Movie Database Analysis

1. Introduction

This report provides an in-depth analysis of a movie database using **PostgreSQL** queries. The insights cover various aspects, including top-rated movies, industry trends, language-based analysis, and financial performance.

---

2. Key Findings

2.1 Top 5 Movies by IMDB Rating


```
SELECT title, industry, release_year, imdb_rating
FROM movies
ORDER BY imdb_rating DESC
```

LIMIT 5;

Insight:

- Identifies the highest-rated movies across industries (Hollywood, Bollywood, etc.).
- Helps in understanding audience preferences.

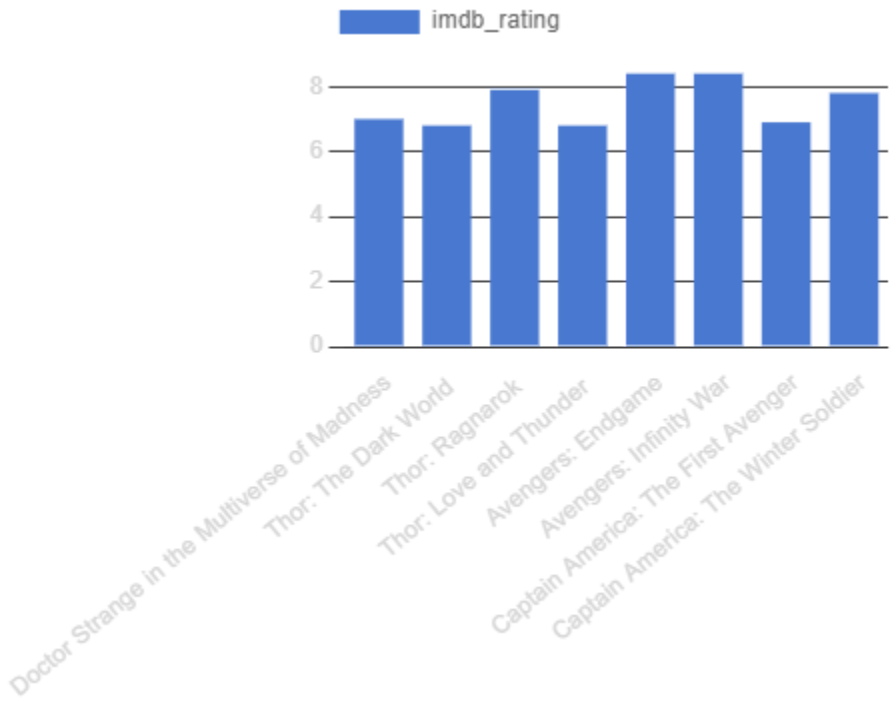
2.2 Movie Releases Over the Years


```
SELECT release_year, COUNT(movie_id) AS total_movies_released
FROM movies
GROUP BY 1
ORDER BY 1 ASC;
```

Insight:

- Tracks movie production trends year-wise.
- Useful for analyzing industry growth/decline.

2.3 Marvel Studios Movies




```
SELECT title, imdb_rating, release_year, studio
FROM movies
WHERE studio = 'Marvel Studios';
```

Insight:

- Lists all Marvel movies with ratings.
- Helps in evaluating Marvel's cinematic universe performance.

2.4 Avengers Series Movies

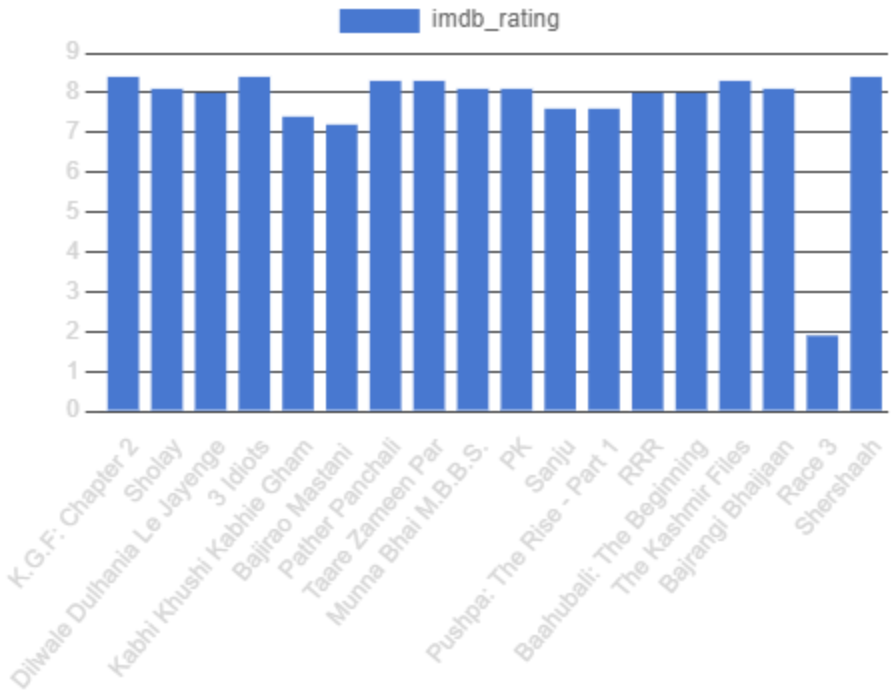
title character varying (100)	release_year integer	imdb_rating numeric (5,2)
Avengers: Endgame	2019	8.40
Avengers: Infinity War	2018	8.40
Captain America: The First Avenger	2011	6.90

```
SELECT title, release_year, imdb_rating
FROM movies
WHERE title LIKE '%Avenger%';
```

Insight:

- Focuses on the Avengers franchise.
- Shows release timeline and audience reception (via ratings).

2.5 Bollywood Movies Analysis




```
SELECT title, release_year, imdb_rating
FROM movies
WHERE industry = 'Bollywood';
```

Insight:

- Provides an overview of Bollywood films in the dataset.
- Useful for comparing with Hollywood metrics.

2.6 Movies Released in 2022

title character varying (100)	release_year integer
K.G.F: Chapter 2	2022
Doctor Strange in the Multiverse of Madness	2022
Thor: Love and Thunder	2022
RRR	2022
The Kashmir Files	2022

```
SELECT title, release_year
FROM movies
WHERE release_year = 2022;
```

Insight:

- Highlights recent movies for trend analysis.

2.7 High-Rated Movies (Post-2020, IMDB > 8)

```
SELECT title, release_year, imdb_rating
FROM movies
WHERE release_year > 2020 AND imdb_rating > 8
ORDER BY 3 DESC;
```

Insight:

- Identifies critically acclaimed recent films.

2.8 Best Movies (2015–2020, IMDB > 8)

```
SELECT title, release_year, imdb_rating
FROM movies
WHERE release_year BETWEEN 2015 AND 2020 AND imdb_rating > 8
```

```
ORDER BY 3 DESC;
```

Insight:

- Helps in understanding high-quality productions in a specific period.

2.9 Thor Movie Series (Ordered by Year)

```
SELECT title, release_year
FROM movies
WHERE LOWER(title) LIKE '%thor%'
ORDER BY 2 DESC;
```

Insight:

- Chronicles the Thor franchise evolution.

2.10 Most Common Movie Languages

```
SELECT l.name, COUNT(DISTINCT movie_id) AS number_of_movies
FROM movies m
LEFT JOIN languages l ON l.language_id = m.language_id
GROUP BY 1
ORDER BY 2 DESC;
```

Insight:

- Shows language distribution in the dataset.
- Useful for localization strategy.

2.11 Bengali & Telugu Movies

```
-- Bengali Movies
SELECT l.name, m.title
FROM movies m
LEFT JOIN languages l ON l.language_id = m.language_id
WHERE l.name = 'Bengali';

-- Telugu Movies
SELECT m.title, l.name
FROM movies m
LEFT JOIN languages l ON l.language_id = m.language_id
WHERE l.name = 'Telugu';
```

Insight:

- Regional film industry analysis.

2.12 Revenue of Hindi Movies (Converted to Millions)

```
SELECT title,
   CASE
      WHEN unit = 'Thousands' THEN ROUND(revenue/1000, 2)
      WHEN unit = 'Billions' THEN ROUND(revenue*1000, 2)
      ELSE revenue
   END AS revenue_millions
FROM movies m
INNER JOIN finantials f ON f.movie_id = m.movie_id
INNER JOIN languages l ON l.language_id = m.language_id
WHERE l.name = 'Hindi'
ORDER BY 2 DESC;
```

Insight:

- Financial performance of Hindi cinema.

2.13 Highest-Grossing Hindi Movie

```
SELECT title,
   CASE
      WHEN unit = 'Thousands' THEN ROUND(revenue/1000, 2)
      WHEN unit = 'Billions' THEN ROUND(revenue*1000, 2)
      ELSE revenue
   END AS revenue_millions
FROM movies m
INNER JOIN finantials f ON f.movie_id = m.movie_id
INNER JOIN languages l ON l.language_id = m.language_id
WHERE l.name = 'Hindi'
ORDER BY 2 DESC
LIMIT 1;
```

Insight:

- Identifies the most successful Hindi film.

2.14 Database Time Range

```
SELECT
   MAX(release_year) AS recent_year,
   MIN(release_year) AS starting_year
FROM movies;
```

Insight:

- Defines the dataset's temporal scope.

2.15 Movies Exceeding Average IMDB Rating

```
SELECT title, imdb_rating
FROM movies
WHERE imdb_rating > (SELECT AVG(imdb_rating) FROM movies)
ORDER BY 2 DESC;
```

Insight:

- Highlights above-average-rated movies.

---

3. Conclusion

This analysis provides actionable insights into:
✔ **Top-performing movies** (ratings & revenue).
✔ **Industry trends** (Marvel, Bollywood, regional cinema).
✔ **Financial success metrics** (highest-grossing films).
✔ **Language-based distribution** (Hindi, Bengali, Telugu).

Next Steps:

- Further segmentation by genre.
- Comparative analysis between studios.
- Audience sentiment correlation with revenue.