# IMDb Movies Analysis Report

This report provides a structured analysis of IMDb's Top Rated Movies dataset, covering dataset description, operations performed, key insights, and actionable recommendations.

# 1. Dataset Description

**1.1 Source:** IMDb Top-Rated Movies dataset (user-rated, CSV file).

**1.2 Columns:**

- Title – Movie title
- Director – Director of the movie
- Stars – Main actors/actresses
- Tags – Genres/keywords (comma-separated)
- IMDb Rating – Viewer rating (0–10 scale)
- Meta Score – Critic rating (0–100 scale)
- Votes – Audience vote count

**1.3 Data Quality:**

- No major missing values.
- Votes contained shorthand (K/M) → cleaned and converted to numeric.
- Ratings and scores cast to numeric for proper analysis.

# 2. Operations Performed

## *2.1 Data Cleaning & Preparation*

- Normalized Votes (e.g., 5K → 5000, 2M → 2000000).
- Casted IMDb Rating and Meta Score to double type.
- Split Tags and Stars into individual values for frequency analysis.

## *2.2 Descriptive Analytics*

- Summary statistics for ratings and scores.
- Top/bottom movies by IMDb and Meta Score.
- Distinct directors, stars, and tag counts.

## *2.3 Relationship & Comparative Analysis*

- Correlation between IMDb Rating and Meta Score.
- Score gap analysis (IMDb vs Meta Score).
- Directors with consistently high/low ratings.
- Popularity by tags and stars.

## *2.4 Visualization*

- Histograms (IMDb Ratings).
- Scatter plots (IMDb vs Meta Score).
- Word clouds (tags, title words).
- Bar charts (top movies, directors, stars, tags).
- Donut and polar charts (rating distributions, categories).

## 3. Key Insights

### *3.1 Ratings & Scores*

• IMDb Ratings mostly fall between 7.5–9.0.
• High correlation with Meta Score, but notable gaps exist for some movies.
• Consistently high-performing directors have minimum ratings ≥ 8.

### *3.2 Director & Star Insights*

• Certain directors (e.g., Christopher Nolan, Steven Spielberg) dominate top ratings.
• Stars with frequent appearances often align with higher-rated films.
• Directors with 3+ movies show a wide spread in consistency (std dev analysis).

### *3.3 Tag/Genre Trends*

• Drama and Action tags are most common.
• Tags like Biography and History often carry higher IMDb averages.
• Combined tags (Action + Drama) appear in many top movies.

### *3.4 Audience & Critics Alignment*

• Correlation indicates alignment, but large score gaps highlight cases where audiences and critics disagree.
• Some highly rated IMDb movies received lower Meta Scores and vice versa.

## 4. Recommendations

### *4.1 For Filmmakers*

• Focus on consistently high-performing genres (Drama, Action, Biography).
• Collaborating with high-reputation directors/stars increases likelihood of top ratings.

### *4.2 For Platforms (e.g., streaming services)*

• Highlight directors with strong consistency (all films rated ≥ 8).
• Promote movies where both audience and critics align highly.

### *4.3 For Analysts*

• Explore predictive models to estimate rating success of upcoming releases.
• Build clustering models using Tags, Ratings, and Votes to profile audience preferences.
• Study critic–audience divergence for deeper insights into perception gaps.