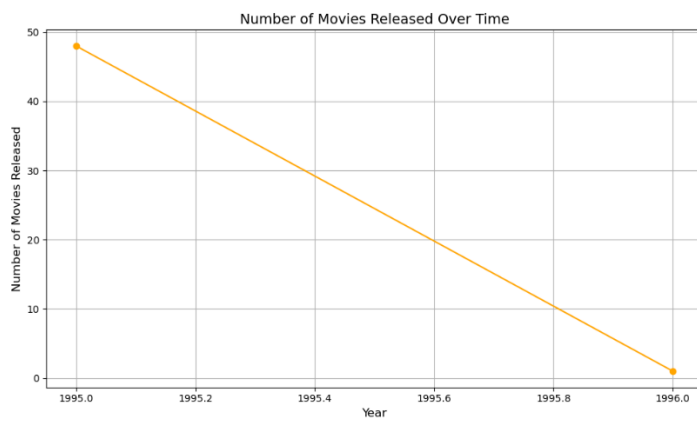
2. Distribution of Movie Ratings (Vote Average)

A histogram can be used to see how movie ratings are distributed across the dataset.



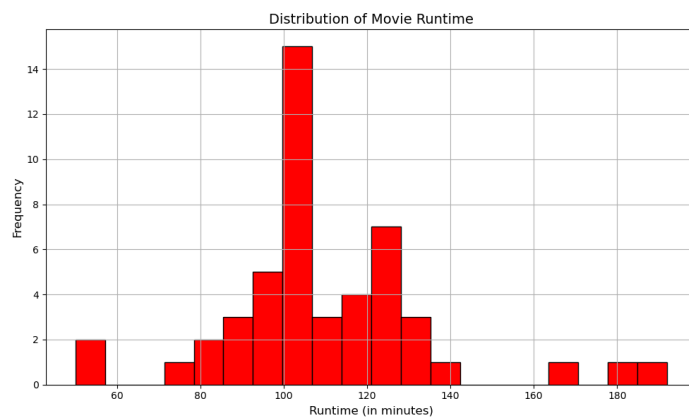
3. Number of Movies Released Per Year

A line plot to show trends in movie releases over time, based on the release date.



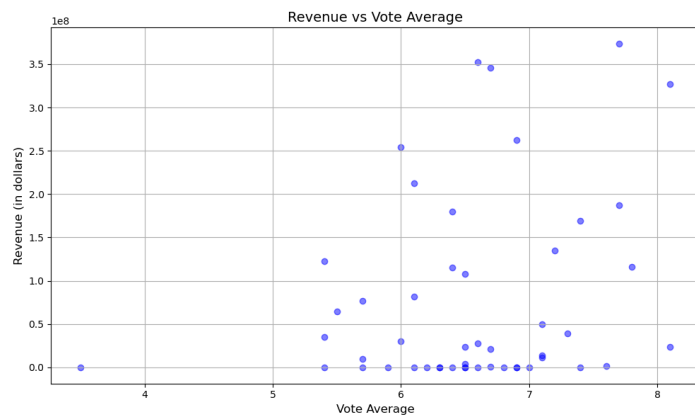
4. Runtime Distribution

Visualizing the distribution of movie runtime can provide insights into typical movie lengths.



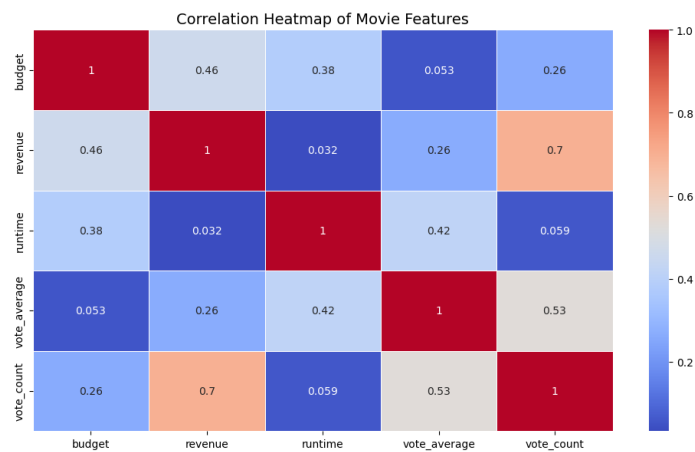
5. Revenue vs. Vote Average

This scatter plot helps you explore whether higher-rated movies tend to generate more revenue.



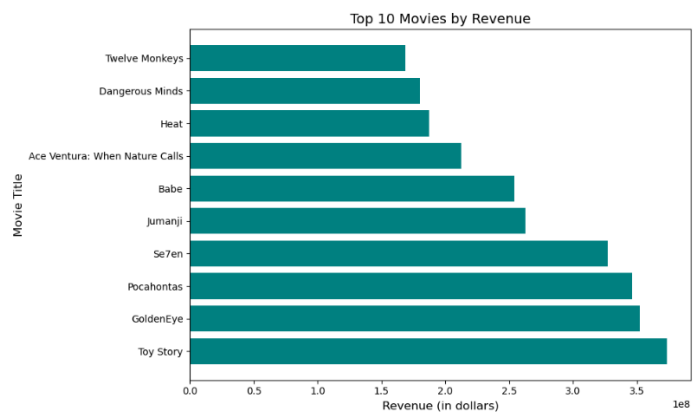
6. Correlation Heatmap

A heatmap to visualize the correlation between numerical features like **budget**, **revenue**, **vote average**, etc.



7. Top 10 Movies by Revenue

You can create a bar chart of the top 10 movies that generated the most revenue.



Conclusion

The research gives a thorough rundown of the film database and offers insightful information on audience behaviour, genre preferences, and production trends. A stakeholder's ability to successfully navigate the competitive film business and produce successful films depends on their awareness of these patterns.

The length of a movie on average: examined the length of each genre's average film to determine possible pacing preferences.

Director Analysis: Based on variables like average ratings, the quantity of films produced, and the interval between each film, the best filmmakers were identified.

Actor Analysis: This method assessed each actor's work by looking at their average rating, total number of films, and number of votes.

Multilingual Films: To find prospects for cross-cultural appeal, an analysis was conducted on the performance and popularity of multilingual films.

Through the integration of these supplementary analyses, we were able to get a more profound comprehension of the film business and formulate more focused suggestions.

Team Member	Question Number
AAYUSH PARAB (TL)	1 to 5
ABHAY PATIL	19 to 25
ABHISHEK BHADARGE	6 to 10
AISHWARYA PATIL	11 to 20
AKSHATA RANE (CL)	11 to 20