CSCI 5408

DATA MANAGEMENT AND WAREHOUSING

LAB ASSIGNMENT - 1

Submitted By: Kenil Shaileshkumar Patel (kenil.patel@dal.ca)
Banner ID: B00954251
Submitted On: September 20, 2023

Gitlab Repository Link

https://git.cs.dal.ca/kenil/csci5408_f23_b00954251_kenil_patel/-/tree/main/Lab1

Table of Contents

Query – 1

Check how many unique actors are present in the IMDB dataset.

```
SELECT COUNT(DISTINCT CONCAT(first_name, ' ', last_name)) AS
Total_No_Of_Unique_Actors FROM lab_1.actors;
```

In this query, first name and last name of actors are concatenated. Then it counts the distinct concatenated names and we get the Total number of unique actors from the actor's table. Total_No_Of_Unique_Actors is an alias to label the result of this count for better readability.

Solution:-

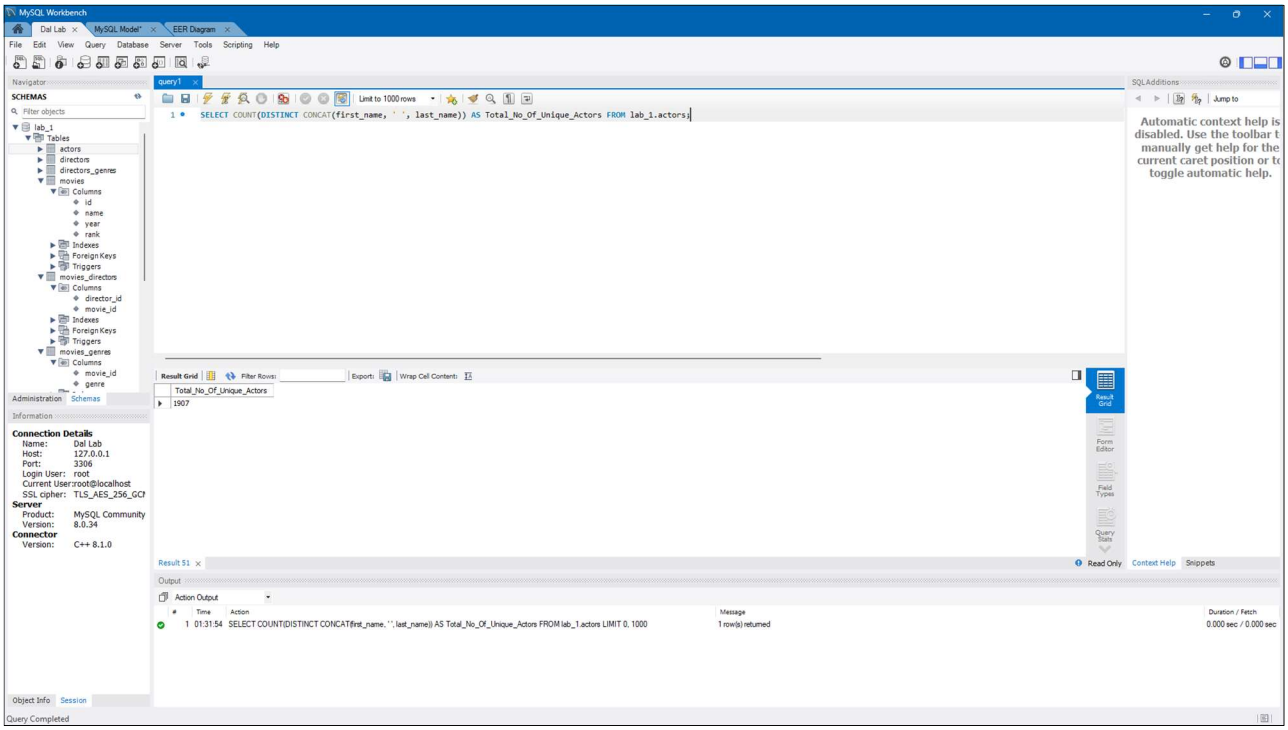

Figure 1 Output of Query-1

Query – 2

Check how many movies were released between the years 1990s and 2000.

SELECT COUNT(*) AS Count_of_movies_btw_1990_to_2000 FROM lab_1.movies WHERE year BETWEEN 1990 AND 2000;

In this query, it counts the total no of movies that were released in a year between 1990 and 2000. And alias to the count is Count_of_movies_btw_1990_to_2000 for better presentation.

Solution:-

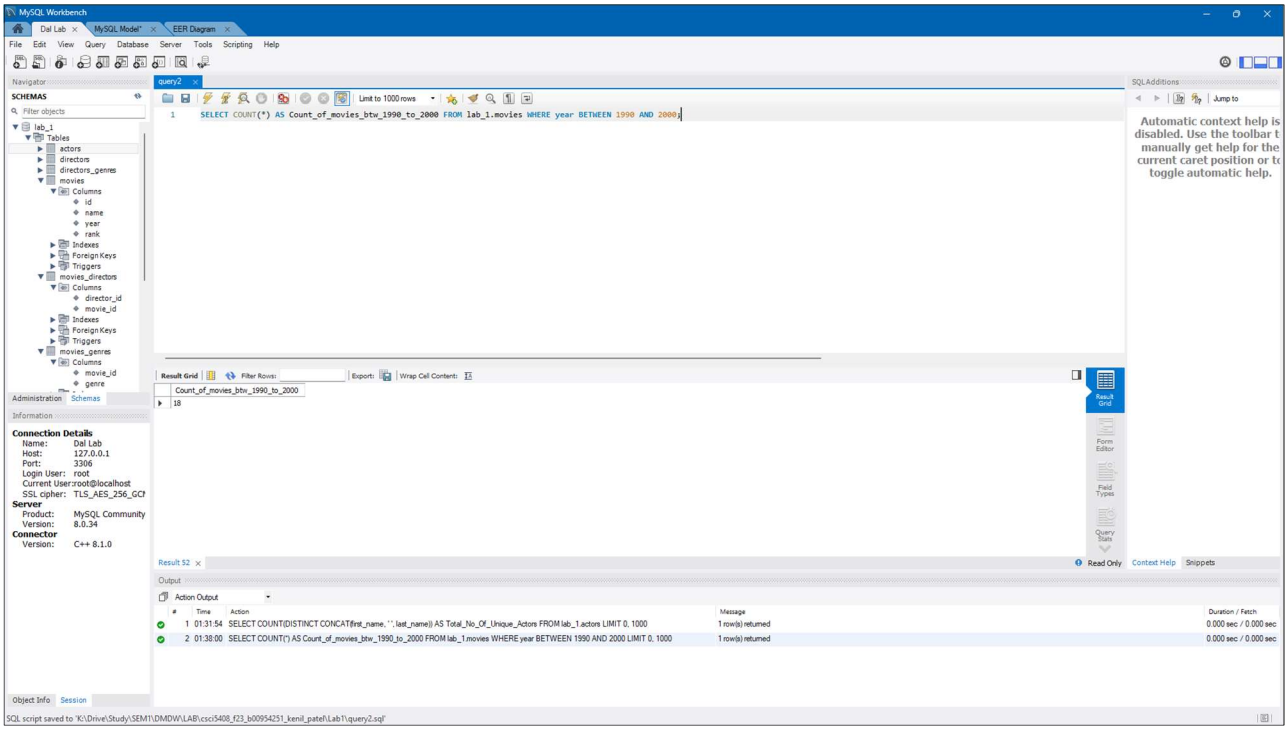

Figure 2 Output of Query-2

Query – 3

Find the list of genres of movies directed by Christopher Nolan.

```
SELECT DISTINCT lab_1.directors_genres.genre FROM lab_1.directors
RIGHT JOIN lab_1.directors_genres ON lab_1.directors.id = lab_1.directors_genres.director_id
WHERE CONCAT(lab_1.directors.first_name, ' ', lab_1.directors.last_name) = "Christopher
Nolan";
```

This query extracts the distinct genres which are associated with the director "Christopher Nolan". I have done the right join on directors and directors_genres based on directors IDs and filtering results on the basis of specific directors.

Solution:-

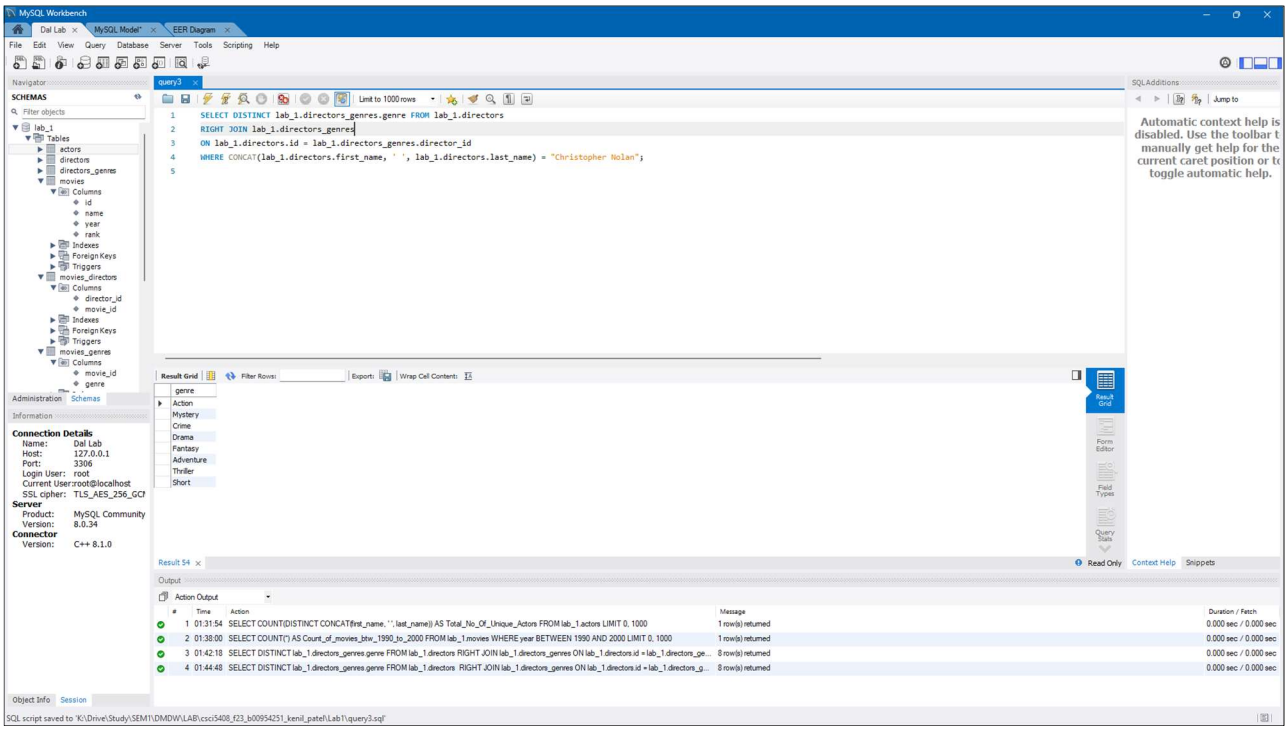

Figure 3 Output of Query-3

Query – 4

Find the list of all directors and the movie name which are ranked between 8 to 9 and have a genre of Sci-Fi and Action.

```
SELECT M.name AS movie_name, CONCAT(D.first_name, " ", D.last_name) AS director_name
FROM lab_1.movies M
JOIN lab_1.movies_directors MD ON M.id = MD.movie_id
JOIN lab_1.directors D ON MD.director_id = D.id
WHERE M.id IN (
   SELECT MG.movie_id
   FROM lab_1.movies_genres MG
   WHERE MG.genre IN ('Action', 'Sci-Fi')
   GROUP BY MG.movie_id
   HAVING COUNT(DISTINCT MG.genre) = 2
) AND (M.rank BETWEEN 8 AND 9);
```

First, in the sub-query, extract all movie_id from the movies_genres table which has action and sci-fi genre both in a movie. After, getting movie_id, I joined on movies_directors table and got the directors_id and another joined the director's table to get the director's name.

Solution:-

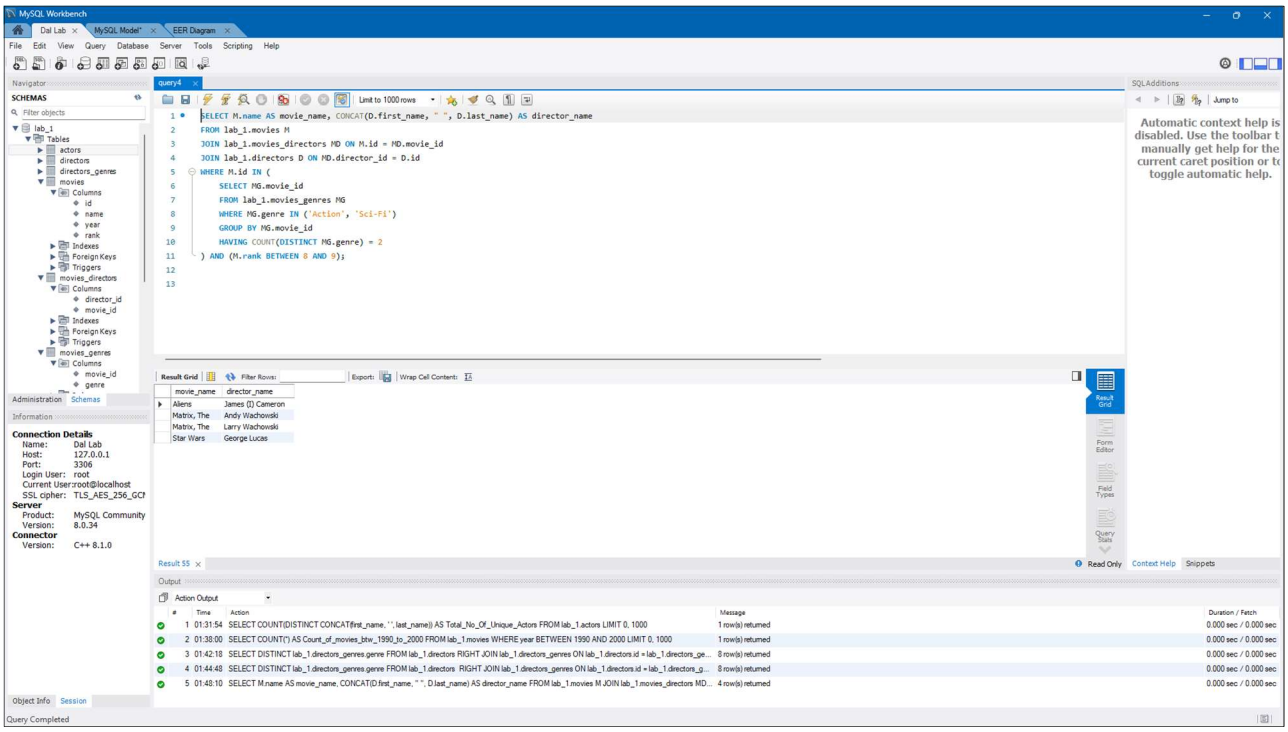

Figure 4 Output of Query-4

Query – 5

Find the name of the movie in which the actor's role is any doctor, and the movie has the highest number of roles of a doctor.

```
SELECT M.name FROM lab_1.movies M
WHERE M.id IN
(SELECT R.movie_id
FROM lab_1.roles R
WHERE R.role LIKE '%Doctor%'
GROUP BY R.movie_id
ORDER BY COUNT(*) DESC
)
LIMIT 1;
```

This query retrieves the movie name for the movie with the highest count of characters that have doctors' roles. And subquery counts the total doctor roles in a movie and returns a movie_id. The outer query selects the movie name from the top row which has the highest count.

Solution:-

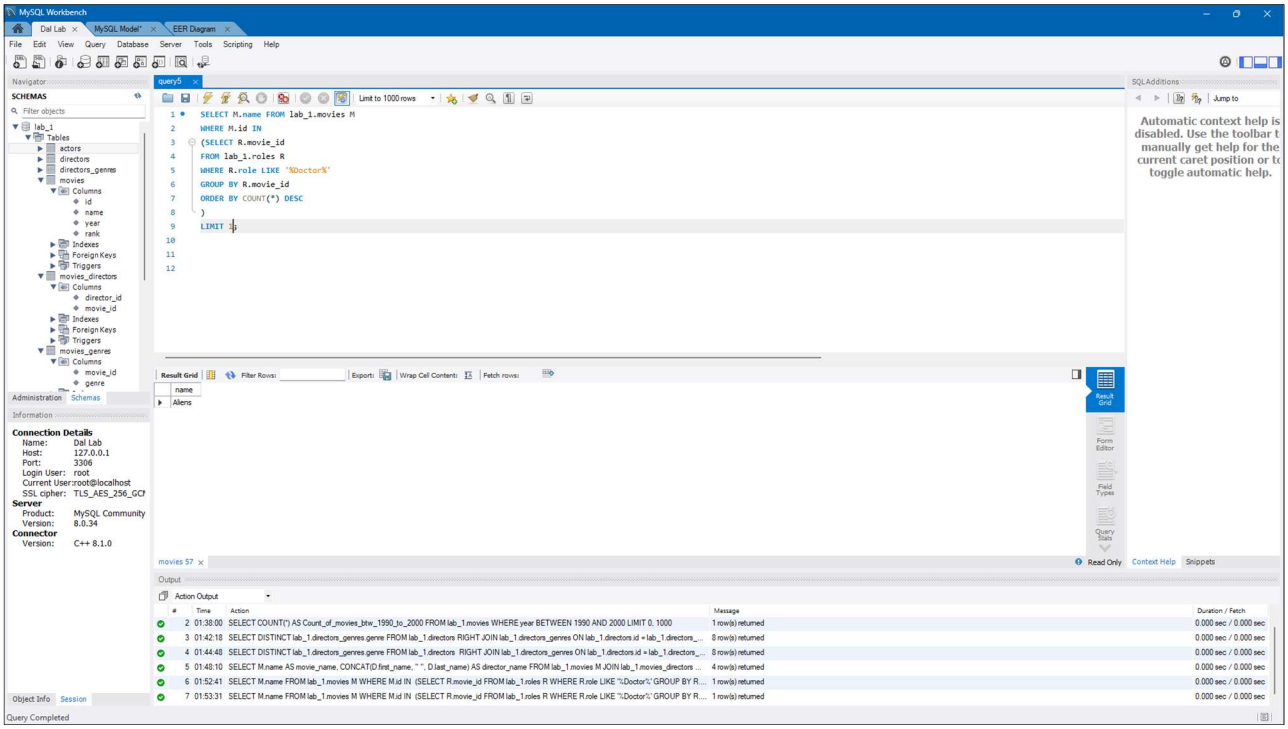

Figure 5 Output of Query-5

Query – 6

Find the list of the movies that start with the letter ‘f’.

SELECT lab_1.movies.name AS movie_name FROM lab_1.movies WHERE lab_1.movies.name LIKE 'F%';

This query finds the names of the movies that start with F. So in LIKE it checks that the first letter is capital F and the rest can be anything.

Solution:-

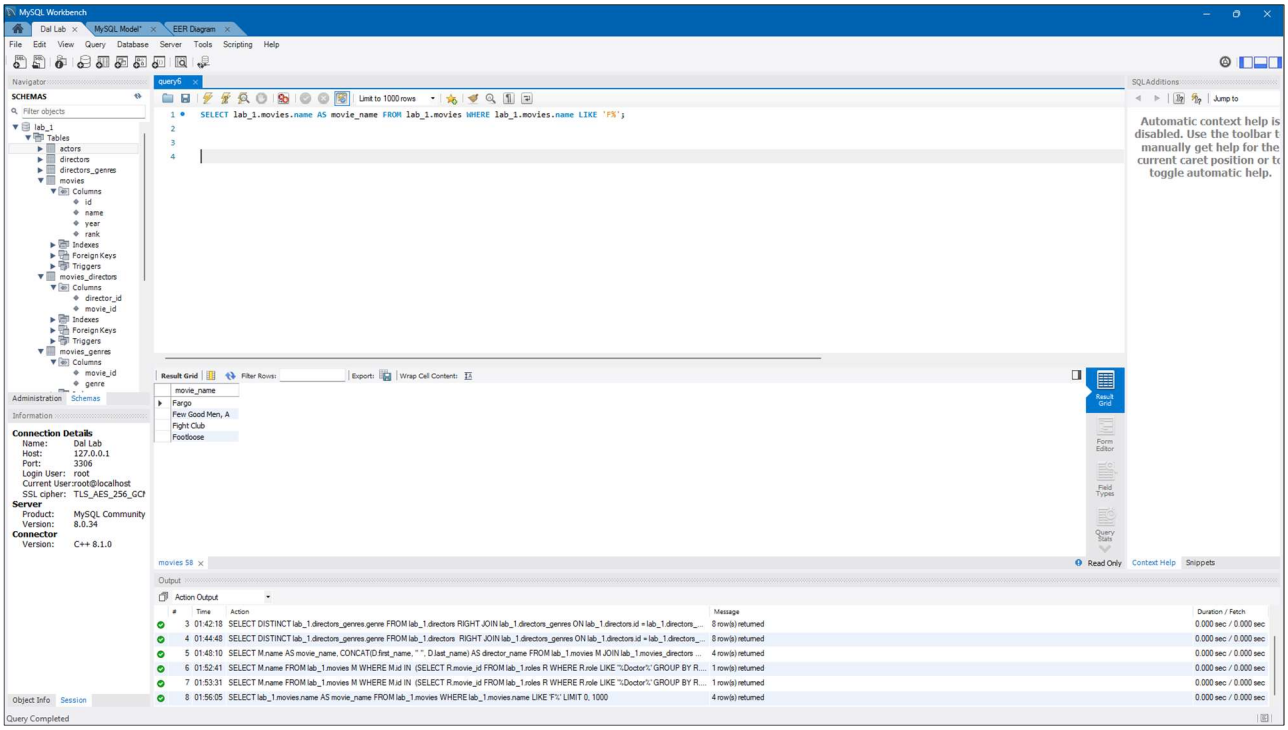

Figure 6 Output of Query-6