Predictive Model Summary

1.Objective:

To build a predictive model to estimate a movie's **IMDB rating** (or popularity, revenue, etc.) based on various features such as genre, director, cast, budget, and more.

2.Tools Used:

- **Python (Jupyter Notebook):**
  For data cleaning, exploratory data analysis (EDA), feature engineering, and model building using libraries like pandas, seaborn, scikit-learn, and matplotlib.
- **Excel:**
  Used for initial data inspection, manual corrections, and summary reporting of descriptive statistics.

3.Data Overview:

- Dataset: IMDB-Movie-Data.csv
- Number of movies: ~1,000
- Features: Title, Genre, Director, Cast, Year, Runtime, Rating, Revenue, Votes, etc.

Data Preprocessing:

- Handled missing values (e.g., revenue, rating) and clean the data

Key Insights:

- Revenue and vote count are strong predictors of IMDB rating.
- Movies with top-tier directors and well-known casts tend to score higher and Longer runtimes slightly correlate with higher ratings.
- Genre also plays a significant role, especially Drama and Action.