IMDB

MOVIES ANALYSIS REPORT

Introduction

The popularity of movies is an ever-evolving phenomenon influenced by various factors such as genre, director, actors, and audience preferences. This report presents an in-depth analysis of the IMDB dataset comprising 1,000 of the most popular movies released between 2006 and 2016. Through this analysis, we aim to uncover patterns, trends, and insights into what makes a movie successful in terms of audience ratings and revenue. The findings will be supported by data visualizations and statistical summaries, offering a comprehensive overview of the movie industry trends during this period.

Project Overview and Objectives

The primary objective of this project is to analyse the characteristics of popular movies over a decade to identify patterns and insights that influence their success. By exploring attributes such as runtime, genre, director, ratings, votes, and revenue, the analysis aims to answer specific questions, such as:

1. Which years saw the highest audience engagement?
2. What are the characteristics of top-rated and high-revenue movies?
3. How does a movie's rating affect its revenue?
4. What trends exist in terms of movie runtime, genre, and yearly production?

These insights are valuable for industry professionals, researchers, and enthusiasts interested in understanding the dynamics of movie popularity and success.

Dataset Information

The dataset used for this analysis comprises 1,000 movies listed on IMDB from 2006 to 2016. The dataset includes the following attributes:

- **Title**: Name of the movie.
- **Genre**: Movie genres, often including multiple categories.
- **Description**: A brief synopsis of the movie.
- **Director**: The movie's director.
- **Actors**: Lead actors in the movie.
- **Year**: Release year of the movie.
- **Runtime (Minutes)**: Total runtime in minutes.
- **Rating**: IMDB rating on a scale of 1 to 10.
- **Votes**: Total number of audience votes.
- **Revenue (Millions)**: Box office revenue in millions.
- **Metascore**: Metacritic score, an aggregate rating from critics.

Data Preprocessing and Descriptive Statistics

Preprocessing:

Before delving into the analysis, the dataset underwent several preprocessing steps to ensure accuracy and consistency:

1. Handling Missing Values: Rows with missing values in critical columns such as Revenue and Metascore were removed to maintain the integrity of the analysis.
2. Removing Duplicates: The dataset was checked for duplicate entries, and none were found.
3. Data Formatting: Data types for numerical columns were verified, and formatting inconsistencies were corrected.
4. Feature Engineering: A new column, rating_category, was created to classify movies based on their ratings into categories such as "Excellent," "Good," and "Average."

Descriptive Statistics

- Total Movies: 838 after preprocessing.
- Time Span: 2006 to 2016.
- Average Rating: 6.81.
- Average Runtime: 114 minutes.
- Revenue Range: From $0.18 million to $936.63 million.

Observations from Statistics

- The dataset showcases a rich variety of genres, with a significant number of movies having runtimes between 100 and 125 minutes.
- The majority of movies have ratings between 6.0 and 7.5, indicating a relatively high average audience approval.

---

1. Titles of Movies with Runtime >= 180 Minutes

| Title	Runtime (Minutes)
The Wolf of Wall Street	180
The Hateful Eight	187
La vie d'Adèle	180

Movies like *The Wolf of Wall Street*, *The Hateful Eight*, and *La vie d'Adèle* are among the longest, each with runtimes of 180 minutes or more. Longer runtimes are often associated with detailed storytelling and epic narratives, appealing to niche audiences who prefer immersive cinematic experiences.

---

2. Find which year there was the highest voting

The year 2014 garnered the highest total votes (19.7 million), followed by 2013 and 2012. A bar chart depicting yearly votes reveals a noticeable spike in 2014, indicating a peak in audience engagement.

This period reflects a surge in audience interaction, possibly due to the release of highly anticipated blockbuster movies and advancements in digital platforms encouraging online voting and reviews.

	Year	Votes
0	2014	19709200
1	2013	18944679
2	2012	18033412
3	2010	14881727
4	2011	14777520
5	2015	14116879
6	2016	13550689
7	2008	12795847
8	2009	12023126
9	2007	11727351
10	2006	11366521

Votes By Year
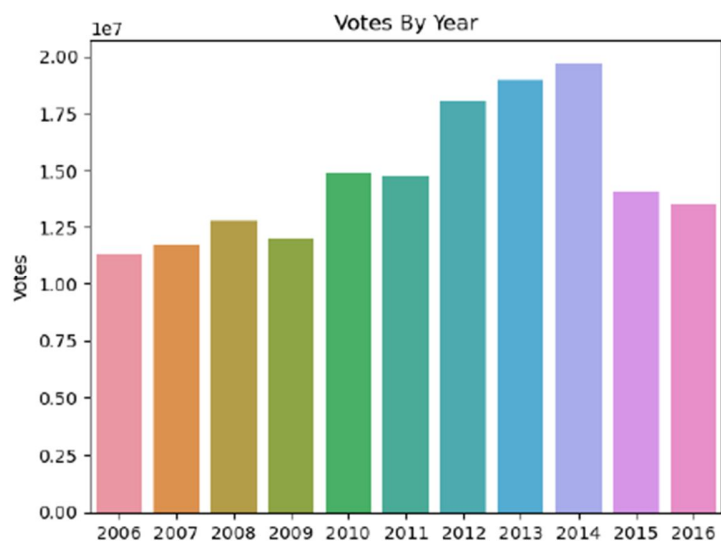
3. Year with the Highest Revenue

	Year	Revenue (Millions)
0	2016	11002.09
1	2015	8799.09
2	2014	7945.33
3	2013	7399.11
4	2012	6826.39
5	2010	5989.65
6	2011	5341.09
7	2009	5208.39
8	2008	4681.22
9	2007	4095.26
10	2006	3577.48

Revenue By Year
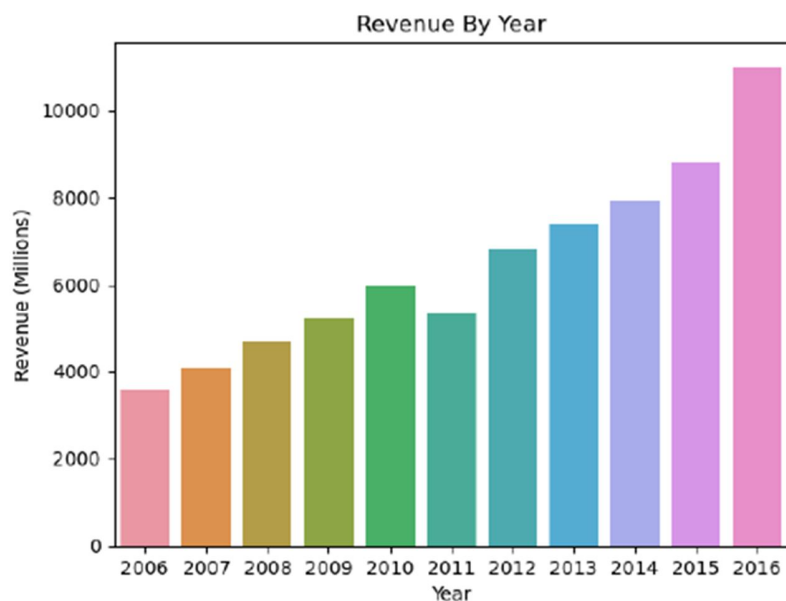
The year 2016 recorded the highest revenue, totaling $11.0 billion. The bar plot for yearly revenue clearly shows 2016 as a standout year, far surpassing previous years.

The growing trend of franchise movies and large-scale productions likely contributed to this peak. Movies like *Finding Dory* and *Rogue One* played a pivotal role in driving revenue.

4. Average Rating for each Director

|	Director	Rating
0	Christopher Nolan	8.68
1	Olivier Nakache	8.60
2	Makoto Shinkai	8.60
3	Florian Henckel von Donnersmarck	8.50
4	Aamir Khan	8.50
...	...	...
519	Sam Taylor-Johnson	4.10
520	Joey Curtis	4.00
521	George Nolfi	3.90
522	James Wong	2.70
523	Jason Friedberg	1.90

Directors like Christopher Nolan (8.68), Olivier Nakache (8.60), and Makoto Shinkai (8.60) have the highest average ratings. From the table of directors, it is clear that these filmmakers consistently deliver critically acclaimed movies.

Consistently high ratings for these directors underline their ability to deliver quality cinema that resonates with audiences. Their storytelling and directorial style contribute significantly to their high average ratings.

5. Top 10 Lengthiest Movies

Movies such as *The Hateful Eight* (187 minutes) and *Interstellar* (169 minutes) dominate the list. A horizontal bar chart illustrates how these movies compare in runtime.

Lengthy movies often feature intricate plots and rich character development, catering to dedicated fans of specific genres. Their extended duration allows for more elaborate storytelling.

Top 5 Lengtly Movies

---
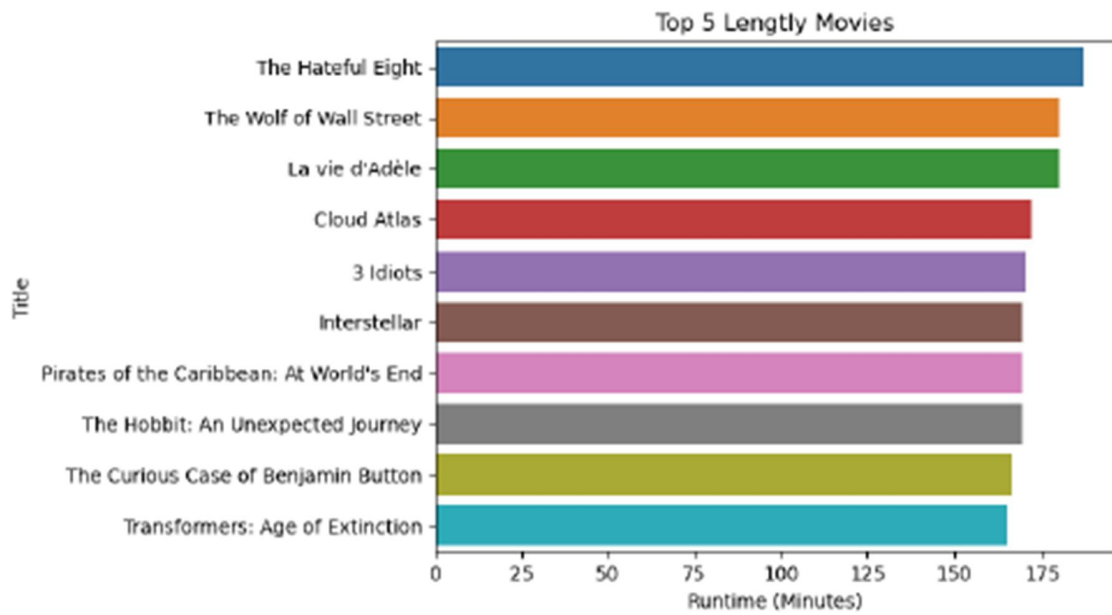
6. Number of Movies Released Per Year

Number of Movies Per Year
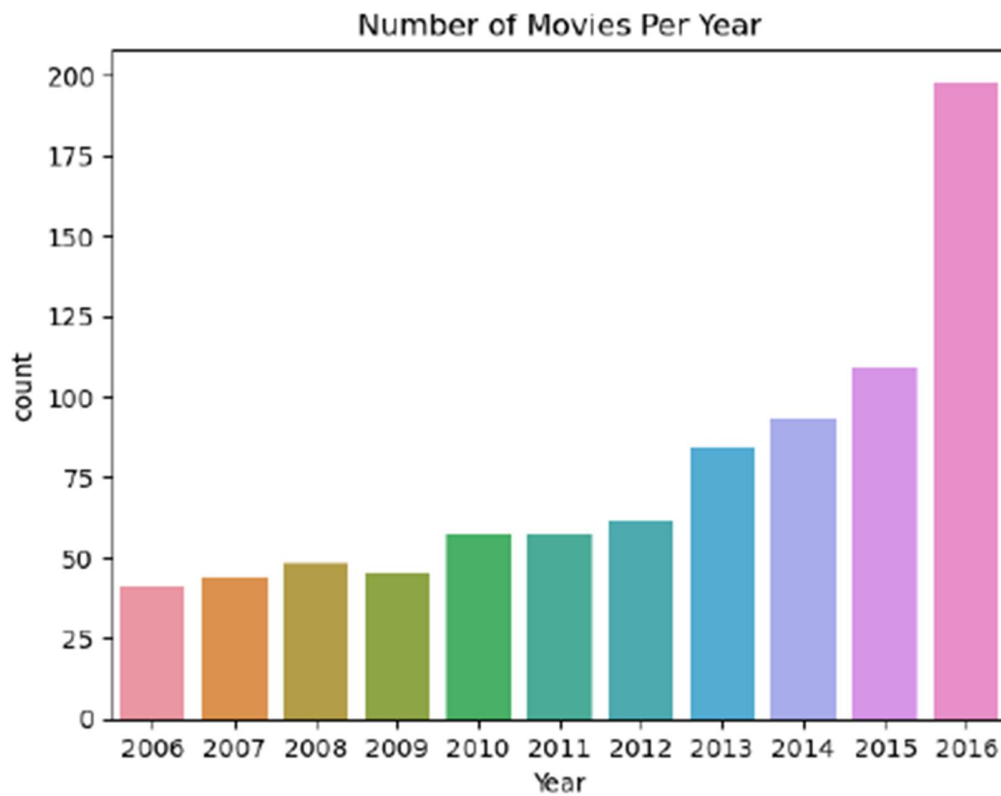
The year 2016 saw the highest number of movie releases (198), followed by 2015 (109). A count plot vividly displays the growth in movie production over the years, with a sharp increase in 2016.

The increase in movie production over the years reflects a growing demand for diverse cinematic content. This trend highlights the expanding global film industry.

---

7. Most Popular Movie Based on Revenue

Star Wars: Episode VII - The Force Awakens is the highest-grossing movie, with $936.63 million in revenue. This finding is supported by a prominent position in the revenue table and corresponding visuals.

Blockbuster franchises with established fanbases dominate revenue charts. The significant revenue of this movie underscores the commercial potential of sequels and spin-offs.

---

8. Top 10 Highest-Rated Movies and Their Directors

Title	Rating	Director
The Dark Knight	9.0	Christopher Nolan
Inception	8.8	Christopher Nolan
Interstellar	8.6	Christopher Nolan
Kimi no na wa	8.6	Makoto Shinkai
The Intouchables	8.6	Olivier Nakache
The Prestige	8.5	Christopher Nolan
The Departed	8.5	Martin Scorsese
The Dark Knight Rises	8.5	Christopher Nolan
Whiplash	8.5	Damien Chazelle
The Lives of Others	8.5	Florian Henckel von Donnersmarck

Movies like *The Dark Knight* (Christopher Nolan) and *Kimi no Na wa* (Makoto Shinkai) lead the ratings. Exceptional storytelling and critical acclaim drive high ratings. These movies often leave a lasting impact on audiences, cementing their place in cinematic history.
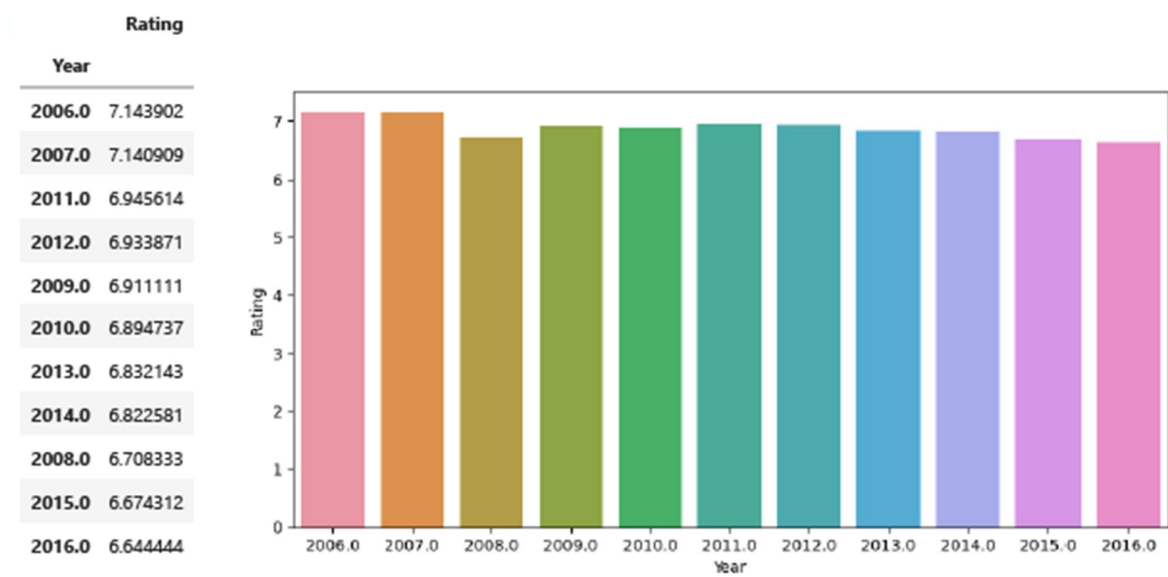
---

9. Top 10 Highest Revenue Movies
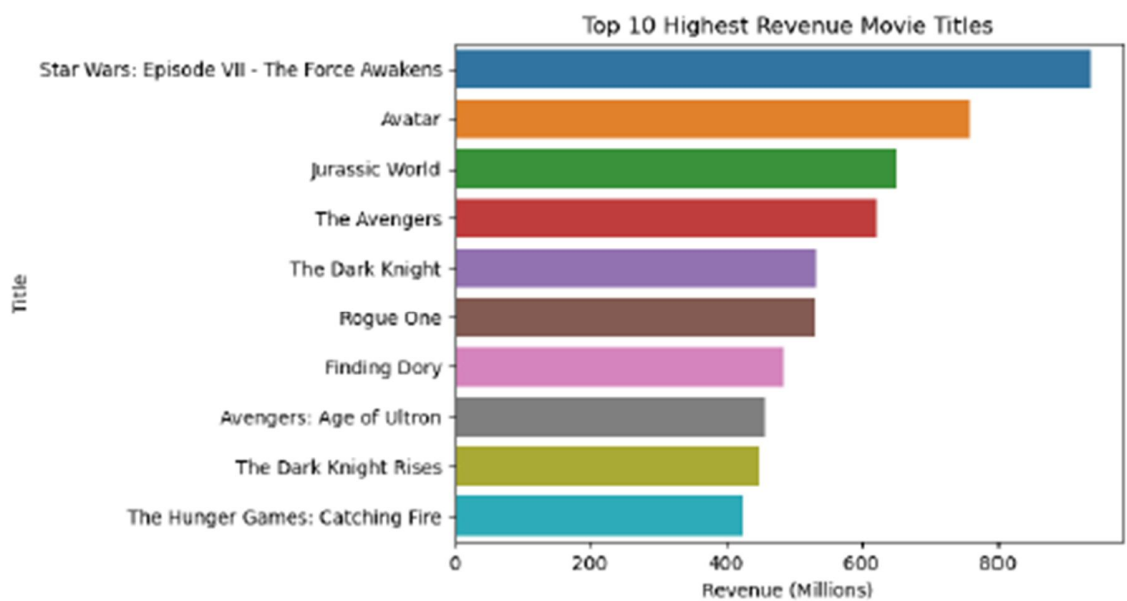
The top revenue-generating movies include *Star Wars: Episode VII*, *Avatar*, and *Jurassic World*. A revenue chart showcases their dominance over other movies. Movies with a combination of strong narratives and cutting-edge visuals attract larger audiences globally. These movies often rely on advanced technology and extensive marketing campaigns.
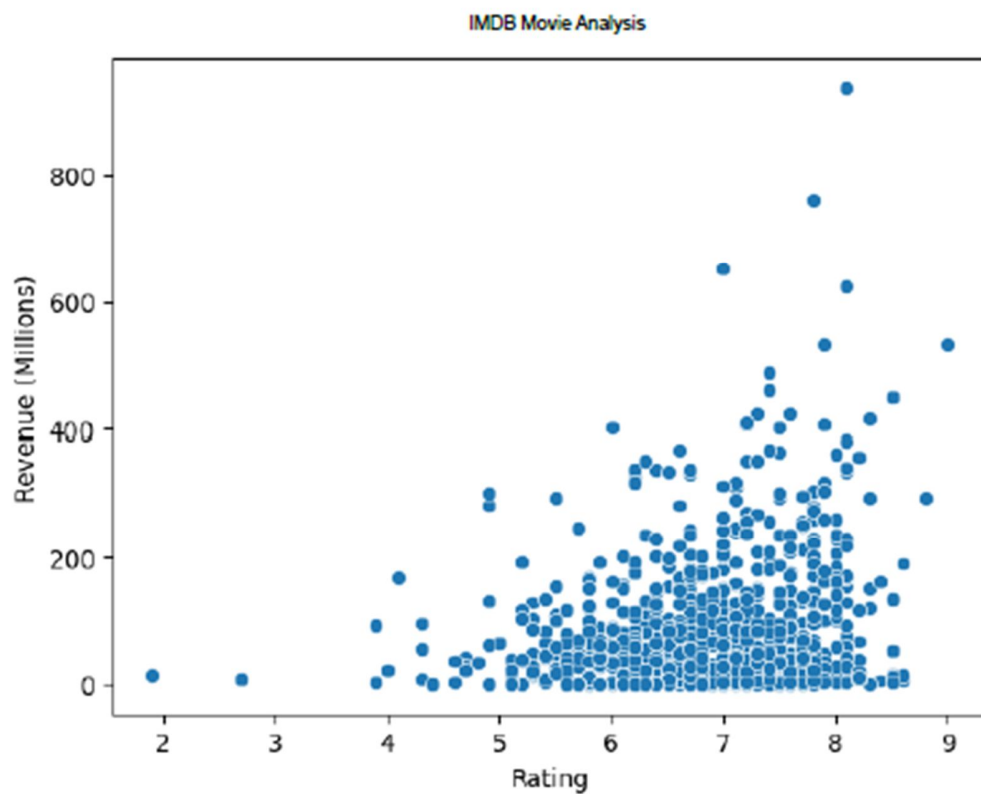
---

10. Average Rating Per Year

The highest average ratings were recorded in 2006 (7.14), while 2016 had the lowest (6.64). A bar plot reveals a gradual decline in average ratings over the years.

Early years of the dataset highlight more critically acclaimed movies, whereas recent years show diversification in genres and production styles, possibly catering to a broader audience base.

Year	Rating
2006.0	7.143902
2007.0	7.140909
2011.0	6.945614
2012.0	6.933871
2009.0	6.911111
2010.0	6.894737
2013.0	6.832143
2014.0	6.822581
2008.0	6.708333
2015.0	6.674312
2016.0	6.644444

11. Does Rating Affect Revenue?

A positive correlation exists between ratings and revenue, as depicted in the scatter plot showing a clear upward trend. High ratings often reflect positive word-of-mouth and critical reviews, boosting box office performance. This relationship emphasizes the importance of quality content in commercial success.

12. Movie Classification Based on Ratings

Movies were categorized into "Excellent" (ratings >= 7.0), "Good" (ratings >= 6.0), and "Average" (ratings < 6.0) to better understand audience preferences. A pie chart was created to depict the distribution of these categories, showing that over 60% of the movies belong to the "Excellent" category, followed by "Good" and then "Average."

This categorization reveals that the dataset predominantly features movies with high audience approval, indicating a strong inclination towards quality cinema. The dominance of "Excellent" rated movies underscores a general preference for well-made films that meet or exceed audience expectations, which could serve as a benchmark for filmmakers aiming for success.

Conclusion

This analysis of 1,000 popular IMDB movies from 2006 to 2016 highlights several key trends in the film industry. High-grossing and highly rated movies often belong to renowned directors, feature elaborate storytelling, or are part of successful franchises. The positive relationship between ratings and revenue underscores the importance of audience satisfaction and critical acclaim in driving movie success. As the industry continues to evolve, these insights provide valuable guidance for filmmakers, producers, and marketers aiming to create impactful cinema.