* **Name: Tanvi Nitin Paygude**
* **Elevate Labs Internship Task 3**
* **Dataset Used: netflix\_titles.csv (File Attached in the Repository)**

**Title: SQL Data Analysis Task –Netflix Dataset**

1. **Objective**

**The objective of this task is to analyse customer data using SQL by performing various queries such as selections, aggregations, grouping, subqueries, and creating views. The goal is to understand customer behaviour based on demographics and spending patterns.**

1. **Tools Used**

**Platform: MySQL**

**Dataset: Mall\_Customers.csv**

**Language: SQL**

1. Dataset Structure

CREATE TABLE mytable(

show\_id VARCHAR(5) NOT NULL PRIMARY KEY,

title VARCHAR(104) NOT NULL,

country VARCHAR(123),

date\_added DATE,

release\_year INTEGER NOT NULL,

duration VARCHAR(10)

);

INSERT INTO mytable(show\_id, title, country, date\_added, release\_year, duration)

VALUES

('s41', 'Interstellar', 'United States', '2014-11-07', 2014, '169 min'),

('s42', 'Black Mirror', 'United Kingdom', '2011-12-04', 2011, '5 Seasons'),

('s43', 'Joker', 'United States', '2019-10-04', 2019, '122 min'),

('s44', 'Friends', 'United States', '1994-09-22', 1994, '10 Seasons'),

('s45', 'Gladiator', 'United States', '2000-05-05', 2000, '155 min'),

('s46', 'Sherlock', 'United Kingdom', '2010-07-25', 2010, '4 Seasons'),

('s47', 'The Social Network', 'United States', '2010-10-01', 2010, '120 min'),

('s48', 'Peaky Blinders', 'United Kingdom', '2013-09-12', 2013, '6 Seasons'),

('s49', 'Parasite', 'South Korea', '2019-05-30', 2019, '132 min'),

('s50', 'Money Heist', 'Spain', '2017-05-02', 2017, '5 Seasons'),

('s51', 'The Irishman', 'United States', '2019-11-01', 2019, '209 min'),

('s52', 'Ozark', 'United States', '2017-07-21', 2017, '4 Seasons'),

('s53', '1917', 'United Kingdom', '2019-12-25', 2019, '119 min'),

('s54', 'The Queens Gambit', 'United States', '2020-10-23', 2020, '1 Season'),

('s55', 'Tenet', 'United States', '2020-08-26', 2020, '150 min'),

('s56', 'The Witcher', 'United States', '2019-12-20', 2019, '3 Seasons'),

('s57', 'Whiplash', 'United States', '2014-10-10', 2014, '107 min'),

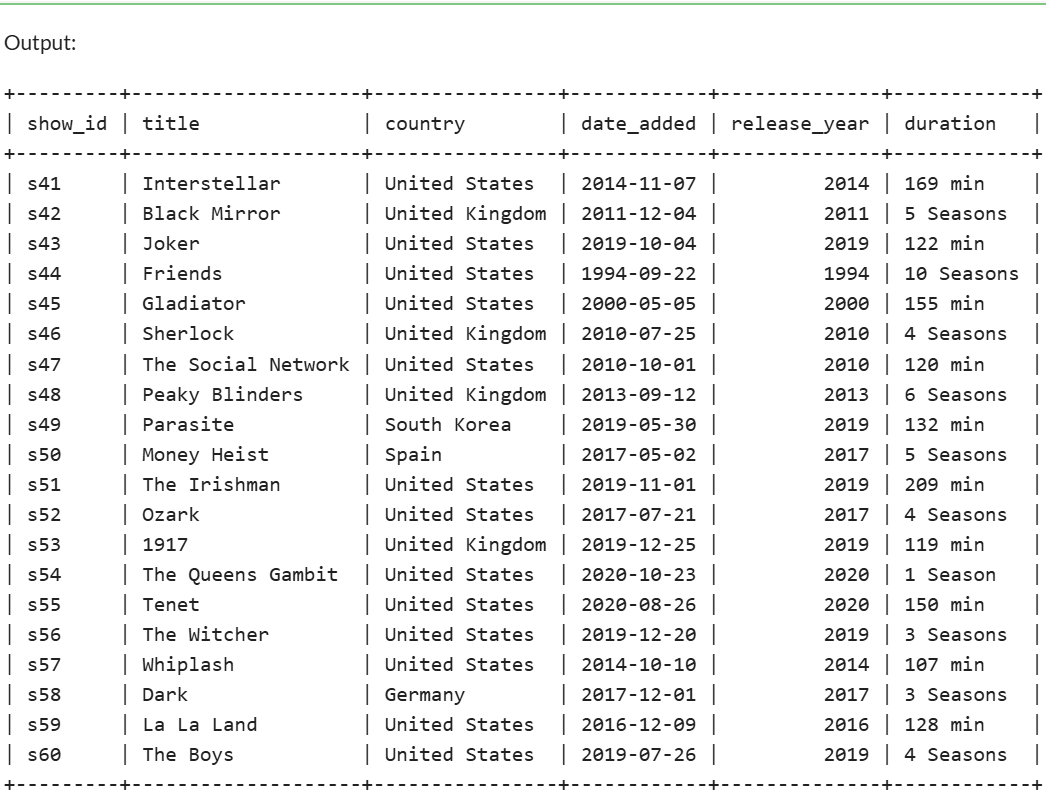
('s58', 'Dark', 'Germany', '2017-12-01', 2017, '3 Seasons'),

('s59', 'La La Land', 'United States', '2016-12-09', 2016, '128 min'),

('s60', 'The Boys', 'United States', '2019-07-26', 2019, '4 Seasons');

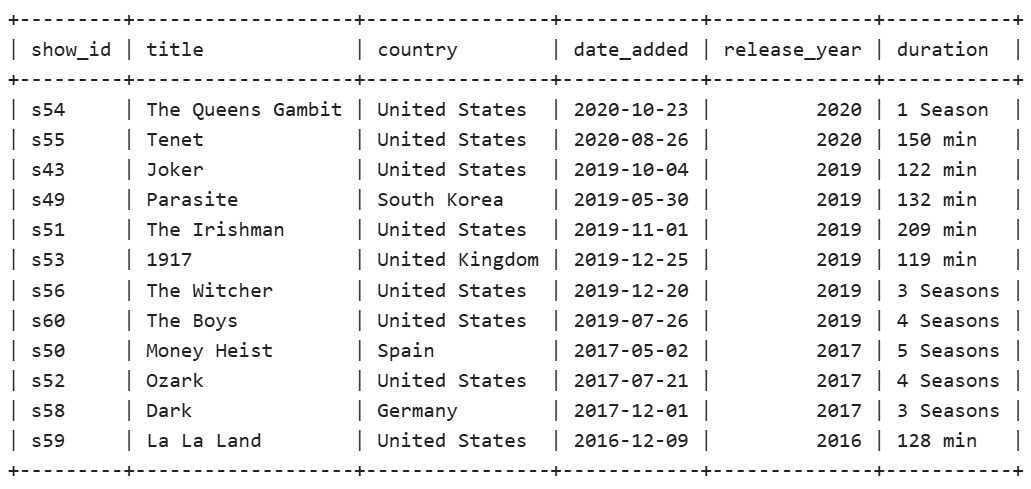
select\*from mytable

Output:



* 1. Basic Queries

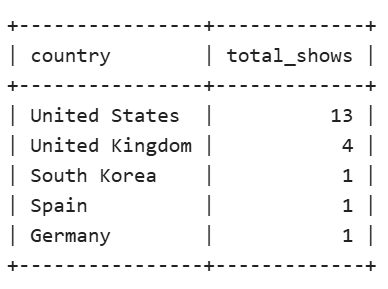
SELECT \* FROM mytable WHERE release\_year > 2015 ORDER BY release\_year DESC;

Output  
  


* 1. Group By with Aggregates

SELECT country, COUNT(\*) AS total\_shows FROM mytable GROUP BY country ORDER BY total\_shows DESC;

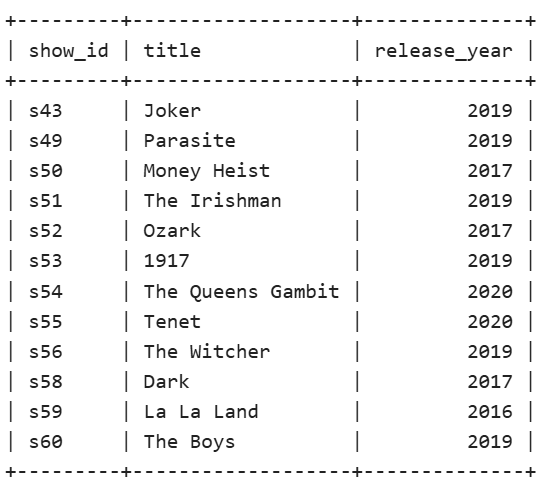
Output



* 1. Subqueries

SELECT show\_id, title, release\_year FROM mytable WHERE release\_year > ( SELECT AVG(release\_year) FROM mytable);

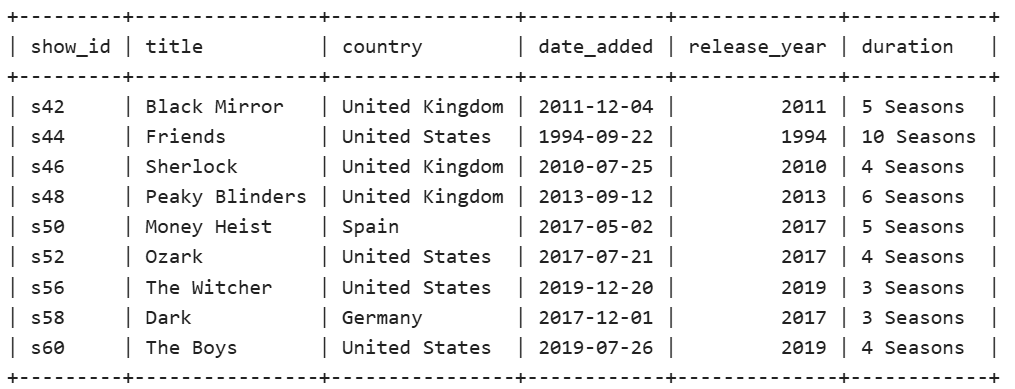
Output



4. Using the Views

VIEW all\_tv\_shows AS SELECT show\_id, title, country, date\_added, release\_year, duration FROM mytable WHERE duration LIKE '%Seasons';

select \* from all\_tv\_shows



5.Creating Index

SELECT \* FROM mytable

WHERE release\_year = 2019 AND country = 'United States';

