**Netflix Movies and TV Shows Data Analysis using SQL**

This project involves a comprehensive analysis of Netflix's movies and TV shows data using SQL. The goal is to extract valuable insights and answer various business questions based on the dataset.

## Objectives

* Analyze the distribution of content types (movies vs TV shows).
* Identify the most common ratings for movies and TV shows.
* List and analyze content based on release years, countries, and durations.
* Explore and categorize content based on specific criteria and keywords.

## Business Problems

1. Count the number of Movies vs TV Shows

2. Find the most common rating for movies and TV shows

3. List all movies released in a specific year (e.g., 2020)

4. Find the top 5 countries with the most content on Netflix

5. Identify the longest movie

6. Find all the movies/TV shows by director 'Reginald Hudlin'!

7. Count the number of content items in each genre

8. Find each year and the numbers of content release in India on netflix.

return top 5 year with highest content release!

9. List all movies that are documentaries

10. Find how many movies actor 'Salman Khan' appeared in last 10 years!

11.Categorize the content based on the presence of the keywords 'kill' and 'violence' in

the description field. Label content containing these keywords as 'Bad' and all other

content as 'Good'. Count how many items fall into each category.

## Schema

create database netflix;

use netflix;

CREATE TABLE netflix\_data

(

show\_id VARCHAR(10) primary key not null,

type VARCHAR(10) not null,

title VARCHAR(250) not null,

director VARCHAR(550) not null,

casts VARCHAR(1050) not null,

country VARCHAR(550) not null,

release\_year INT not null,

rating VARCHAR(15) not null,

duration VARCHAR(15) not null,

listed\_in VARCHAR(250) not null,

description VARCHAR(550) not null

);