Data Management, Warehousing, And Analytics

Lab1: Introduction to MySQL
(Summer 2023)

Submitted by: Arihant Dugar (B00917961)

GitLab repo: https://git.cs.dal.ca/dugar/csci5408_s23_b00917961_arihant_dugar/-/tree/main/Lab1

Summary:

In Lab1 MySQL, we covered several important topics related to MySQL database management. Firstly, we learned about the installation process for MySQL and how to set up a database. We also explored different types of SQL queries, including Data Definition Language (DDL) and Data Manipulation Language (DML) queries, and how they are used to create, modify, and retrieve data from a database.

Constraints, such as primary keys, foreign keys, and unique constraints, were discussed in detail, and their importance in maintaining data integrity was highlighted. We also looked at subqueries and how they can be used to retrieve data from a database using multiple queries. Joins were another important topic covered in the lab. We learned about different types of joins, such as inner join, left join, and right join, and how they can be used to combine data from multiple tables.

Finally, we explored how to import and export data from a MySQL database. We learned about different file formats that can be used for importing and exporting data, and the various options available for customizing these processes.

Steps Performed:

1. Download and install MySQL workbench [1].
2. Created a new connection in localhost to work on the lab exercise.
3. Created a new schema in localhost named 'imdb'.
4. Downloaded the IMDB dataset provided for the lab exercise [2].
5. Imported the IMDB.SQL according to the steps conveyed into MySQL workbench.
6. Verified the data present in the schema for the imported file.
7. Started working on queries for the lab exercise and executed each query and took a screenshot of its output.
8. Added the technical details along with the screenshot in the lab report.

Lab Exercise:

1. Check how many directors are present in iMDB

   The number of director's present is 34

Below is the query used to fetch the count –

```
SELECT count(*) FROM imdb.directors;
```

Figure 1.1: The query to fetch the number of directors present in iMDB

The output in MySQL workbench –

count(*)
34

Figure 1.2: The result for the query mentioned in figure 1.1

2. Check how many movies are released post-year 2000

   The number of movies released post-year 2000 is 10

   Below is the query to fetch the count –

```
SELECT COUNT(*) AS movie_count
FROM imdb.movies
WHERE year > 2000;
```

Figure 2.1: The query to fetch the number of movies released post-year 2000

The output in MySQL workbench –

movie_count
10

Figure 2.2: The result for the query mentioned in figure 2.1

3. Find the list of genres of movies directed by Andrew Adamson

   We join movies_genres, movies_directors and directors on movie id and director id respectively to fetch the distinct movie genres for a particular director [3].

   Below is the query to fetch the list of distinct genres of movies by Andrew Adamson-

```
SELECT DISTINCT genre FROM imdb.movies_genres mg
INNER JOIN imdb.movies_directors md ON md.movie_id = mg.movie_id
INNER JOIN imdb.directors d ON md.director_id = d.id
WHERE first_name = "Andrew" AND last_name = "Adamson"
```

Figure 3.1: The query to fetch list of genres of movies directed by Andrew Adamson

The output for the above query is –

genre
Adventure
Animation
Comedy
Family
Fantasy
Romance

Figure 3.2: The result for the query mentioned in figure 3.1

4. List of directors whose movies are ranked between 7 to 8 ranking

   The below query fetches the director's details and concatenate the first name and last name and returns the director's name whose movies are ranked between 7 and 8 ranking –

```
SELECT DISTINCT concat(first_name , " ", last_name) as name from imdb.movies_directors md
INNER JOIN imdb.directors d ON md.director_id = d.id
INNER JOIN imdb.movies m ON m.id = md.movie_id
WHERE m.rank >= 7 AND m.rank <= 8
```

Figure 4.1: The query to fetch list of directors whose movies are ranked between 7 to 8 ranking

The output is as below –

name
John (I) Landis
Ron Howard
Rob Reiner
Oliver (I) Stone
Ron Clements
John Musker
Sofia Coppola
Ethan Coen
Joel Coen
Steven Soderbergh
Mike (I) Judge
Darren Aronofsky
John (I) Hughes
Guy Ritchie
David Koepp

Figure 4.2: The result for the query mentioned in figure 4.1

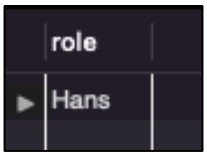
5. Find the role of Julliet Akinyi in Lost in Translation movie

   The below query fetches the role details for the given person –

```
SELECT role FROM imdb.roles r
INNER JOIN imdb.actors a ON a.id = r.actor_id
INNER JOIN imdb.movies m ON r.movie_id = m.id
WHERE a.first_name = "Julliet" AND a.last_name = "Akinyi" AND m.name = "Lost in Translation";
```

Figure 5.1: The query to fetch the role of Julliet Akinyi in Lost in Translation movie

   The output states that the role for Julliet Akinyi is Hans –



role
Hans

Figure 5.2: The result for the query mentioned in figure 5.1

6. List of the movies that contain the letter ‘o' in any position

   There is a total of 20 movies that contain the letter ‘o’ in any position. The query to fetch the data is as below –

```
SELECT * FROM imdb.movies WHERE name LIKE '%o%';
```

Figure 6.1: The query to fetch list of the movies that contain the letter ‘o' in any position

   And the output is –

id	name	year	rank
17173	Animal House	1978	7.5
18979	Apollo 13	1995	7.5
109093	Fargo	1996	8.2
111813	Few Good Men, A	1992	7.5
116907	Footloose	1984	5.8
130128	Godfather, The	1972	9
147603	Hollow Man	2000	5.3
176711	Kill Bill: Vol. 1	2003	8.4
176712	Kill Bill: Vol. 2	2004	8.2
194874	Lost in Translation	2003	8
210511	Memento	2000	8.7
237431	O Brother, Wher…	2000	7.8
238072	Ocean's Eleven	2001	7.5
238695	Office Space	1999	7.6
256630	Pirates of the Ca…	2003	NULL
257264	Planes, Trains &…	1987	7.2
267038	Pulp Fiction	1994	8.7
276217	Reservoir Dogs	1992	8.3
297838	Shawshank Red…	1994	9
314965	Stir of Echoes	1999	7

Figure 6.2: The result for the query mentioned in figure 6.1

References:

[1] "MySQL :: Download MySQL Workbench." MySQL, Oracle Corporation[Online]. Available: https://dev.mysql.com/downloads/workbench/. [Accessed: May 11, 2023].

[2] “IMDB_dataset”. IMDb Movie Dataset [SQL]. Available: https://dal.brightspace.com/d2l/le/content/271677/viewContent/3623394/View

[3] "SQL Inner Join.", W3Schools [Online]. Available: https://www.w3schools.com/sql/sql_join_inner.asp . [Accessed: May 11, 2023].