IMDB MOVIES ANALYSIS USING SQL

presented by
Nithin N K
Vishnu Vardhan
Siva prakash
Venkatesh

Project Team ID: PTID-CDA-DEC-25-1024
PRSQL-02: IMDB MOVIES

Project overview

- Analyze IMDB movie and director data using SQL
- Support decision-making through data-driven insights
- Identify popular, bankable movies and directors
- Understand trends in directors and movies

Dataset overview

- Database includes: IMDB
- Tables:

    Movies

    Directors
- Key Columns:

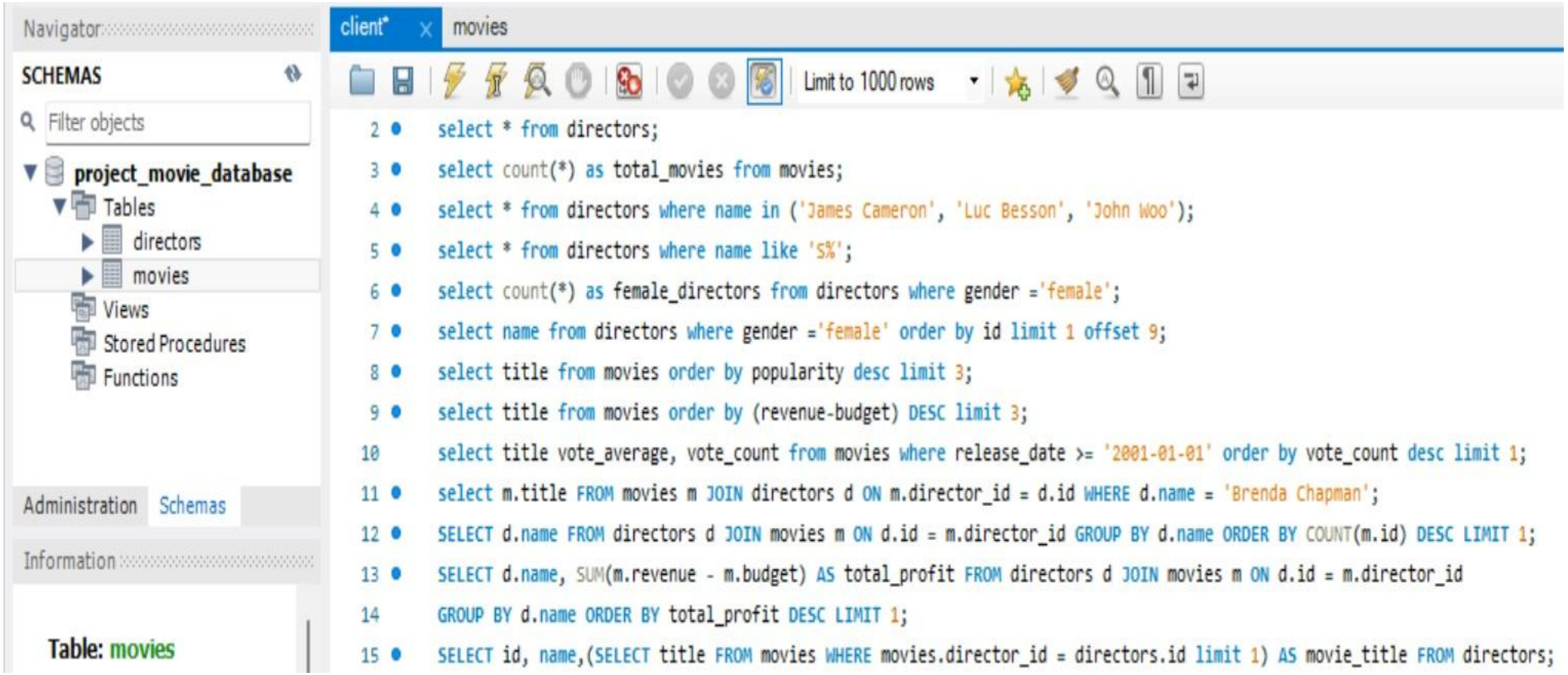
Movies: id, title, release_date, vote_average, revenue, budget, diretor_id.

Directors: id, name, gender.

Problem Statement

- Movie & Director information retrieval
- Director analysis (female directors, top directors)
- Popular & bankable movies
- Bankable directors

IMDB Movies Database Analysis Using SQL



Data Extraction Queries

- Getting Movies data

```
Select * from movies;
```

- Outcomes : The query retrieves all records and columns from the movies table.
- Insights: The dataset contains both budget and revenue and movies with higher budgets generally show high revenue potential

- Getting Directors data

```
Select * from directors;
```

- Outcomes: The query retrieves all records and columns from the directors table.
- Insights : Contains list of professional movie directors.

Total Movies in Dataset

- Movies count in IMDB

```
Select count(*) as total_movies from
movies;
```

- Outcomes : The movies table contains 47 records.
- Insights : The data set is small but manageable.

- Identifying selected directors in IMDB

```
Select * from directors where  name in
("James Cameron", "Luc Besson","John
Woo");
```

- Outcomes: The query returned 3 records
- Insights: Their presence confirms that the dataset includes globally recognized directors

Directors starting with S & Female director count

- Director's name starting with S.

```
select * from directors where name
like "s%";
```

- Outcomes: Successfully returned directors whose names start with “s”.
- Insights: the dataset have names starting with the letter “s”.

- female directors in IMDB

```
select count(*) as female_directors from
directors where gender ='female';
```

- Outcomes: The query returns 625 female directors in the database.
- Insights : The result supports diversity and inclusion reporting in film industry analysis.

Insights on Female Directors and Movie Popularity

- 10th first women directors

```
Select name from directors where  gender = "female"order by id limit 1 offset 9;
```

- Outcomes: The query returns “Hoyt Yeatman” as the 10th female director (based on ordered IDs).
- Insights: The query identifies the 10th female director in the data set .

- 3 most popular movies

```
select original_title,popularity from movies order by popularity desc limit 3;
```

- Outcomes : The query returns top 3 most popular movies.
- Insights : Jurassic world has significantly higher popularity than the other movies.

Box Office Performance and Audience Engagement

- 3 Most bankable movies

```
select original_title,revenue from
movies order by revenue desc limit 3 ;
```

- Outcomes: Identified top 3 bankable movies.
- Avatar, Titanic, The avengers
- Insights : Avatar has the highest revenue among all movies and titanic ranks second,showing long term box office strength

- Most awarded average vote since the January 1st, 2000?

```
Select title,vote_average,vote_count from movies
where release_date >= '2000-01-01’order  by
vote_count desc limit 1;
```

- Outcomes :Avatar has a vote average of 7.2.
- Insights : Avatar has the highest audience engagement among movies released after 2000.

Director Contribution and Movie Output Analysis

- Movie(s) were directed by Brenda Chapman.

```
select m.title from movies m join directors d
on m.director_id = d.id where d.name =
'Brenda Chapman';
```

- Outcomes : Brenda Chapman has no movies recorded.
- Insights : The query correctly uses inner join to link movies with directors using director_id.

- Director made the most movies.

```
Select d.name, count(m.id) as movie_count from
directors d join movies m  on d.id = m.director_id
group by d.name order  by movie_count desc
limit 1;
```

- Outcomes : identified Director with the highest number of movies.
- Insights : Gore Verbinski has directed the most movies in the dataset (3 movies).

Most Bankable Director Based on Profitability

- Which director is the most bankable.

```
Select d.name, sum(m.revenue - m.budget) as total_profit from
directors d join movies m  on d.id = m.director_id group by d.name
order by total_profit desc limit1;
```

- Outcomes : Identified director with the highest total profit.
- Insights : james cameron generates the highest total profit, indicating exceptional box-office success relative to budgets.

Conclusion

- In this project, we used SQL to analyze the IMDB movies database and uncover meaningful insights about movies and directors. By running structured queries, we explored movie popularity, box office performance, director productivity, profitability, and gender representation.
- The analysis helped identify the most popular and bankable movies, showing how audience engagement and revenue indicate commercial success. We also highlighted top-performing and profitable directors while gaining insights into female director representation.
- Overall, this project shows how SQL can turn raw data into useful insights that support data-driven decisions in the entertainment industry.