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

**Overview**

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. The following README provides a detailed account of the project's objectives, business problems, solutions, findings, and conclusions.

**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 Addressed**

**1. Count the Number of Movies vs TV Shows**

**Objective**: Determine the distribution of content types on Netflix.

**2. Find the Most Common Rating for Movies and TV Shows**

**Objective**: Identify the most frequently occurring rating for each type of content.

**3. List All Movies Released in a Specific Year (e.g., 2020)**

**Objective**: Retrieve all movies released in a specific year.

**4. Find the Top 5 Countries with the Most Content on Netflix**

**Objective**: Identify the top 5 countries with the highest number of content items.

**5. Identify the Longest Movie**

**Objective**: Find the movie with the longest duration.

**6. Find Content Added in the Last 5 Years**

**Objective**: Retrieve content added to Netflix in the last 5 years.

**7. Find All Movies/TV Shows by a Specific Director**

**Objective**: List all content directed by 'Rajiv Chilaka'.

**8. List All TV Shows with More Than 5 Seasons**

**Objective**: Identify TV shows with more than 5 seasons.

**9. Count the Number of Content Items in Each Genre**

**Objective**: Count the number of content items in each genre.

**10. Find the Top 5 Years with the Highest Average Content Release in India**

**Objective**: Calculate and rank years by the average number of content releases by India.

**11. List All Movies that are Documentaries**

**Objective**: Retrieve all movies classified as documentaries.

**12. Find All Content Without a Director**

**Objective**: List content that does not have a director listed.

**13. Find How Many Movies a Specific Actor Appeared in Over the Last 10 Years**

**Objective**: Count the number of movies featuring 'Salman Khan' in the last 10 years.

**14. Find the Top 10 Actors in Movies Produced in India**

**Objective**: Identify the top 10 actors with the most appearances in Indian-produced movies.

**15. Categorize Content Based on Keywords**

**Objective**: Categorize content as 'Bad' if its description contains 'kill' or 'violence' and 'Good' otherwise, and count the number of items in each category.

**Findings and Conclusion**

* **Content Distribution**: The dataset contains a diverse range of movies and TV shows with varying ratings and genres.
* **Common Ratings**: Insights into the most common ratings provide an understanding of the content's target audience.
* **Geographical Insights**: The top countries and the average content releases by India highlight regional content distribution.
* **Content Categorization**: Categorizing content based on specific keywords helps in understanding the nature of content available on Netflix.

This analysis provides a comprehensive view of Netflix's content and can help inform content strategy and decision-making.

**Acknowledgments**

The dataset for this project was sourced from Kaggle and highlighted by the YouTube channel [**Zero Analyst**](https://www.youtube.com/@zero_analyst) in one of their project videos. While the dataset idea originates from this source, all SQL queries for the analysis were written independently for this project.

**Author**

* **Parveen Jalwal**
* **LinkedIn**: <https://www.linkedin.com/in/parveen-jalwal-201a2a302>