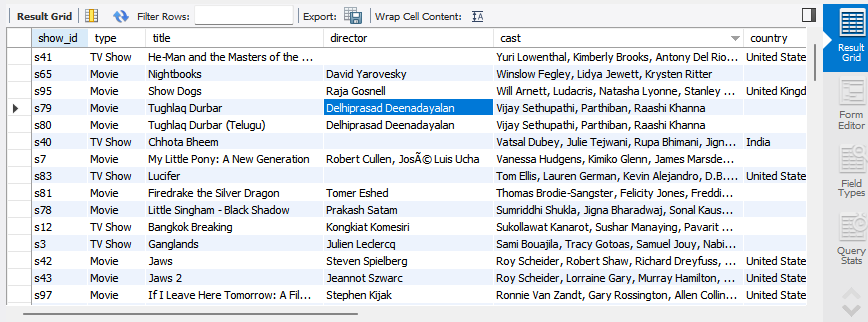
## **SQL Aggregate Functions- WEEK 15**

**1. Data Dive (10 pts): NETFLIX\_SHOWS/TITLES**

-- selecting all rows and columns from the database

SELECT \* FROM netflix\_titles.netflix\_titles;



Difficulties:- The data is very complex and it might generate errors in the process of analyzing if proper commands are not used.

-some columns eg ‘the description column” makes the data complicated in terms of

Executing various database queries and it might bring about complication when

Retrieving a specific row or column from the database.

- some information is missing from the database, ie, from the “Director column”

And this can lead to difficulty when analyzing the whole data.

Interesting thing:- The data is well organized sequentially in terms of column and rows making it

Easier to retrieve a specific set of data .

- The similarity of some column data makes it easier to organize and retrieve

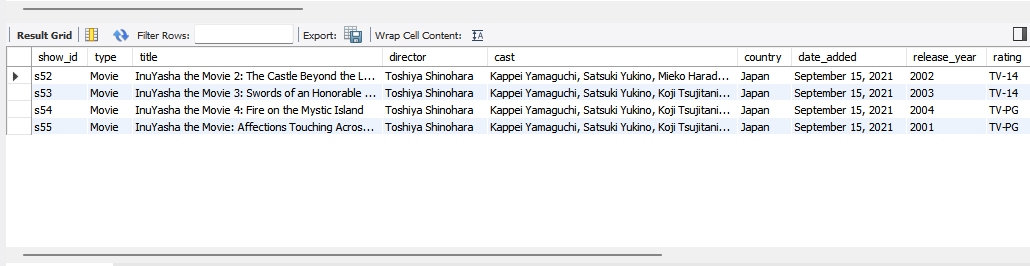
Various sets of data.

**2. Data Fun (20 pts): Use simple SQL queries to play with the data.**

-- selecting all rows and columns from the database

SELECT \* FROM netflix\_titles.netflix\_titles

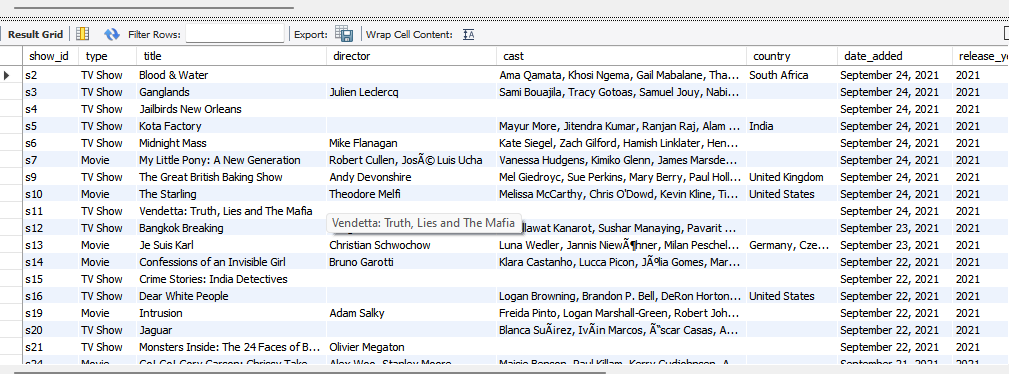
where director = "Toshiya Shinohara";



-- selecting all rows and columns from the database

SELECT \* FROM netflix\_titles.netflix\_titles

where release\_year = "2021";



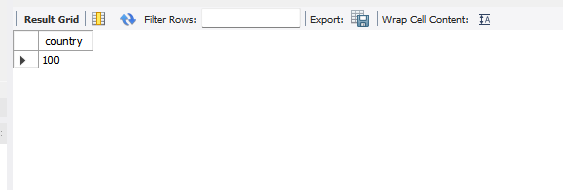
* Find 2 cool facts hidden within the data (e.g., most popular interests).
* Most of the production was done in the year 2021
* All the production was made on the month of september
* Use basic SQL queries like (COUNT, AVG, and SUM) to understand more about the data you have.

-- COUNTING THE TOTAL NUMBER OF COUNTRIES FROM THE PRODUCTION

use netflix\_titles;

SELECT COUNT(\*) country

FROM netflix\_titles;

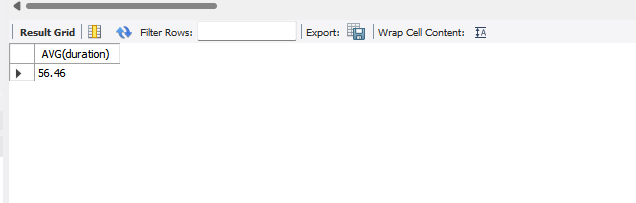


-- CALCULATING THE AVERAGE DURATION

use netflix\_titles;

SELECT AVG(duration)

FROM netflix\_titles;



**3. Ask Away (30 pts):**

* **Formulate 2 questions about the data (e.g., what are popular shows in different countries?).**
* **Write basic SQL queries (WHERE, ORDER BY) to find answers.**

-- This SQL query selects the rating column from the netflix\_titles

-- database, counts the number of titles associated with each unique

-- rating category, and presents the results in descending

-- order of title counts.

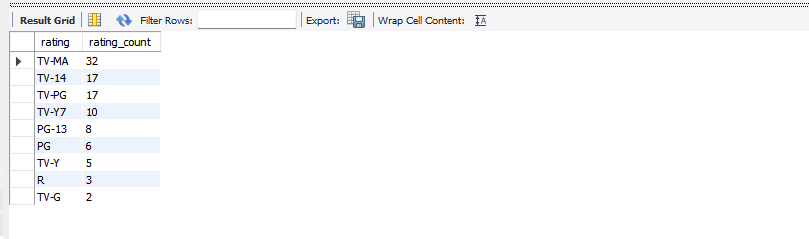
use netflix\_titles;

SELECT rating, COUNT(\*) AS rating\_count

FROM netflix\_titles

GROUP BY rating

ORDER BY rating\_count DESC;



* Share what you learned from the answers.

TV-MA - had the highest rating

TV-G - had the lowest rating

-- This query allows us to identify popular TV shows

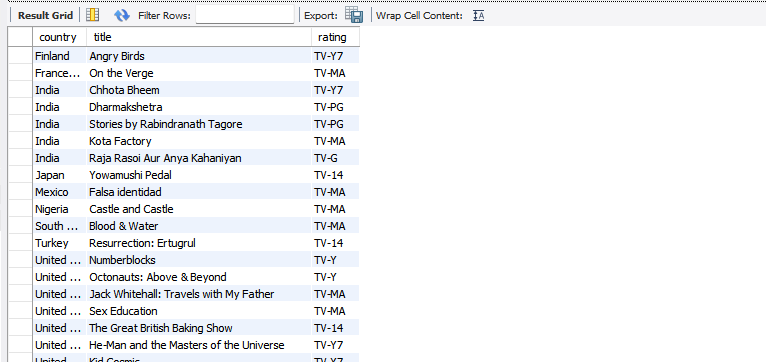
-- in different countries based on viewer ratings.`

SELECT country, title, rating

FROM netflix\_titles

WHERE type = 'TV Show'

ORDER BY country, rating DESC



* Share what you learned from the answers.

-The query reveals which TV shows are highly rated and thus popular in various countries.