

Saturday, 24 May

NETFLIX

SHOWS AND MOVIES



NETFLIX DATA ANALYSIS PROJECT USING SQL




Overview

This project aims to analyze Netflix's content using structured SQL queries to derive meaningful business insights. The analysis covers content type distribution, country-wise trends, director and actor performance, content ratings, and viewer-relevant classifications.

What I Learned from the Netflix SQL Analysis Project



**SQL Skills
Improved**



**Real-World
Data Analysis
Thinking**



**Tools &
Concepts
Practiced**



**Analytical
Confidence
Gained**

Goals

- Explore trends and patterns across Netflix's movies and TV shows.
- Identify popular genres, ratings, and contributing countries.
- Discover content gaps (e.g., missing directors) and top contributors (e.g., directors, actors).
- Classify content based on keywords and viewer perceptions.

Tools Used

- Database: PostgreSQL
- Language: SQL
- Dataset: Netflix content database (includes show_id, type, title, director, casts, country, date_added, release_year, rating, duration, listed_in, description)

Key Business Questions Solved

- **1** Count of Movies vs TV Shows
 - Reveals content distribution across types using GROUP BY.
- **2** Most Common Rating per Type
 - Uses RANK() to extract the top rating per content type.
- **3** Movies Released in 2020
 - Simple filter to show content by year and type.
- **4** Top 5 Countries by Content Count
 - Breaks down multi-country fields using STRING_TO_ARRAY and ranks by content volume.
- **5** Longest Movie on Netflix
 - Parses duration to find the longest movie using SPLIT_PART.
- **6** Content Added in the Last 5 Years
 - Converts date_added string to date for temporal filtering.
- **7** Top 3 Directors by Movie Count
 - Uses UNNEST and TRIM to extract director names and aggregate.

Key Business Questions Solved

- **8** TV Shows With More Than 5 Seasons
 - Filters TV shows based on season count using string parsing.
- **9** Top 5 Genres on Netflix
 - Unpacks genres using UNNEST, then counts appearances.
- **10** Average Annual Indian Releases
 - Calculates the percentage of Indian content released per year.
- **1 1** Bangladeshi Movies
 - Searches titles using ILIKE '%Bangladesh%'.
- **1 2** Content Without Directors
 - Counts entries with NULL director fields.
- **1 3** Salman Khan Movies (Last 10 Years)
 - Uses UNNEST, TRIM, and filtering by actor and date.
- **1 4** Top 10 Indian Actors by Movie Count
 - Ranks most frequent actors in Indian content using UNNEST.
- **1 5** Classify Content as "Good" or "Bad"
 - Uses CASE with ILIKE on the description to label and count content based on violent keywords.