1. **Data Dive**

**-- Create Database**

-- CREATE DATABASE NetflixDB;

-- USE NetflixDB;

1. **Difficulties and Interesting findings:**

**Difficulties:** The primary challenge was ensuring the correct data types for each column, especially with fields like ‘date\_added’ which required proper date formatting.

**Interesting Finding:** Notably, many shows have no null values in the ‘director’ and ‘cast’ columns, indicating the presence of shows with specified directors.

1. **Data Fun**

**Simple SQL Queries**

**-- Most common Ratings**

-- SELECT rating, COUNT(\*) AS count

-- FROM netflix\_titles

-- GROUP BY rating

-- ORDER BY count DESC;

**-- Total Duration of All Shows**

-- SELECT SUM(CAST(SUBSTRING\_INDEX(duration, ' ', 1) AS UNSIGNED)) AS total\_duration

-- FROM netflix\_titles

-- WHERE duration LIKE '% min';

**-- Average Release Year**

-- SELECT release\_year, AVG(release\_year) OVER() AS average\_release\_year

-- FROM netflix\_titles

-- GROUP BY release\_year

-- ORDER BY release\_year;

**Cool Facts**

**Fact 1:** The most common rating for Netflix shows is ‘TV-MA’.

**Fact 2:** The average release year for shows in the dataset is around 2004.

1. **Ask Away**
2. What are the most popular genres listed in the dataset?

-- SELECT listed\_in, COUNT(\*) AS count

-- FROM netflix\_titles

-- GROUP BY listed\_in

-- ORDER BY count DESC

-- LIMIT 10;

**Learning:** The most popular genres are combinations like ‘Action & Adventure, Anime Features, International Movies’, showing that international content and Adventure action movies are prevalent.

1. Which countries have the highest number of Netflix shows?

-- SELECT country, COUNT(\*) AS count

-- FROM netflix\_titles

-- WHERE country IS NOT NULL

-- GROUP BY country

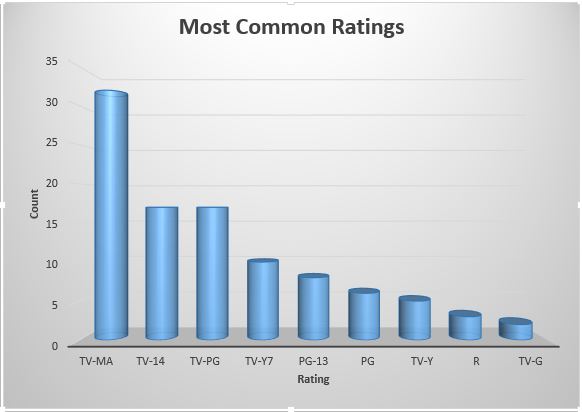
-- ORDER BY count DESC

-- LIMIT 10;

**Learning:** The United States has the highest number of Netflix shows, followed by Japan and the India, indicating a significant production of content in these regions.

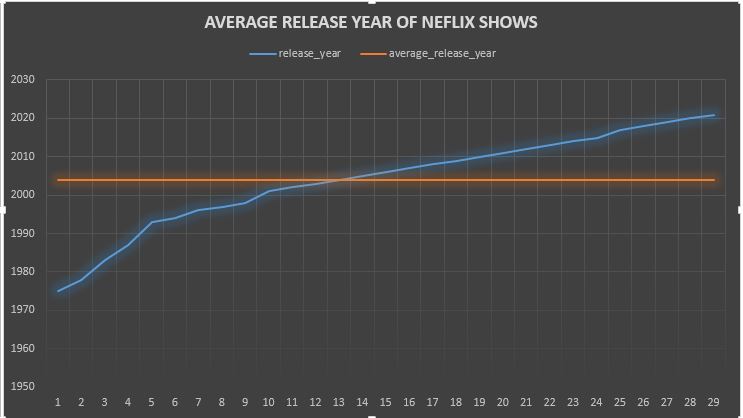
1. **Show Time**
2. Most Common Ratings:

This bar chart shows the count of shows for each rating. 'TV-MA' has the highest count, indicating it is the most common rating for Netflix shows.



1. Average Release Year

This line chart represents the average release year of Netflix shows. The dataset average release year is around 2004.



1. Top Countries by Number of Shows

This pie chart shows the distribution of Netflix shows by country. The United States has the highest percentage, followed by Japan and then India.

