

INTERNSHIP PROJECT

PROJECT TEAM ID: PTID-CDA-MAR-23-138

PROJECT ID: PRSQL-01

PROJECT NAME: IMDB Movies

MEMBER NAME: P SHIVA KUMAR REDDY

TOOLS USED: SQL

Step-1 Connected to the database server using given IP Address, username, password

For the SQL queries have written the queries as follows

USE project_movie_database;

-- seeing all the data in the tables

SELECT * FROM directors;

SELECT * FROM movies;

Step-2

QUERIES

-- a) Can you get all data about movies?

SELECT * FROM movies;

-- b) How do you get all data about directors?

SELECT * FROM directors;

-- c) Check how many movies are present in IMDB.

SELECT COUNT(*) AS total_movies FROM movies;

-- d) Find these 3 directors: James Cameron ; Luc Besson ; John Woo

SELECT * FROM directors WHERE name IN ('James Cameron', 'Luc Besson', 'John Woo');

-- e) Find all directors with name starting with Steven.

SELECT * FROM directors WHERE name LIKE 'Steven%';

-- f) Count female directors.

SELECT COUNT(*) AS female_directors FROM directors WHERE gender = 1;

-- g) Find the name of the 10th first women directors?

SELECT name FROM directors WHERE gender = 1 ORDER BY id LIMIT 9,1;

-- h) What are the 3 most popular movies?

SELECT * FROM movies ORDER BY popularity DESC LIMIT 3;

-- i) What are the 3 most bankable movies?

SELECT * FROM movies ORDER BY revenue DESC LIMIT 3;

-- j) What is the most awarded average vote since the January 1st, 2000?

SELECT

AVG(vote_average) AS avg_vote

```
FROM
  movies
WHERE
  release_date >= '2000-01-01'
ORDER BY avg_vote DESC
LIMIT 1;
```

-- k) Which movie(s) were directed by Brenda Chapman?

-- as it is a join on using common id from both the tables

```
SELECT movies.*
FROM movies
JOIN directors ON movies.director_id = directors.id
WHERE directors.name = 'Brenda Chapman';
```

-- l) Whose director made the most movies?

-- as it is a join on using common id from both the tables

```
SELECT directors.name AS director_name, COUNT(*) AS movie_count
FROM movies
JOIN directors ON movies.director_id = directors.id
GROUP BY directors.name
ORDER BY movie_count DESC
LIMIT 1;
```

-- m) Whose director is the most bankable?

-- as it is a join on using common id from both the tables

```
SELECT directors.name AS director_name, SUM(movies.revenue) AS total_revenue
FROM movies
JOIN directors ON movies.director_id = directors.id
GROUP BY directors.name
ORDER BY total_revenue DESC
LIMIT 1;
```