

# IMDb Movie Data Analysis

To understand and analyze the IMDb movie dataset using SQL. I explored the data to find meaningful insights related to movies, genres, ratings, actors, directors, and production houses.

The dataset had multiple tables - movies, ratings, names, genres, and mapping tables. It contained details like movie titles, release dates, languages, ratings, total votes, production companies, actors, and directors. I imported all data into MySQL and created relationships between tables to perform analysis.

## Key Insights

- Drama turned out to be the most common movie genre.
- Marvel Studios and Warner Bros. were among the top production houses with high-rated and most voted movies.
- Vijay Sethupathi ranked among the top Indian actors based on average rating.
- Tabu and Taapsee Pannu appeared among the top Hindi actresses with the best ratings.
- Thriller movies were divided into categories like Superhit, Hit, One time watch, and Flop based on average ratings.
- I also identified directors who made the most movies and compared their average ratings, total votes, and movie durations.

## Conclusion

This project helped me explore how SQL can be used for real data analysis and finding actionable insights. It gave a deeper understanding of how movie data can reveal trends like which genres perform best, which production houses dominate, and who the top-rated actors and directors are. Overall, this project strengthened my SQL skills and helped me gain confidence in applying them on real datasets.

## Files Submission

- IMDB\_Project\_Final.sql (all queries with comments)