ALI ABDALLAH ABDUL

PLP ACADEMY DATABASE ASSIGNMENT

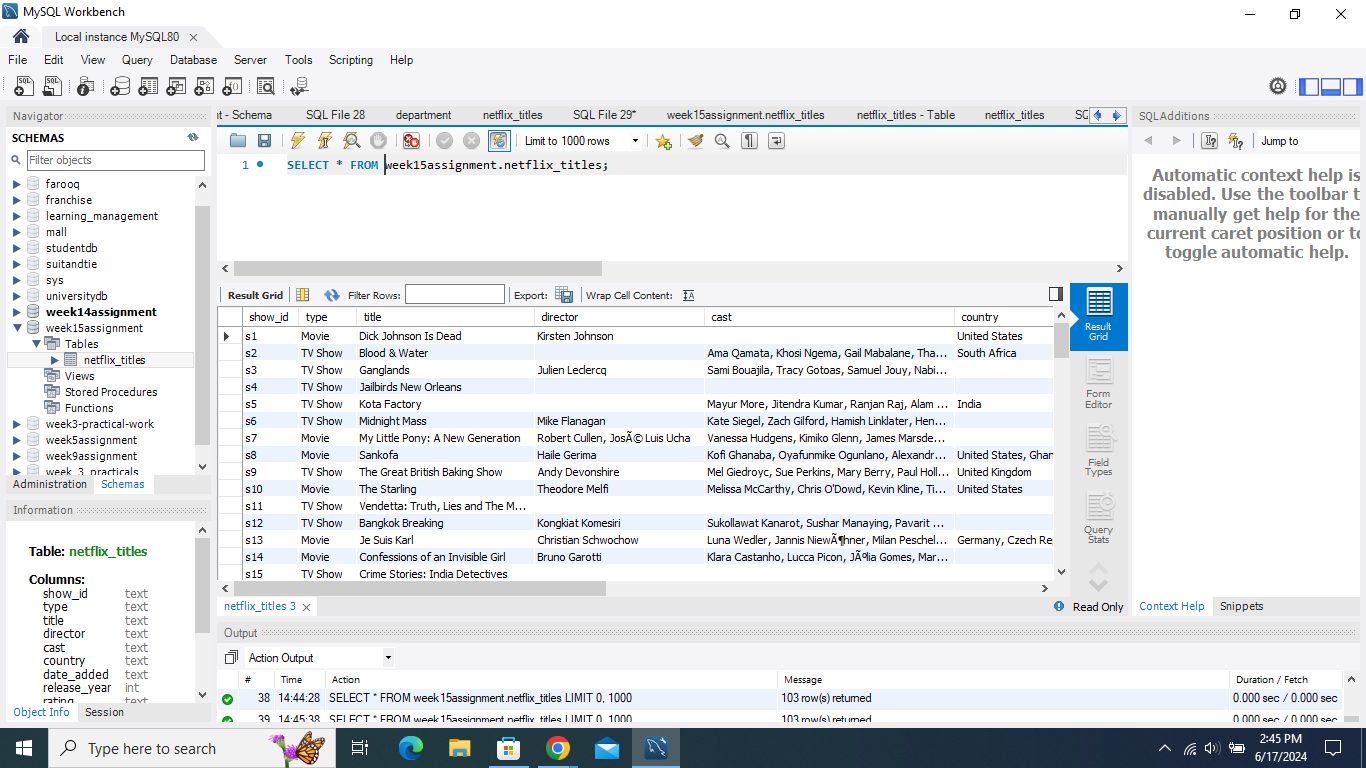
WEEK 18 DATA ANALYSIS

2024 FEB COHORT

**DATASET FOR DATA ANALYSIS NAME: netlixt\_titles**

1. Briefly explain any difficulties and 1 interesting thing you noticed about your chosen dataset.

* It would have been better when data under the Type column are listed alphabetically for uniformity
* It is a big dataset with a table of many columns spread out the full width of the screen. This is making the data reading tedious, and so difficult to draw any meaningful decision from it.
* The data would look more organized with the Show\_id listed at the first column instead of 4th column.



*Figure1: showing table listing movies and TV Shows title, director, cast and country only entries in the dataset*

1. Use simple SQL queries to play with the data.

* How many Movie Shows

We have 57 entries

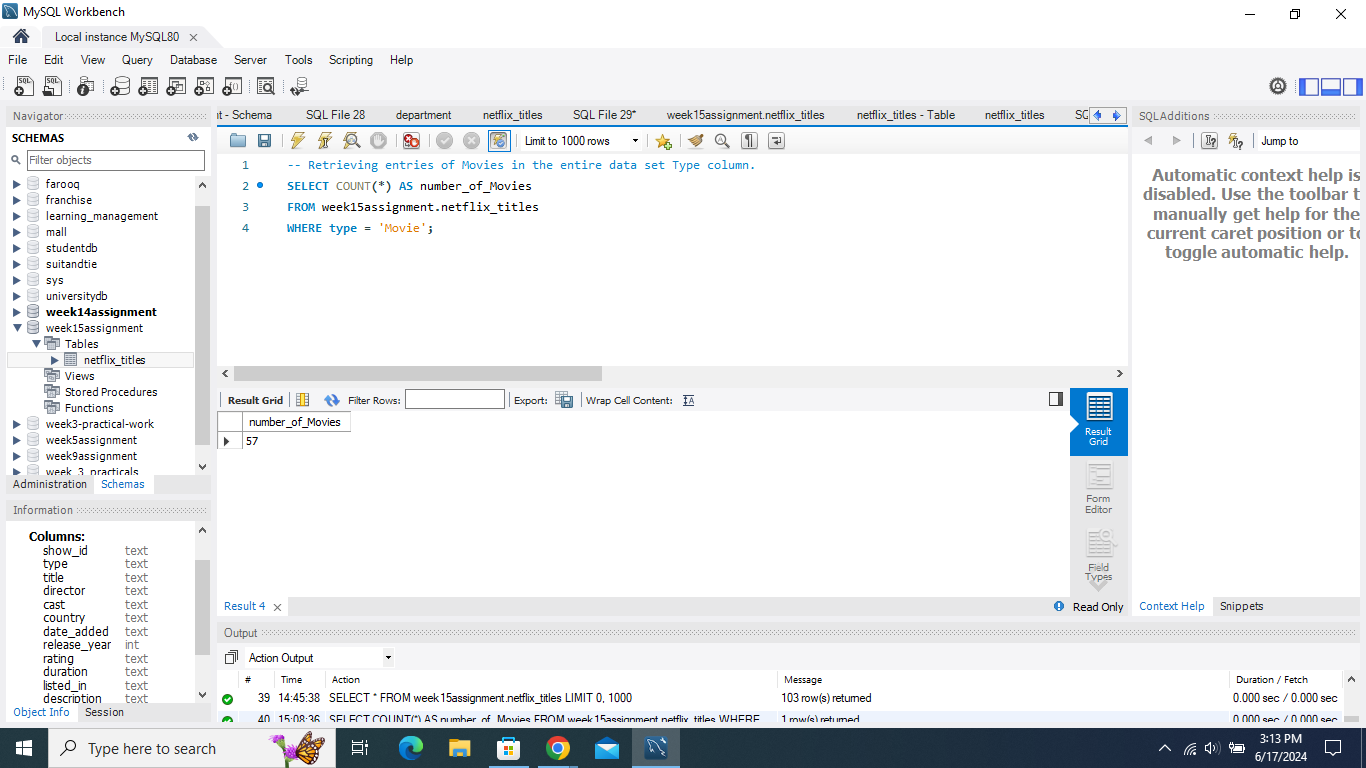
Syntax

*-- Retrieving entries of Movies in the entire data set Type column.*

*SELECT COUNT(\*) AS number\_of\_Movies*

*FROM week15assignment.netflix\_titles*

*WHERE type = 'Movie';*

**

*Figure 2: table showing retrieval of movie entries in the dataset*

* How many TVshows listed in the dataset.

We have 46 entries

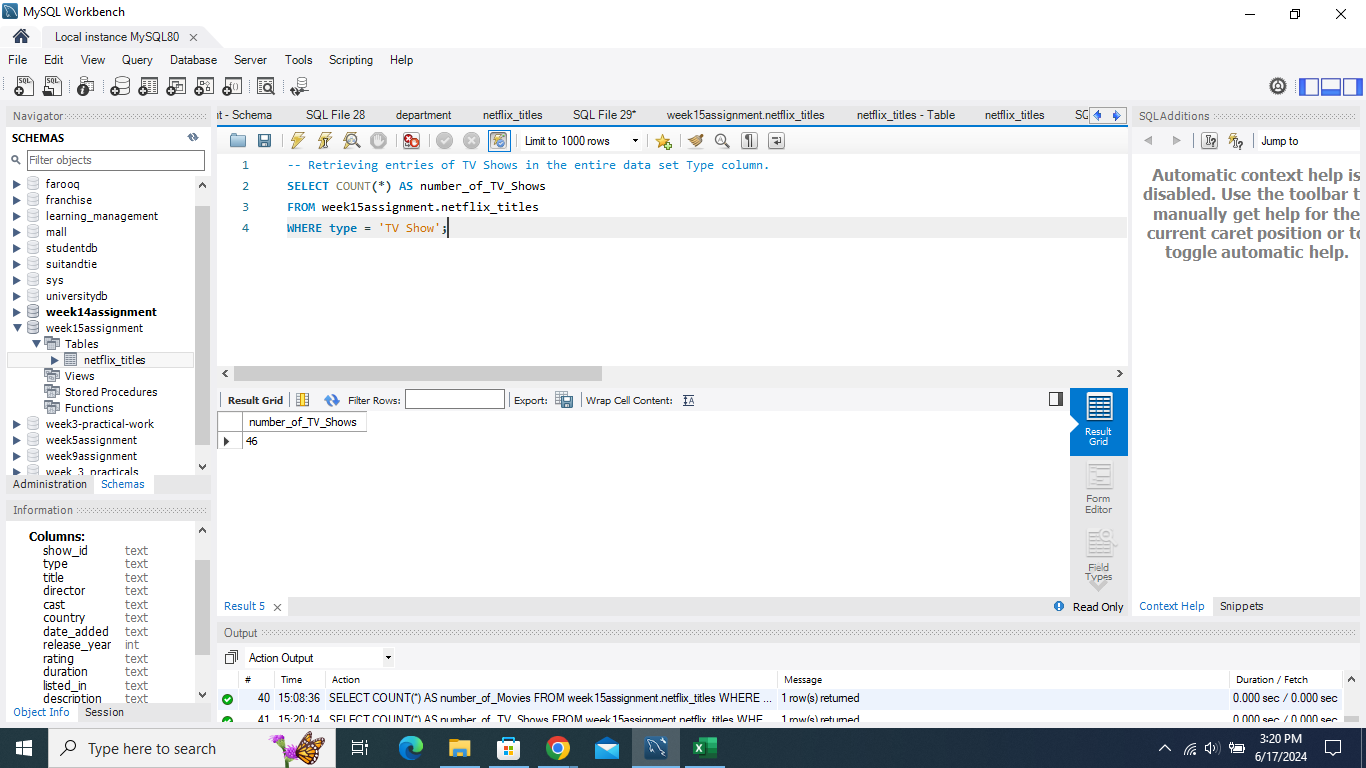
Syntax

*-- Retrieving entries of TV Shows in the entire data set Type column.*

*SELECT COUNT(\*) AS number\_of\_TV\_Shows*

*FROM week15assignment.netflix\_titles*

*WHERE type = 'TV Show’;*



*Figure 3: Showing TV Show entries in the dataset*

* Retrieve Movies and TV Shows directed by Toshiya Shinohora

Syntax

*--Retrieving entries of Movies and TV shows directed by Toshiya Shinohara.*

*SELECT type, title, COUNT(\*) AS Movies\_directed\_by\_Toshiya\_Shinohara*

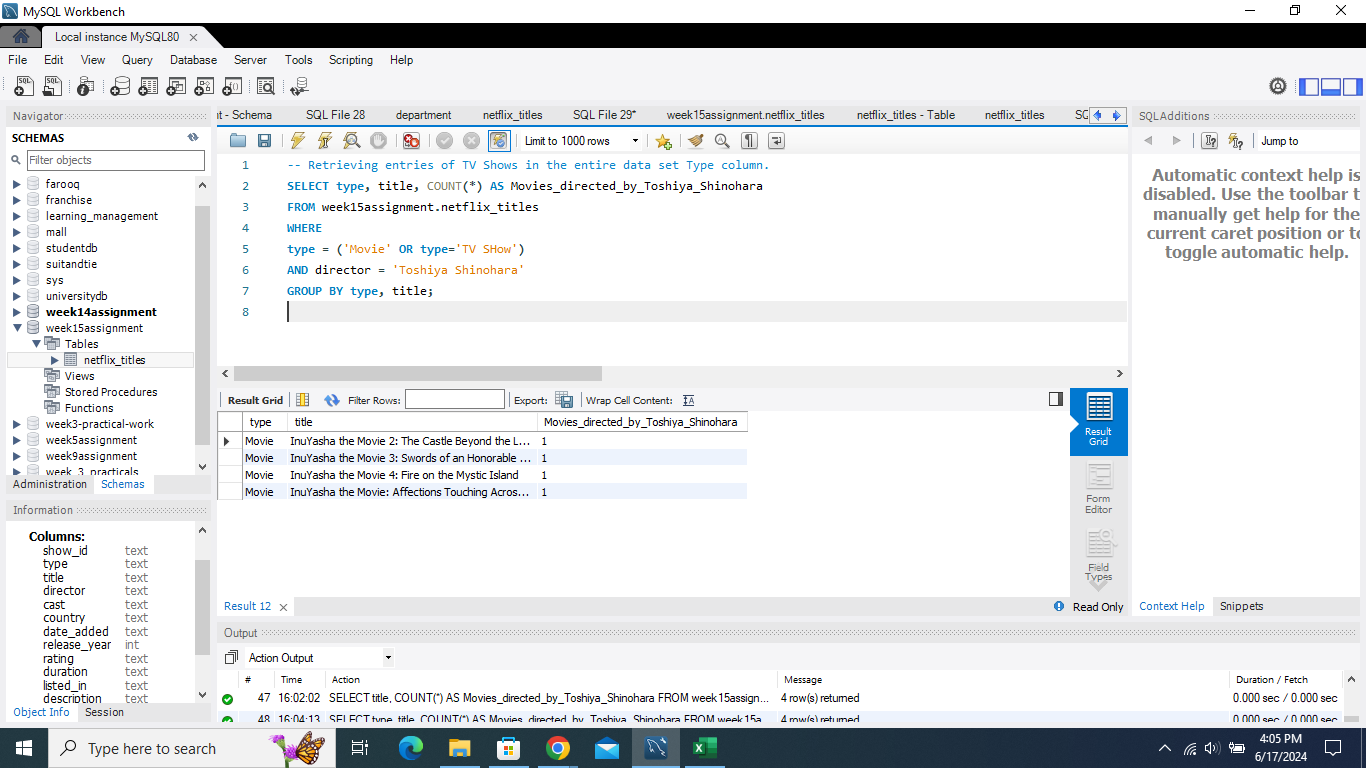
*FROM week15assignment.netflix\_titles*

*WHERE*

*type = ('Movie' OR type='TV Show')*

*AND director = 'Toshiya Shinohara'*

*GROUP BY type, title;*



*Figure 3: Entries of Movies and TV shows directed by Toshiya Shinohara.*

* Show data that list Movies and TV Shows from Japan

Syntax

*-- listing of movies and TV shows,their titles from Japan*

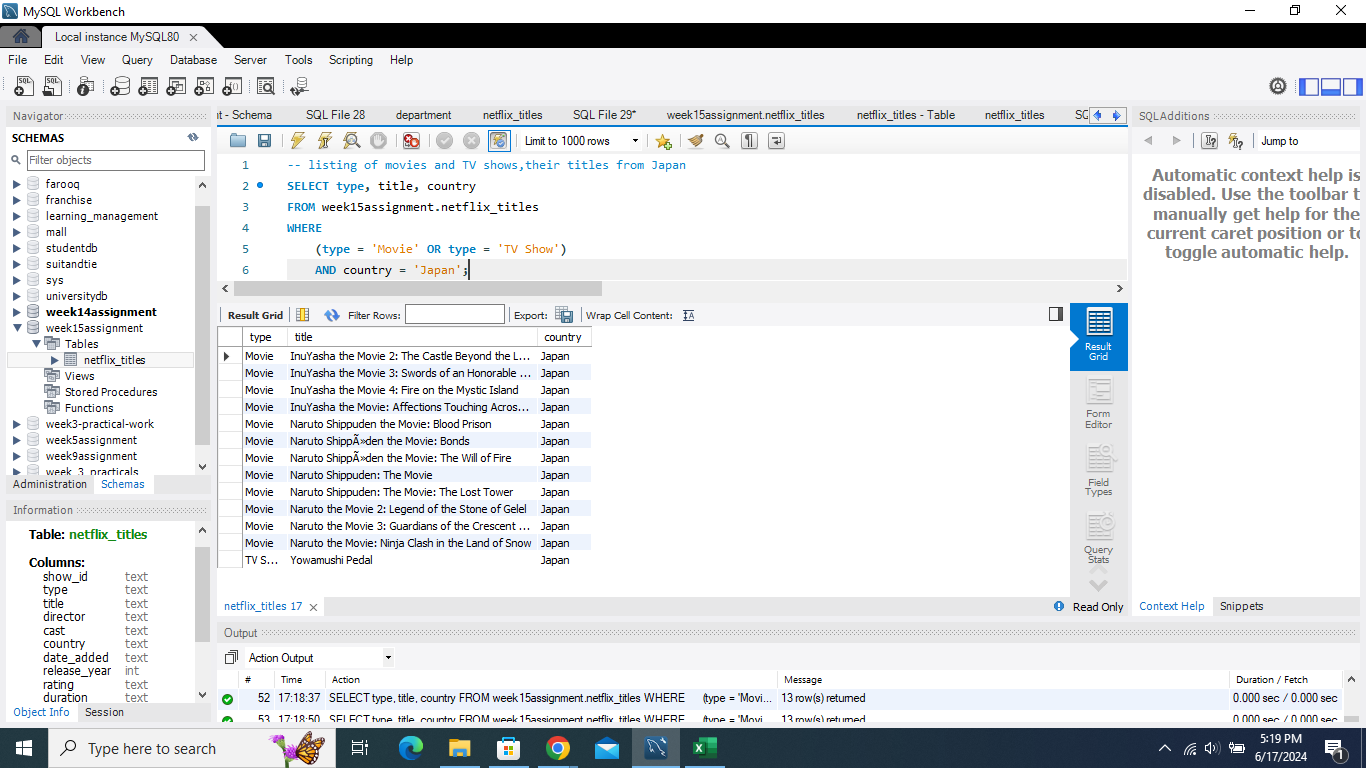
*SELECT type, title, country*

*FROM week15assignment.netflix\_titles*

*WHERE*

*(type = 'Movie' OR type = 'TV Show')*

*AND country = 'Japan';*



*Figure4: Show data that list Movies and TV Shows from Japan*

* List all TV Shows and Movies listed on 24 sept 21

Syntax

*/\* Listing all TV Shows and Movies listed on 24 sept 21,by title,type,show id,release year\*/*

*SELECT show\_id,type, title, director, country, date\_added, release\_year*

*FROM week15assignment.netflix\_titles*

*WHERE*

*(type = 'Movie' OR type = 'TV Show')*

*AND date\_added = '24-Sep-21'*

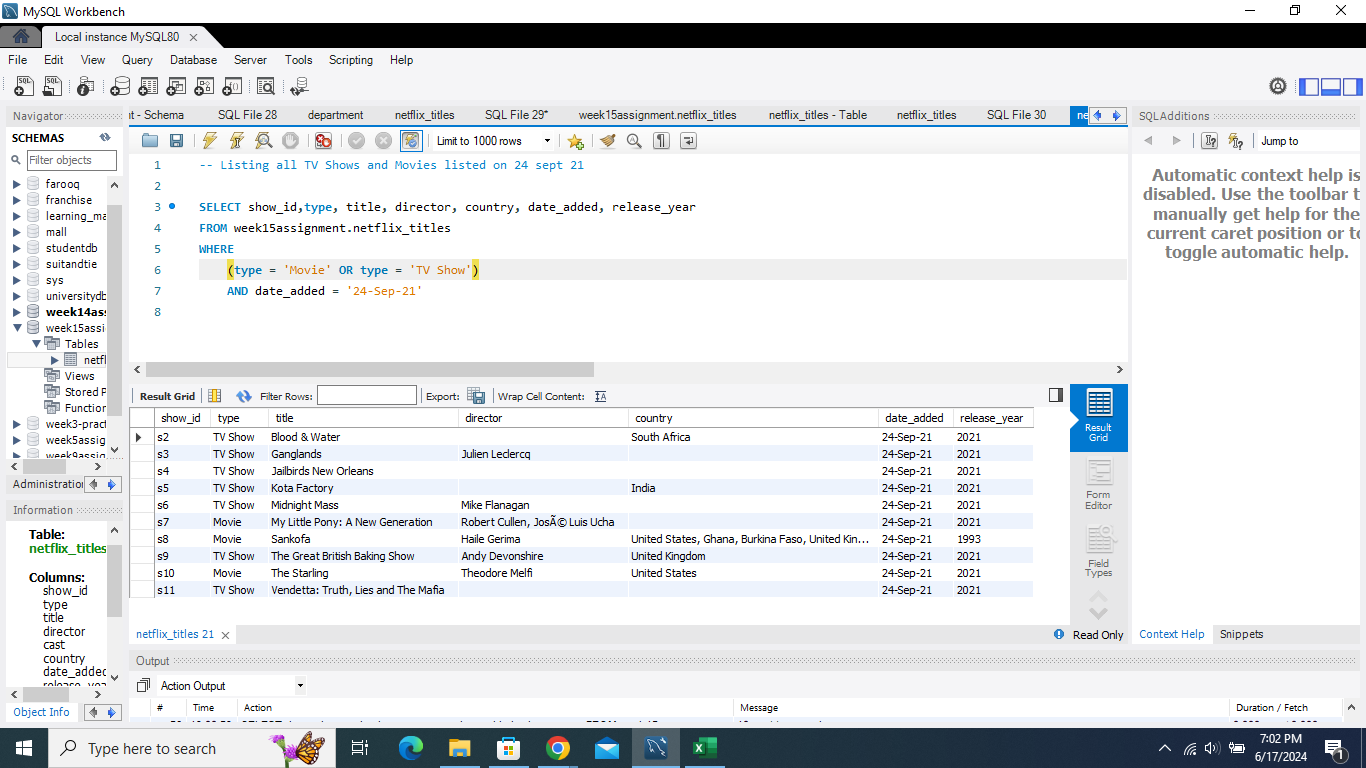


Figure 5: *Listing all TV Shows and Movies listed on 24 sept 21,by title,type,show id,release year*

1. Find 2 cool facts hidden within the data (e.g., most popular interests).

* The oldest movie and when it was released and recorded in the dataset

Syntax

*-- Finding the most oldest Movie in the dataset*

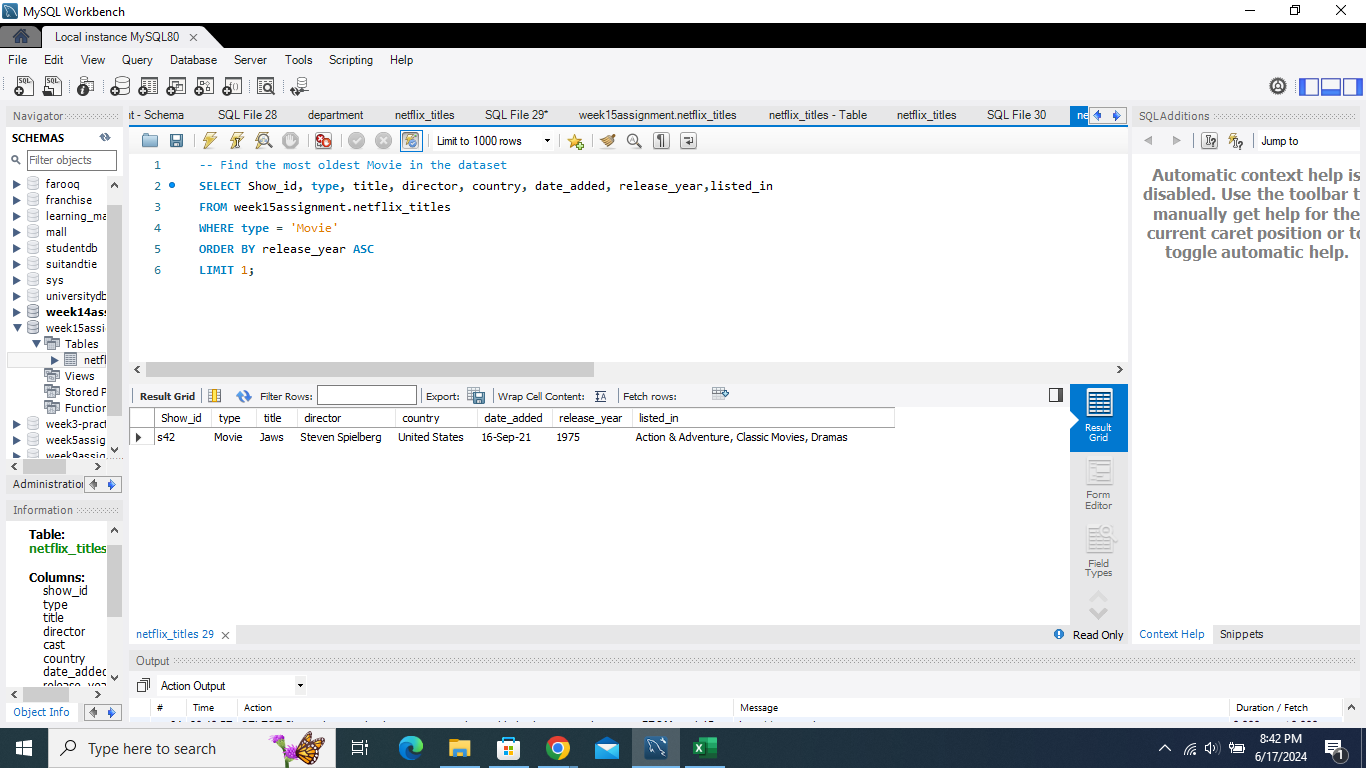
*SELECT Show\_id, type, title, director, country, date\_added, release\_year,listed\_in*

*FROM week15assignment.netflix\_titles*

*WHERE type = 'Movie'*

*ORDER BY release\_year ASC*

*LIMIT 1;*



*Figure 6: Table analysis of the oldest Movie in the dataset.*

* Finding the Oldest TV Show in the dataset

Syntax

*-- Find the most oldest TV show in the dataset*

*SELECT Show\_id, type, title, director, country, date\_added, release\_year,listed\_in*

*FROM week15assignment.netflix\_titles*

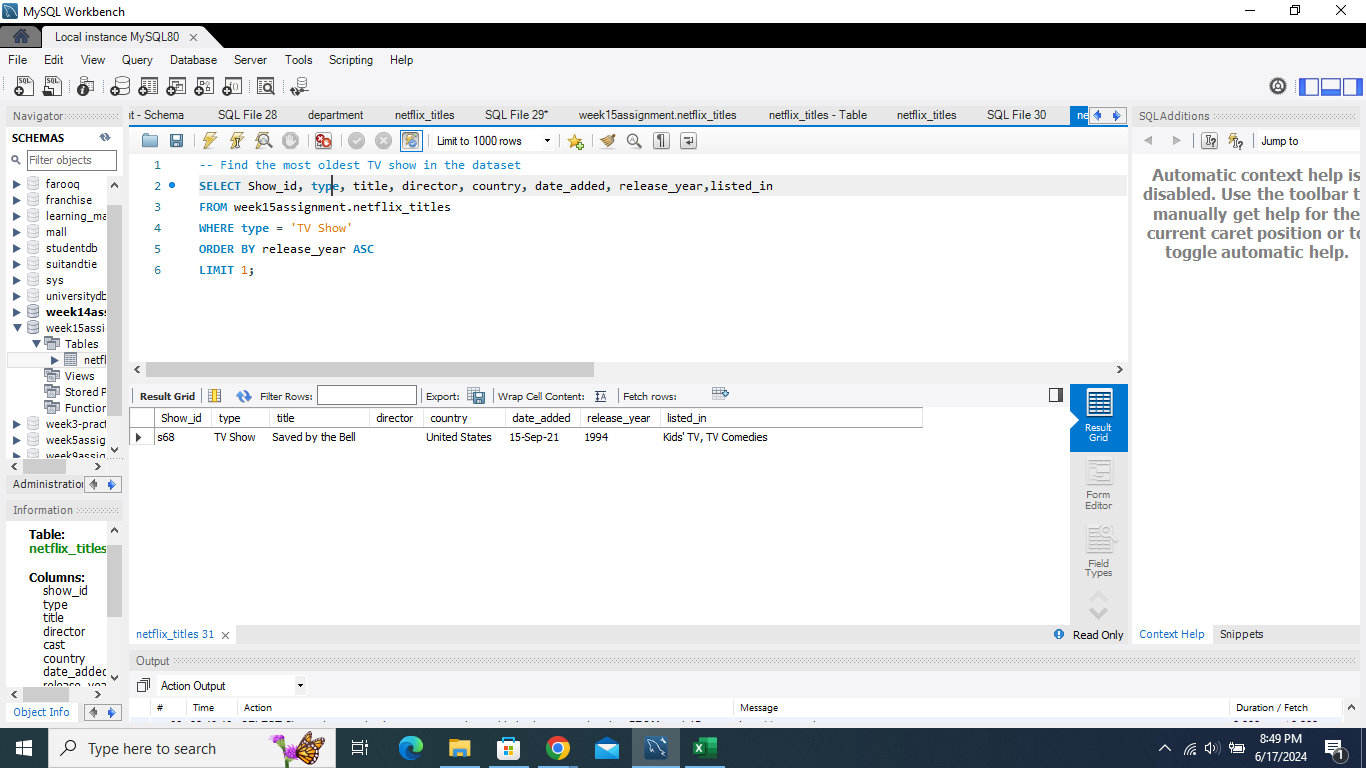
*WHERE type = 'TV Show'*

*ORDER BY release\_year ASC*

*LIMIT 1;*

*Output*

*'s68', 'TV Show', 'Saved by the Bell', '', 'United States', '15-Sep-21', '1994', 'Kids\' TV, TV Comedies'*



*Figure7: Table showing the Oldest TV Show in the dataset*

* Using function COUNT to define the selected rows in the dataset

Syntax

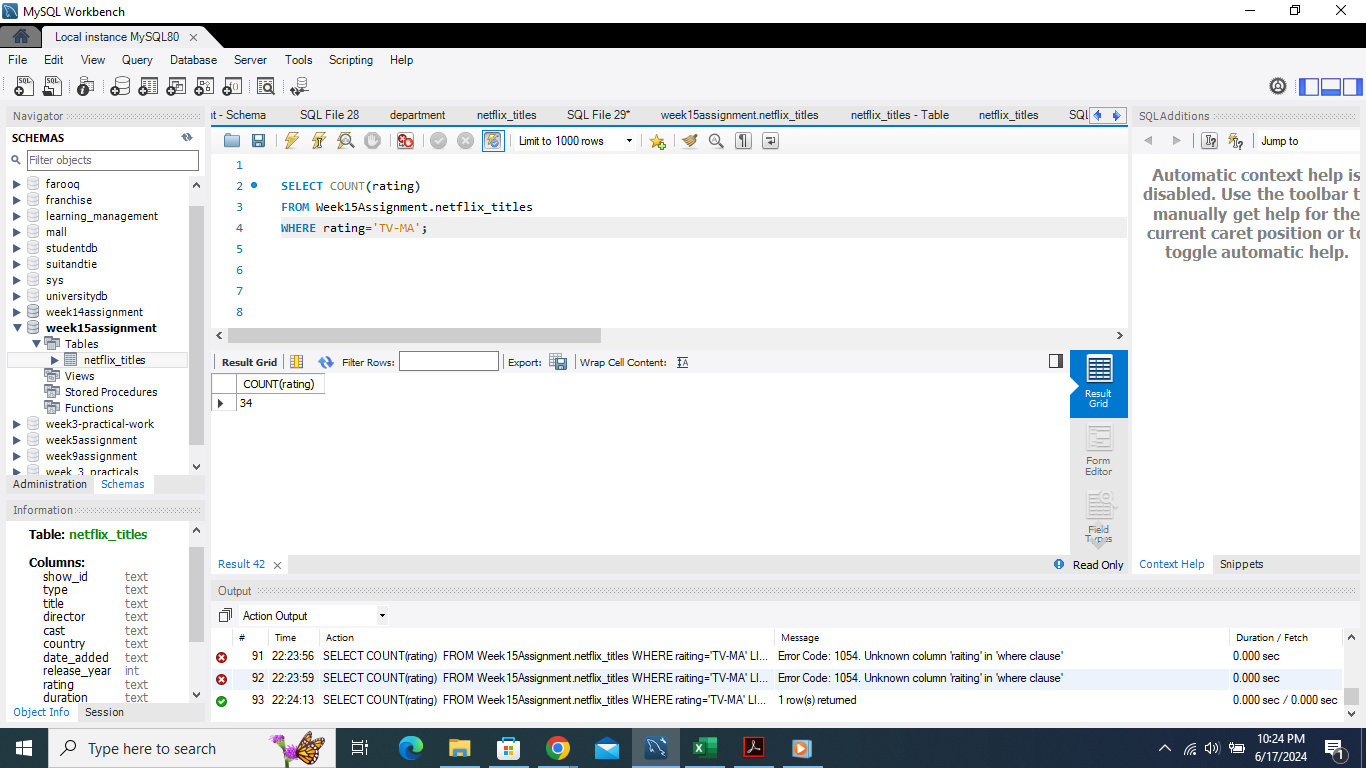
*SELECT COUNT(rating)*

*FROM Week15Assignment.netflix\_titles*

*WHERE rating='TV-MA';*

This is to count the number of entries for TV-MA under rating

Output: '34'



*Figure 8: Table showing the Counting the TV-MA entries under rating column*

* Using COUNT to select the entries for South Africa under country column

Syntax:

*SELECT COUNT(country)*

*FROM Week15Assignment.netflix\_titles*

*WHERE country='South Africa';*

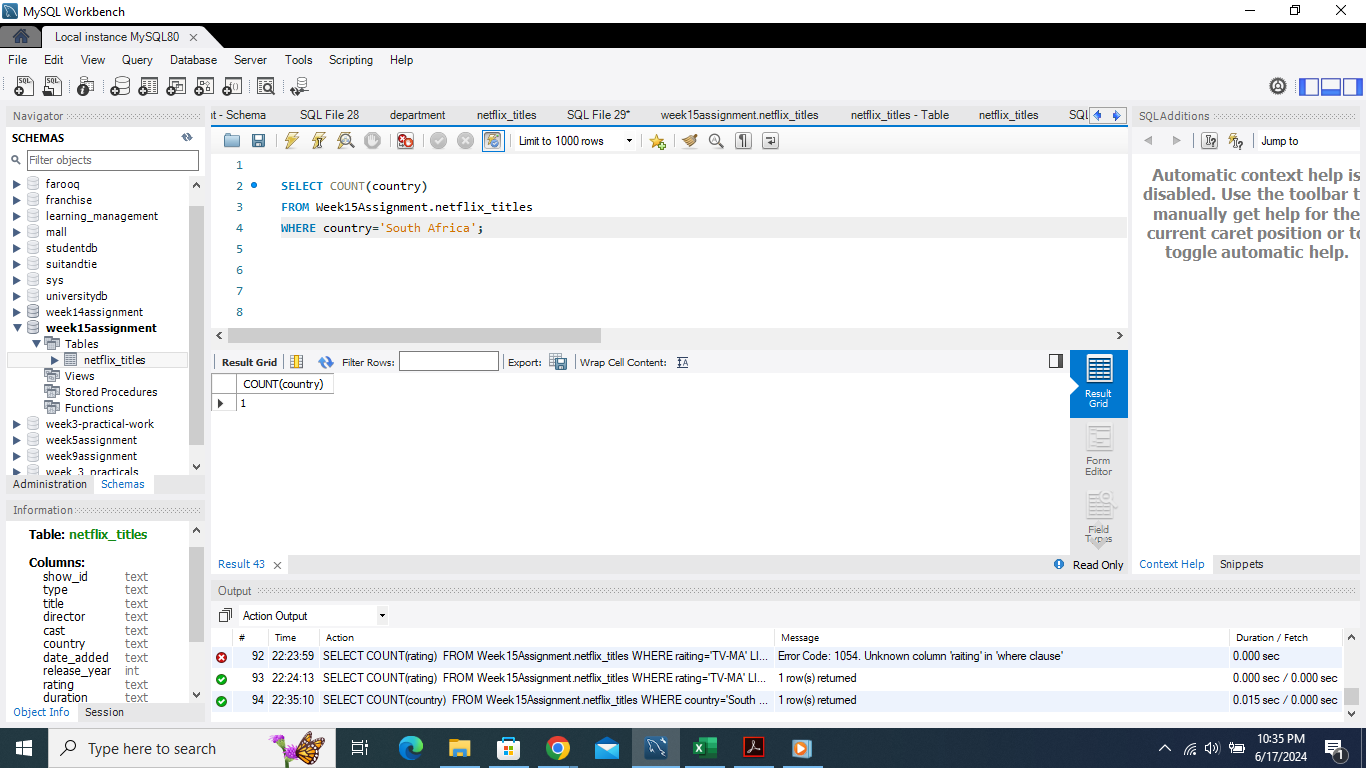
q

Figure 9: Table showing the number of entries of South Africa under Country

1. Ask Away :

Formulate 2 questions about the data (e.g., what are popular shows in different countries?).

1. Which year has the most Movies listed with the rating of TV-MA?
2. How many entries were made in the year 2021 alone ?

3. Write basic SQL queries (WHERE, ORDER BY) to find answers.

(i)Syntax:

*SELECT release\_year, COUNT(\*) AS movie\_count*

*FROM Week15Assignment.netflix\_titles*

*WHERE rating = 'TV-MA' AND type = 'Movie'*

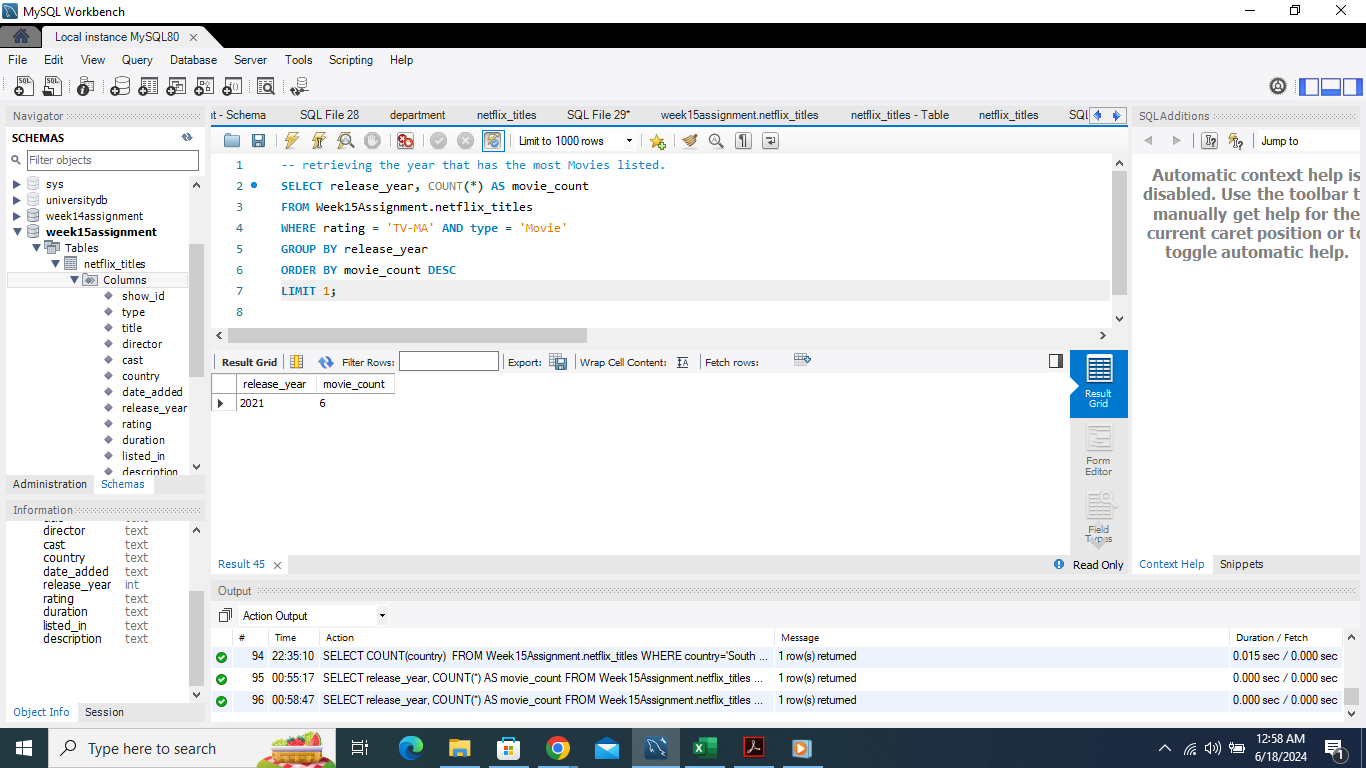
*GROUP BY release\_year*

*ORDER BY movie\_count DESC*

*LIMIT 1;*

*Output*

'2021', '6'



*Figure 10: Table showing the year that has the most Movies listed with the rating of TV-MA*

*This is now showing the title,rating,director and year of the Movies with the rating TV-MA*

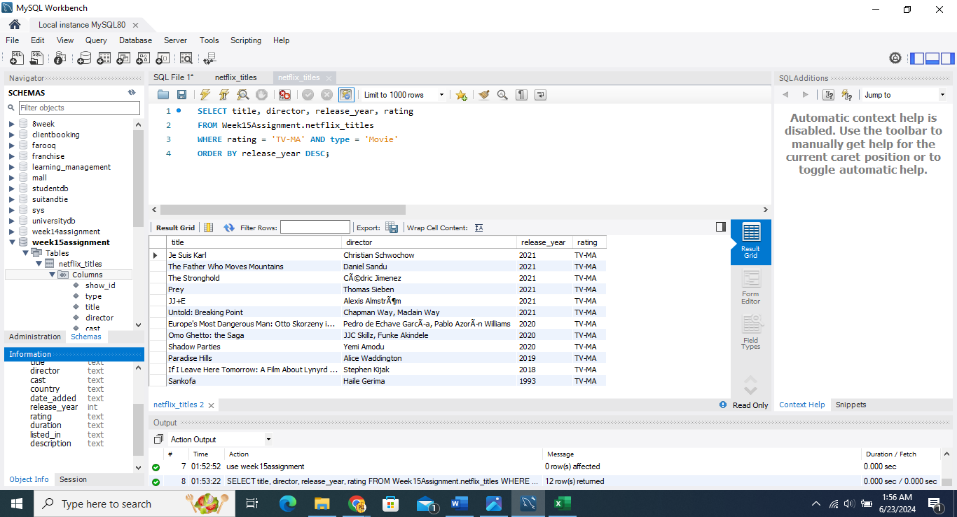
*Syntax*

*SELECT title, director, release\_year, rating*

*FROM Week15Assignment.netflix\_titles*

*WHERE rating = 'TV-MA' AND type = 'Movie'*

*ORDER BY release\_year DESC;*



*This table summarizes the movies with rating TV-MA*

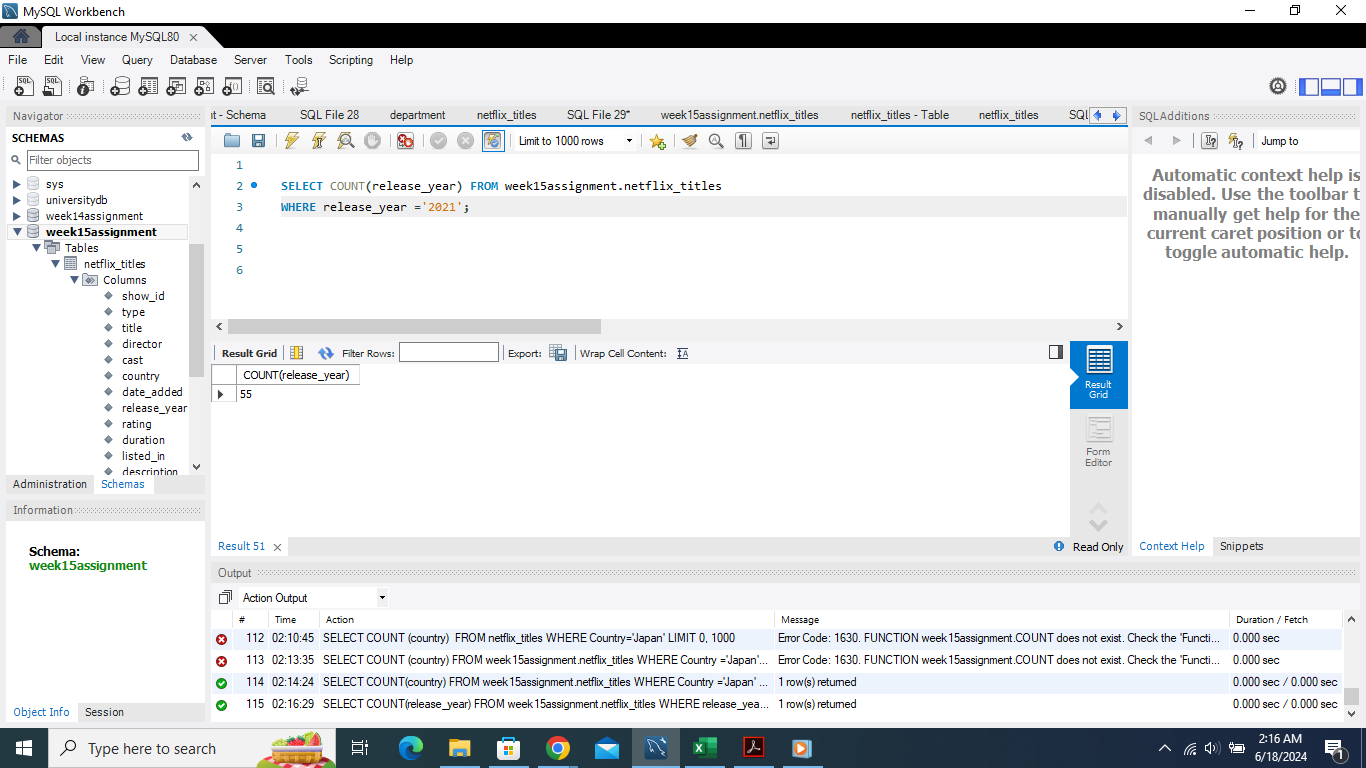
(ii)Syntax:

*-- Counting the number of entries entered in the year 2021*

*SELECT COUNT(release\_year) FROM week15assignment.netflix\_titles*

*WHERE release\_year ='2021';*

Output: '55'



*Figure 11: Table showing the number of entries entered in the year 2021*