Netflix SQL Data Analysis

Tools used : PostgreSQL, SQL, Excel (for exporting output), Google Docs, Python(Matplotlib & Seaborn)

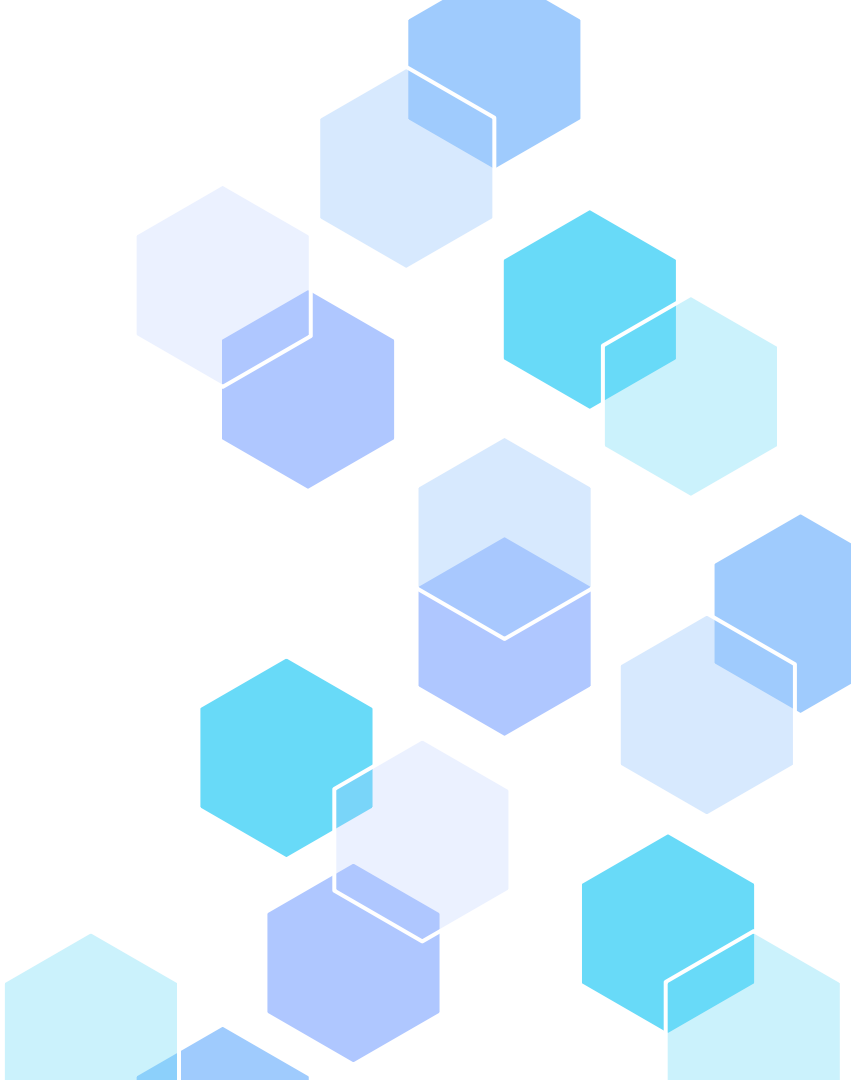
Submitted by – Priyanshu Virendra

Date: September 2025

01

Introduction

Executive Summary



Introduction

This project aims to analyze Netflix's content database using advanced SQL queries in PostgreSQL and visualize insights using Python. The goal is to identify trends, pattern, and opportunities for improvement in Netflix's content strategy by examining data related to content type, release year, ratings, country of origin, directors, and genres. I solved over 15 real world business questions, visualized results using Python charts, and derived actionable insights to support Netflix's data driven decisions.

Business Objective

Netflix want to make smarter content investment decisions. To support that, I was assigned the task of analyzing their historical content data and:

- Understand which content types dominate the platform
- Identify top-performing content categories, ratings, and directors.
- Highlight content trends over time and across regions.
- Identity potentially harmful content using description-based keywords.

Dataset Overview

- . **source:** Sample Netflix Dataset .
- **total overview:** 1,000+ content entries .
- **key rows:**

- Title .
- Director .
- Cast .
- Country .
- Release_Year .
- Rating . Duration .
- Listed_in .
- Type (Movies/TV Show) .
- Description

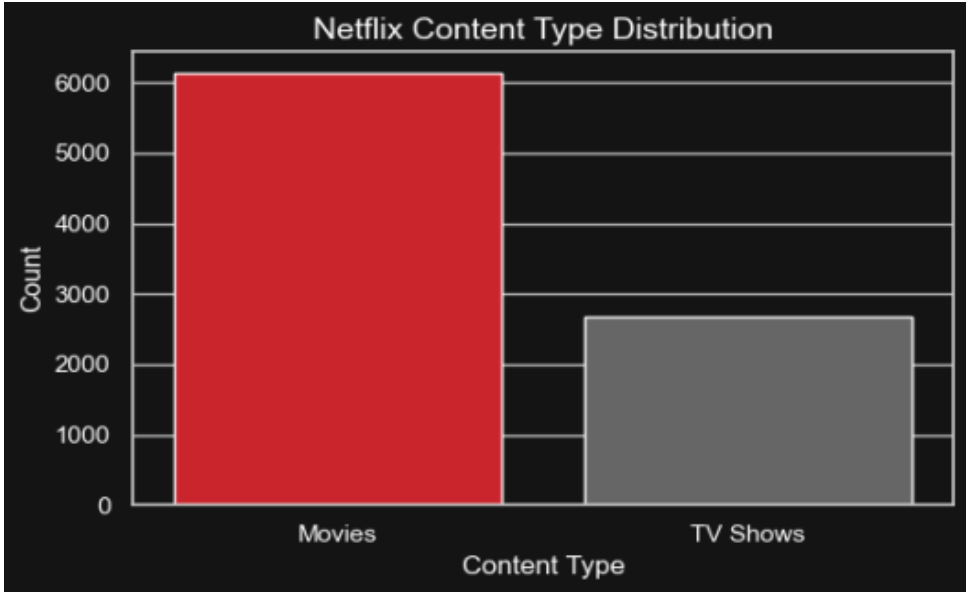
Key Questions Solved and Insights with mini scripts

Number of Movies vs TV Shows

Problem: Netflix needs to balance its content type inventory.

Insights: There are 6131 movies and 2676 TV shows.

Action: continue investing in TV shows as their engagement potential is higher over time.

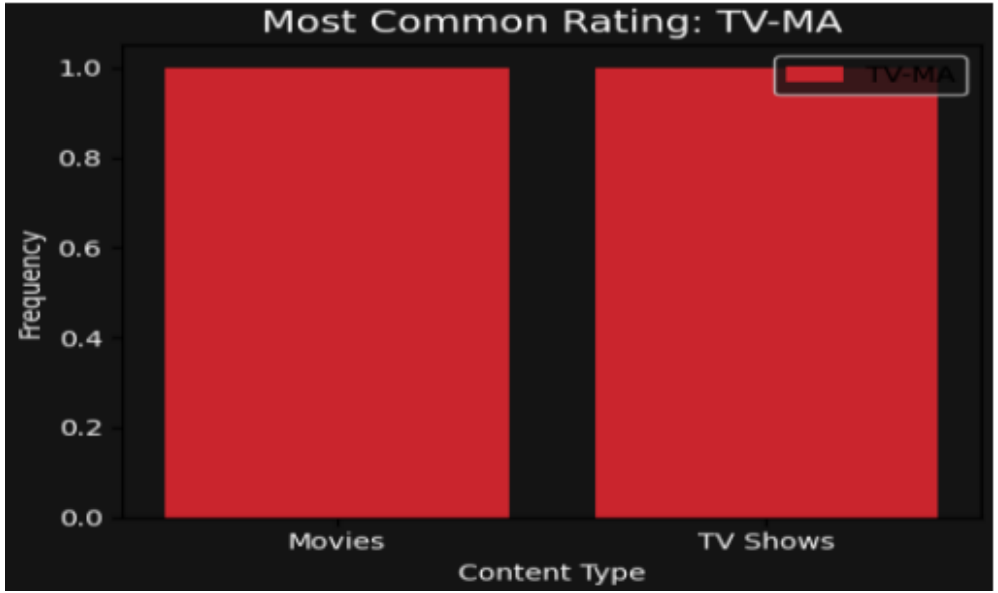


Key Questions Solved and Insights with mini scripts

Most Common Rating

Problem: Is content age-appropriate or skewed toward mature viewers?

Insights: TV-MA is the most frequent rating for both Movies and TV shows.
Action: Monitor and regulate the content balance between mature and general audience



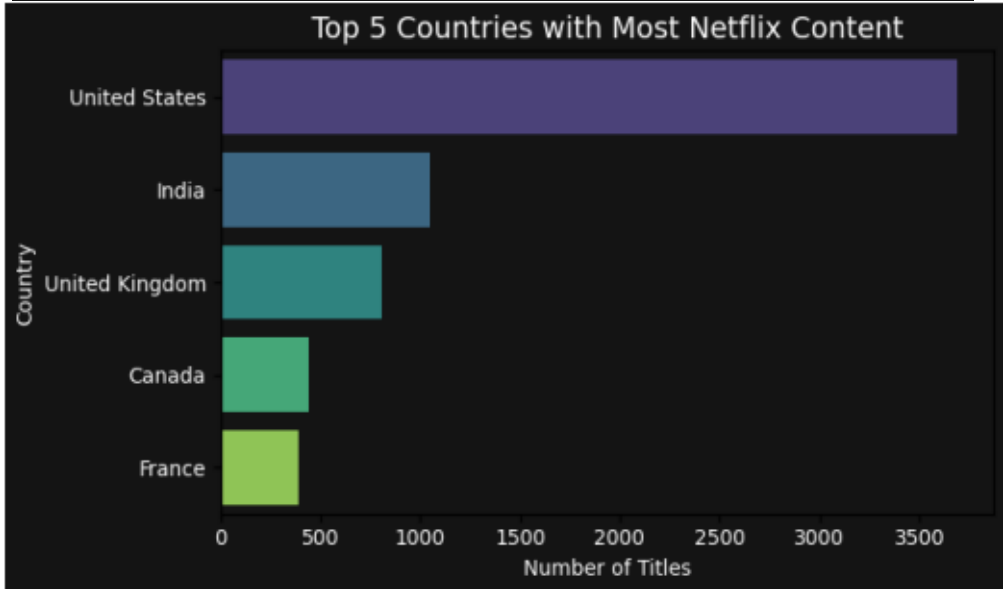
Key Questions Solved and Insights with mini scripts

Top 5 countries with Most Content

Problem: Where is the most of the content coming from geographically?

Insights: United Staes leads (3689), followed by India (1046), UK (804), Canada (445), France (393).

Action: Strengthen collaboration with creators from these topproducing countries.



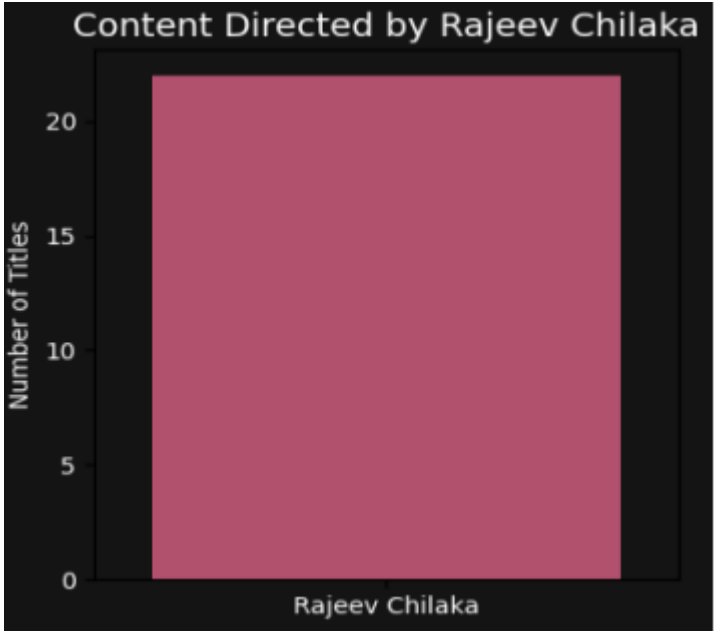
Key Questions Solved and Insights with mini scripts

TV Movies by Rajeev Chilaka

Problem: Evaluate directors with large content volumes.

Insights: Rajeev Chilaka has 22 TV movie listings.

Action: Consider partnerships or content campaigns highlight his work.



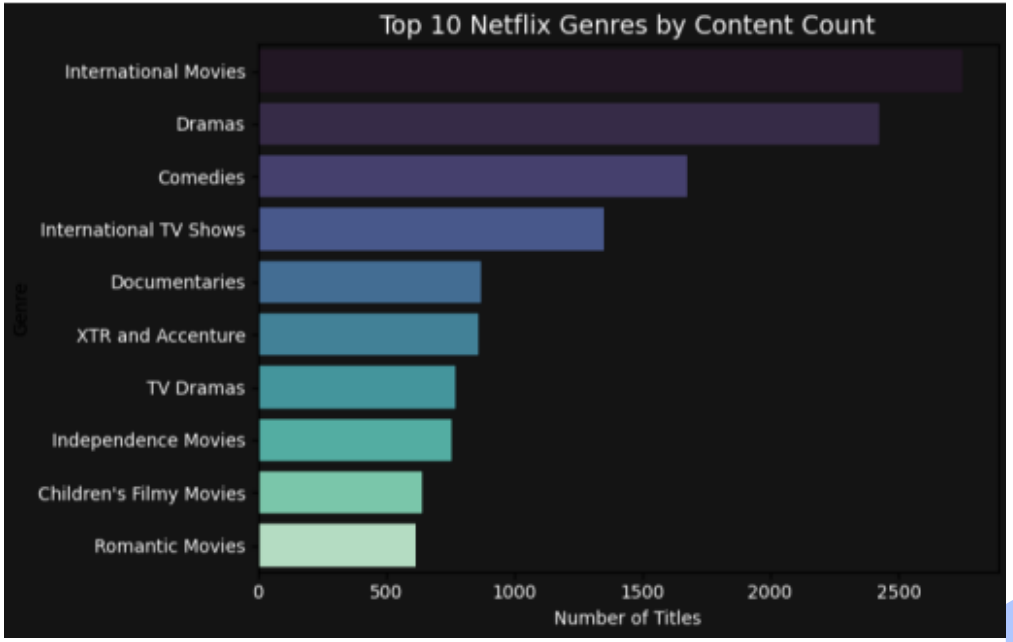
Key Questions Solved and Insights with mini scripts

Content count by Genre

Problem: What genres dominate Netflix?

Insights: International Movies (2752), Dramas (2427), Comedies (1674), and International TV shows (1351) lead.

Action: Expand offering in these genres while exploring underrepresented niches.



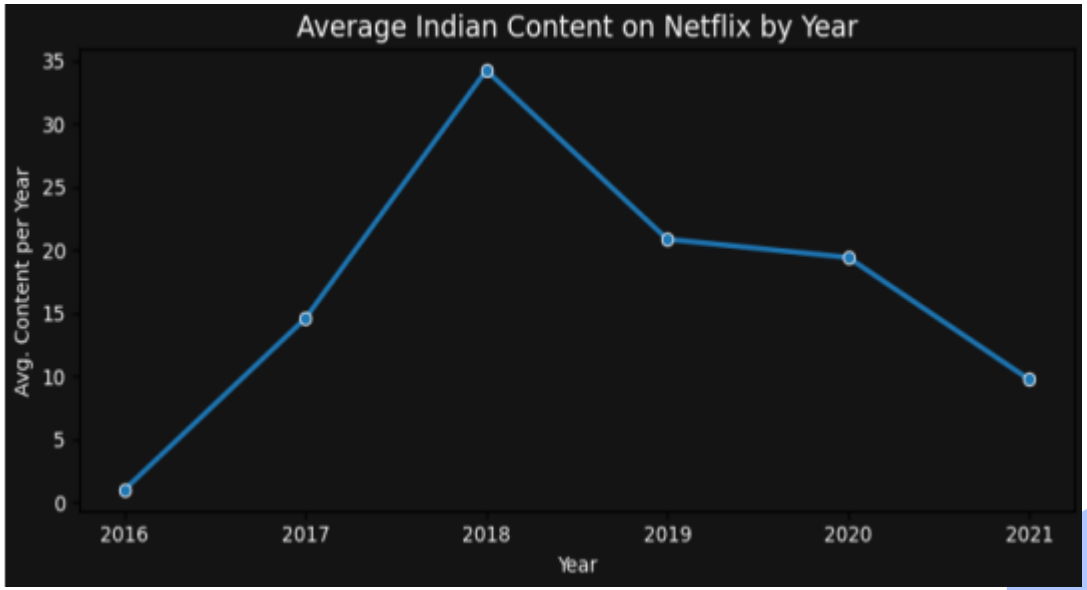
Key Questions Solved and Insights with mini scripts

Yearly Content Trend in India

Problem: How has Indian content grown over the years?

Insights: Peak in 2018 (333 contents), steady contribution from 2017-2021.

Action: Reignite growth by investing in regional and genrediverse Indian productions.



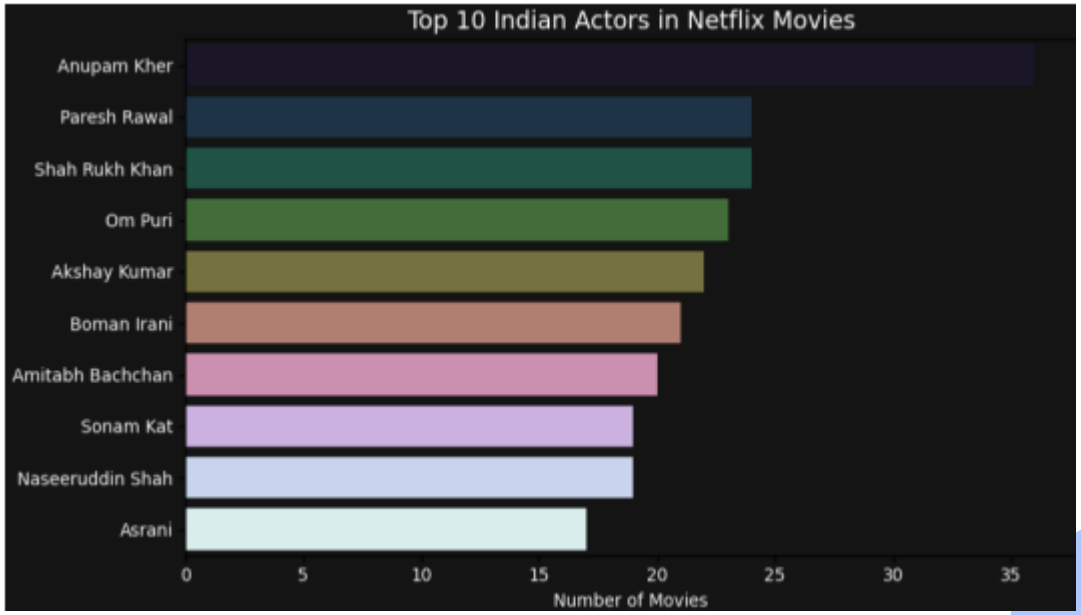
Key Questions Solved and Insights with mini scripts

Top 10 Indian Actors in Netflix Content

Problem: Who are the most featured actors in Indian Netflix films?

Insights: Anupam Kher (36), Paresh Rawal and Shah Rukh Khan (24 each), Om Puri (23), etc.

Action: Feature popular actors in marketing campaigns to increase audience pull.



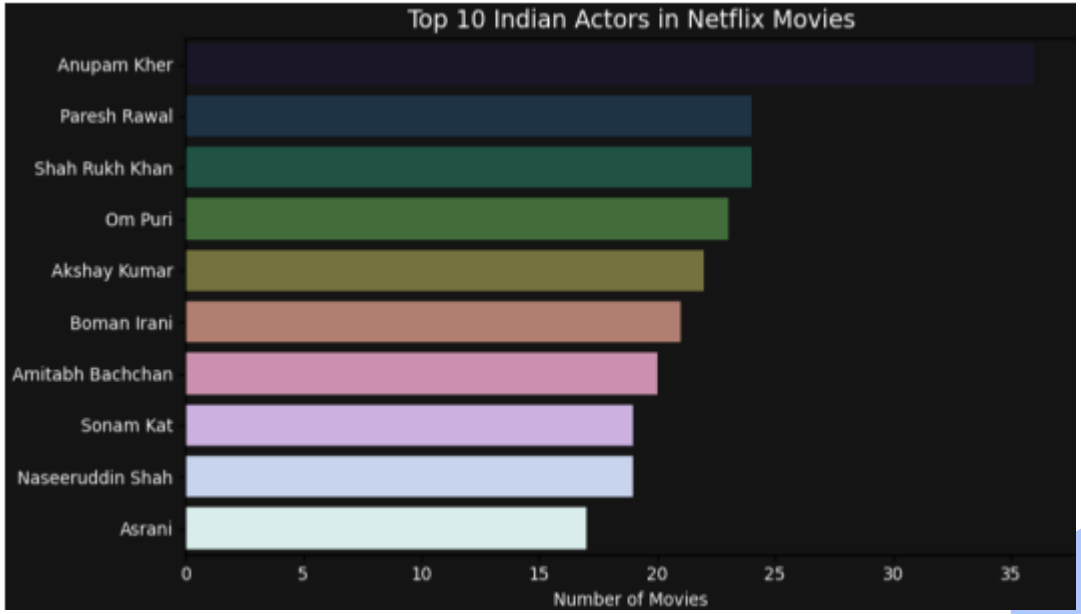
Key Questions Solved and Insights with mini scripts

Keyword Labelling (Good vs Bad Content)

Problem: Is there problematic content description?

Insights: 342 titles are flagged as “bad” (mentioning violence/skill), while 8465 are labeled “good”.

Action: Implemented content labeling policies and highlight positive messaging where applicable.



Recommendations

- Continue expanding high-performing genres like Dramas and International Movies.
- Partner with influential creators and actors, especially in top producing countries.
- Track ratings distribution for better content age segmentation.
- Prioritize production and marketing for TV Shows, especially in India.
- Apply sentiment or keyword-based filtering to regulate content tone and viewer perception.

Conclusions

This project provided a data-driven overview of Netflix's global content strategy using SQL and Python. The insights derived reveal patterns in content type, production geography, and genre trends. By acting on the recommendation, Netflix can make more informed, targeted decisions to improve viewer engagement and maintain content quality.