

# Lights, Camera, Statistics! IMDb Movie Ratings and the Secret Recipe for Success

**Dive deep into the world of IMDb movie ratings to unravel the mysteries behind movie genres, box office triumphs, and other factors that can make or break a film's success.**

In this project, we'll embark on an exciting journey through the IMDb movie ratings universe, where we'll explore the intricate relationships between movie genres, box office performance, and other factors that can turn a modest flick into a blockbuster hit.

Using the power of Python and your newfound knowledge of statistics, you'll become a movie data detective, uncovering the secrets hidden within the ratings, budgets, and revenues. Armed with these insights, you'll be able to offer sage advice to film producers, investors, and distributors, guiding them towards the path of cinematic glory.

Put on your director's hat and get ready to:

- **Demystify the IMDb movie dataset:** You'll clean, preprocess, and make sense of the raw data, transforming it into a treasure trove of insights just waiting to be discovered.
- **Unearth the secrets of movie genres:** Are some genres destined for greatness, while others are doomed to mediocrity? Investigate the relationship between genres and IMDb ratings to reveal the hidden patterns that shape audience preferences.
- **Discover the factors that influence box office success:** Dive into the world of budgets, revenues, and ratings to identify the key ingredients that can turn a film into a box office sensation.
- **Predict the future (of movie ratings):** Harness the power of confidence intervals to make informed predictions about the average ratings of different genres, countries, and time periods.
- **Share your wisdom with the world:** Summarize your findings and offer valuable recommendations to movie industry professionals, helping them make data-driven decisions that lead to success on the silver screen.

Are you ready to become a movie ratings guru and unlock the secrets of cinematic success? Grab your popcorn, dim the lights, and let the adventure begin!

## Project Outline

### 1. Data Preparation and Cleaning

1. Import the dataset and necessary libraries.
2. Clean the dataset by handling incorrect data types and missing values.
3. Extract relevant features for the analysis.

### 2. Exploratory Data Analysis

1. Explore the data
2. Investigate the distribution of **movie ratings**, overall and in relation to **decades**.
3. Analyze the relationships between **ratings**, **critic acclaim**, **Runtime** and **box office success** using correlation and covariance.

### 3. Estimating Average Ratings

1. Apply confidence intervals to estimate the average **box office** for different **decades**.

## 4. Conclusions

1. Summarize the findings from the analysis.

