**DATABASE ASSIGNMENT WEEK 15**

**SQL query to create a database for storing the imported data:**

-- Creating a database that will contain the data

CREATE DATABASE netflix\_shows;

USE netflix\_shows;

-- Viewing the entire dataset

SELECT \* FROM netflix\_titles;

**SQL query to find the total number of TV shows in the first 100 records:**

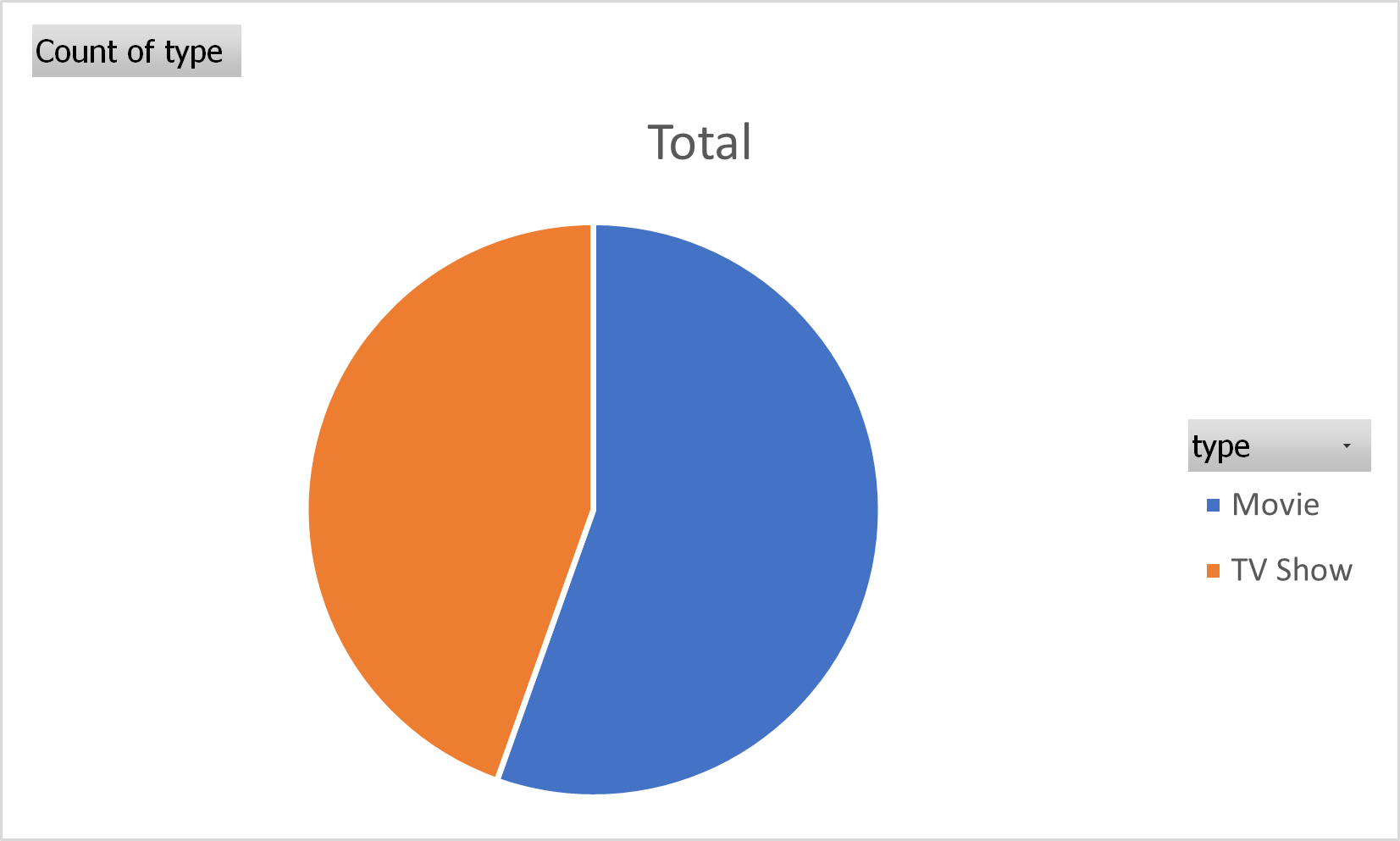
-- Count the total number of tv shows

SELECT COUNT(\*) AS total\_tv\_shows

FROM netflix\_titles

WHERE type = 'TV Show';

The total number of tv shows was 45 out of 100.



**SQL query to determine the average duration of a movie for the first 100 records:**

-- Query to find the average duration of movies

SELECT AVG(CAST(SUBSTRING\_INDEX(duration, ' ', 1) AS UNSIGNED)) AS average\_movie\_duration

FROM netflix\_titles

WHERE type = 'Movie';

The average duration of a movie was 100.7273 minutes.

**Query to count the number of titles for each rating for the first 100 records:**

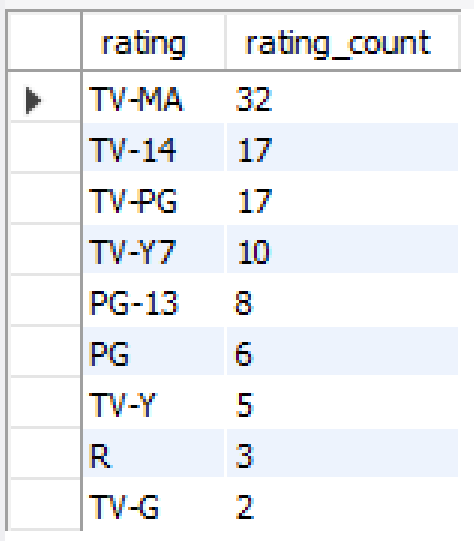
-- Query to find the distribution of content ratings

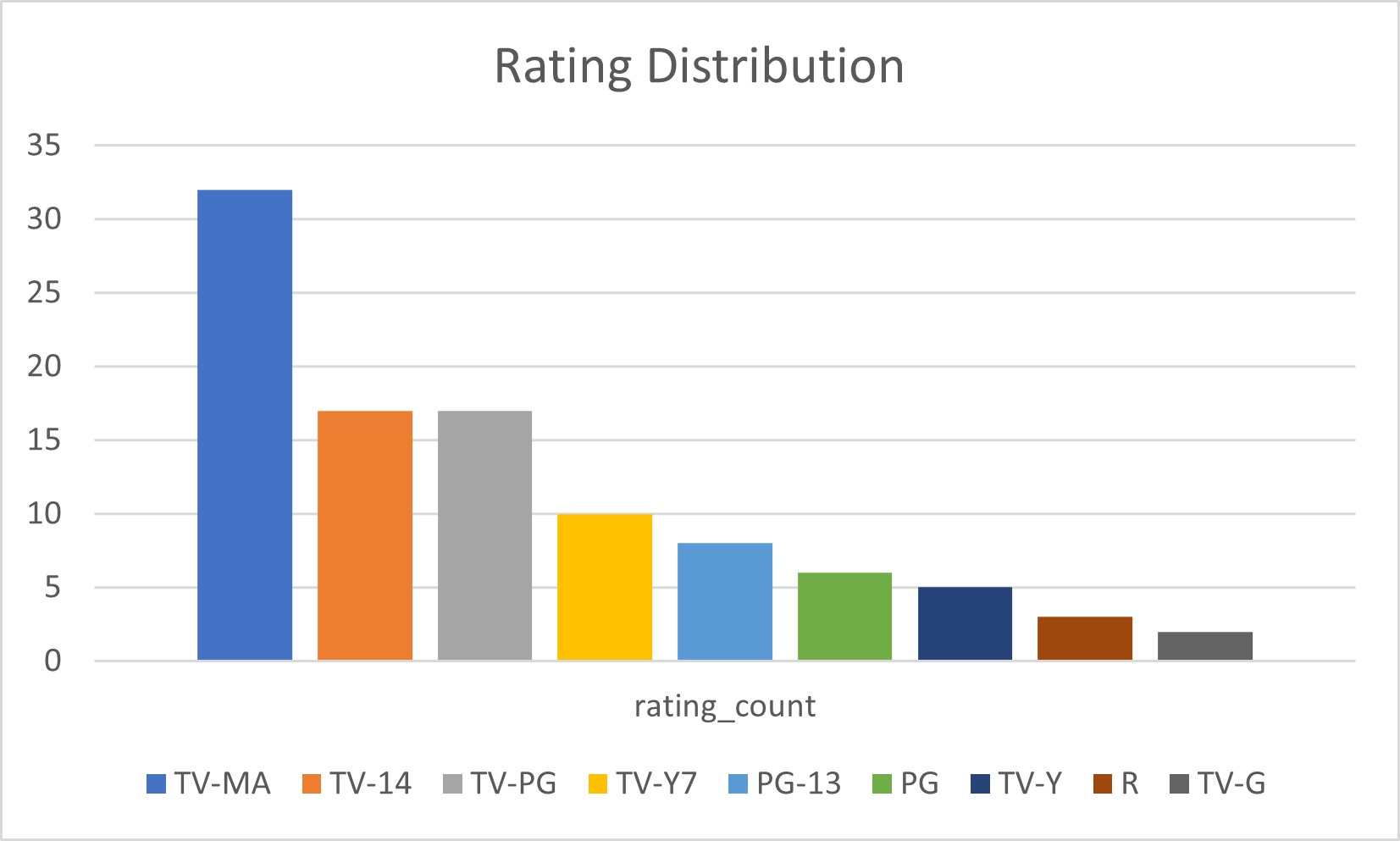
SELECT rating, COUNT(\*) AS rating\_count

FROM netflix\_titles

GROUP BY rating

ORDER BY rating\_count DESC;

The results obtained from the query are shown in the figure below:  


****

**Query to determine the most common genres from the first 100 records :**

-- Query to find the most common genres

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

FROM (

    SELECT TRIM(SUBSTRING\_INDEX(SUBSTRING\_INDEX(listed\_in, ',', numbers.n), ',', -1)) AS genre

    FROM netflix\_titles

    INNER JOIN (SELECT 1 n UNION ALL SELECT 2 UNION ALL SELECT 3 UNION ALL SELECT 4 UNION ALL SELECT 5 UNION ALL SELECT 6 UNION ALL SELECT 7 UNION ALL SELECT 8 UNION ALL SELECT 9 UNION ALL SELECT 10) numbers

    ON CHAR\_LENGTH(listed\_in) - CHAR\_LENGTH(REPLACE(listed\_in, ',', '')) >= numbers.n - 1

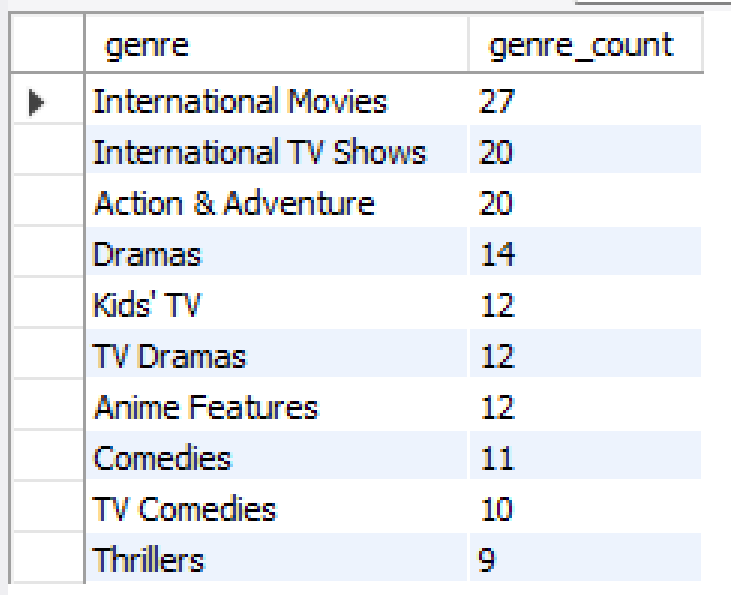
) AS genres\_list

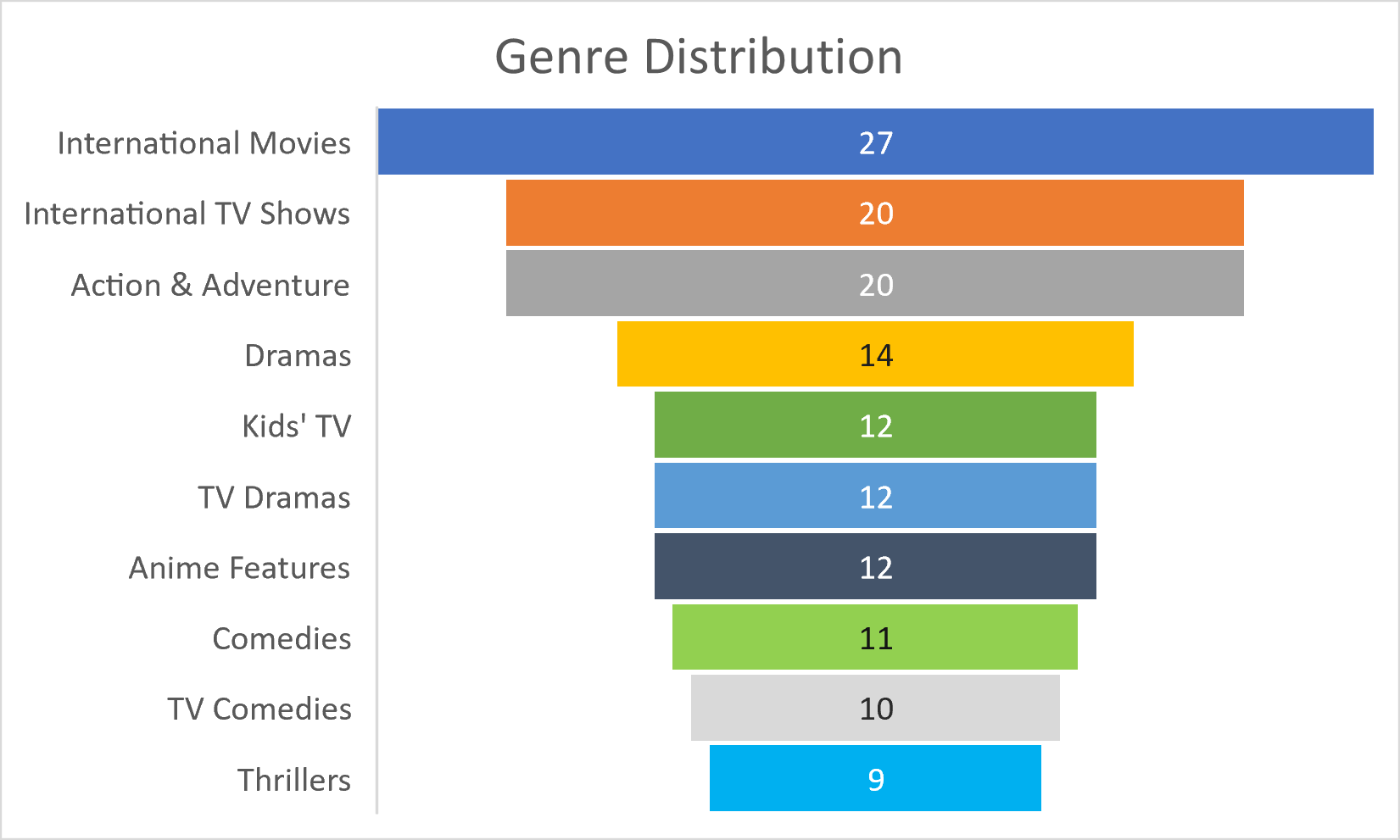
GROUP BY genre

ORDER BY genre\_count DESC

LIMIT 10;

Results obtained:





**Cool facts about the Netflix Titles dataset:**

1. The average duration of a movie, based on the first 100 records, is 100 minutes (1h 40 mins).
2. The content rating with the highest number of titles is TV-MA (TV Mature Audience), with 32 titles out of 100, while TV-G (TV General Audience) has the lowest number of titles i.e 2 out of the first 100.
3. The most common genre is ‘International Movies’ followed by ‘International TV Shows’.

**Question 1: What are the top 5 countries producing the most TV Shows?**

SQL query:

-- Query to find the top 5 countries producing the most TV shows

SELECT country, COUNT(\*) AS tv\_show\_count

FROM netflix\_titles

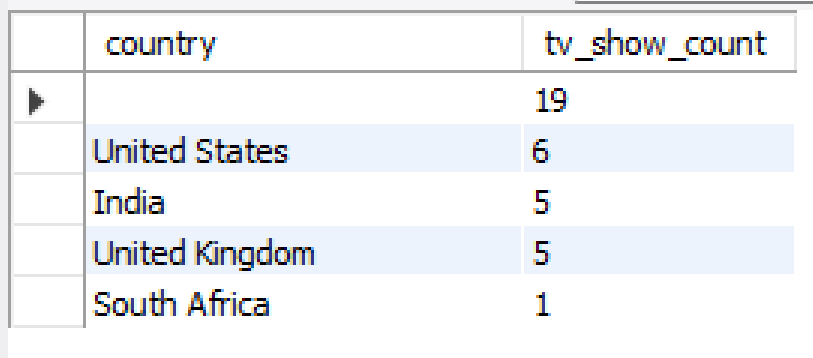
WHERE type = 'TV Show'

GROUP BY country

ORDER BY tv\_show\_count DESC

LIMIT 5;

Results:



The country that produces the highest number of TV shows based on the first 100 records is the United States.

**Question 2: What are the top 5 most recent TV Shows added to Netflix?**

SQL query:

-- Query to find the top 5 most recent TV Shows added to Netflix

SELECT title, release\_year

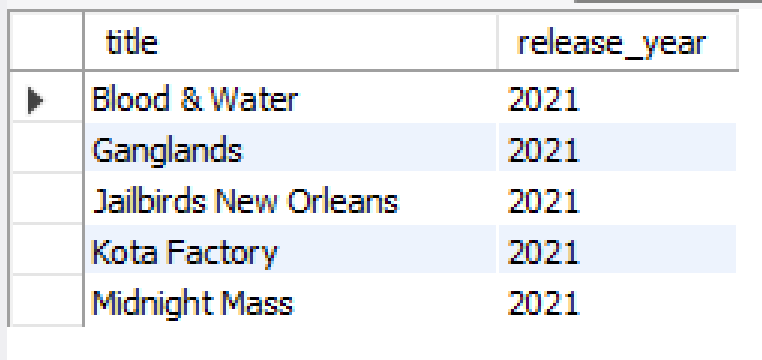
FROM netflix\_titles

WHERE type = 'TV Show'

ORDER BY release\_year DESC

LIMIT 5;

Result:



This query returns the newest additions to Netflix programs, providing a snapshot of the most recent TV Shows available to viewers. This can be useful for keeping up-to-date with new releases.