

# **Netflix Data Analysis**

Using Google BigQuery & Python

**Prepared by:**

Narasimha Kasu

**Date:**

September 3, 2025

The Netflix logo, consisting of the word "NETFLIX" in a bold, red, sans-serif font.

## Executive Summary

This report presents an analysis of the Netflix Movies & TV Shows dataset using SQL in Google BigQuery and Python visualizations. The objective was to explore content distribution, growth trends, genres, and ratings.

### Key Insights:

- Netflix's catalog is dominated by **Movies ( 70%)** compared to TV Shows ( 30%).
- The **United States and India** contribute the largest share of titles.
- Content additions **peaked in 2019**, followed by a decline after 2020.
- **TV-MA and TV-14** are the most common ratings, reflecting a focus on mature audiences.
- **Drama and Comedy** are the most dominant genres, appealing to a wide global audience.

This analysis highlights Netflix's global strategy, emphasizing movie-heavy content, regional dominance by the US and India, and strong focus on mature audiences.

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Tools Used</b>	<b>3</b>
<b>3</b>	<b>Methodology</b>	<b>3</b>
<b>4</b>	<b>Analysis and Results</b>	<b>4</b>
4.1	Movies vs TV Shows . . . . .	4
4.2	Top Countries . . . . .	4
4.3	Content Growth Over Time . . . . .	5
4.4	Ratings Breakdown . . . . .	5
4.5	Top Genres . . . . .	5
<b>5</b>	<b>Conclusion</b>	<b>6</b>
<b>6</b>	<b>References</b>	<b>6</b>

# 1 Introduction

This project analyzes the **Netflix Movies & TV Shows dataset** (around 8,800 titles) using SQL in Google BigQuery and Python visualizations. The goal is to answer key business questions about Netflix's catalog composition, growth patterns, and audience focus.

## 2 Tools Used

- Google BigQuery – SQL queries and data cleaning
- Python (Pandas, Matplotlib, Seaborn) – Data visualization
- Kaggle Notebook – Analysis & publishing

## 3 Methodology

1. Imported the dataset into Google BigQuery.
2. Cleaned the `date_added` column to create `titles_clean`.
3. Wrote SQL queries for distribution, trends, genres, and ratings.
4. Exported results to Kaggle Notebook for visualizations.
5. Created charts and insights.

## 4 Analysis and Results

### 4.1 Movies vs TV Shows

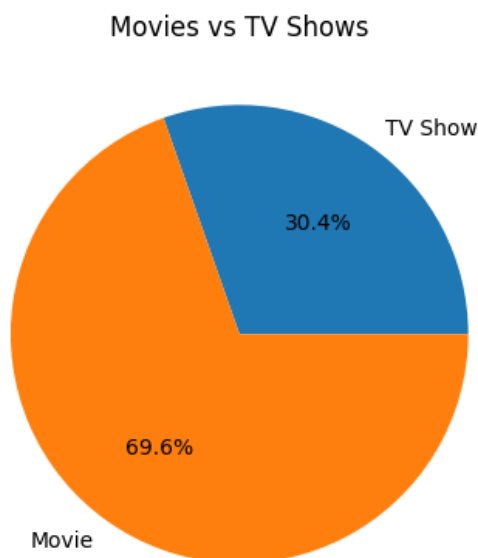


Figure 1: Distribution of Movies vs TV Shows.

**Insight:** Movies dominate Netflix's catalog ( 70%) compared to TV Shows ( 30%).

### 4.2 Top Countries

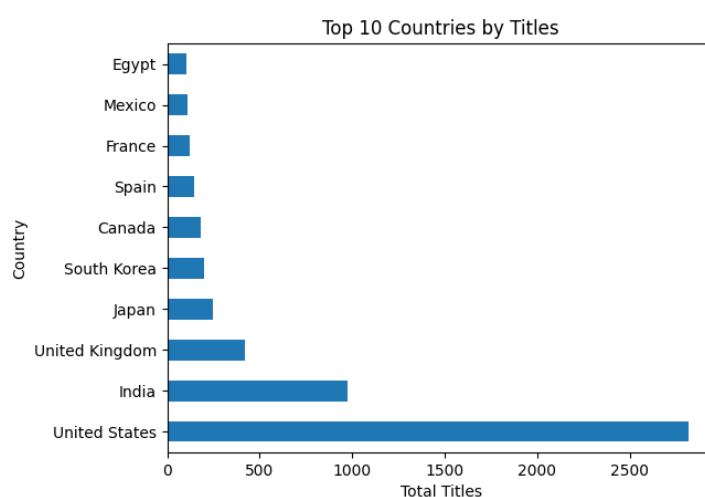


Figure 2: Top contributing countries on Netflix.

**Insight:** The United States and India contribute the largest share of titles.

### 4.3 Content Growth Over Time

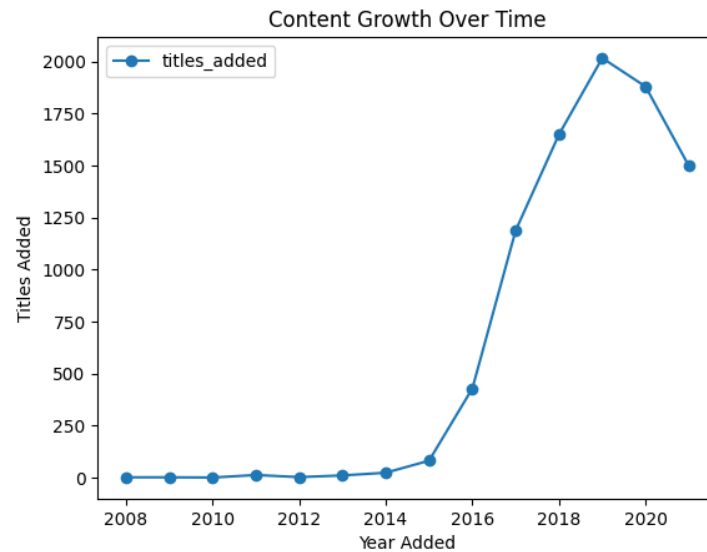


Figure 3: Growth of titles over time.

**Insight:** Content additions peaked in 2019 and slowed after 2020.

### 4.4 Ratings Breakdown

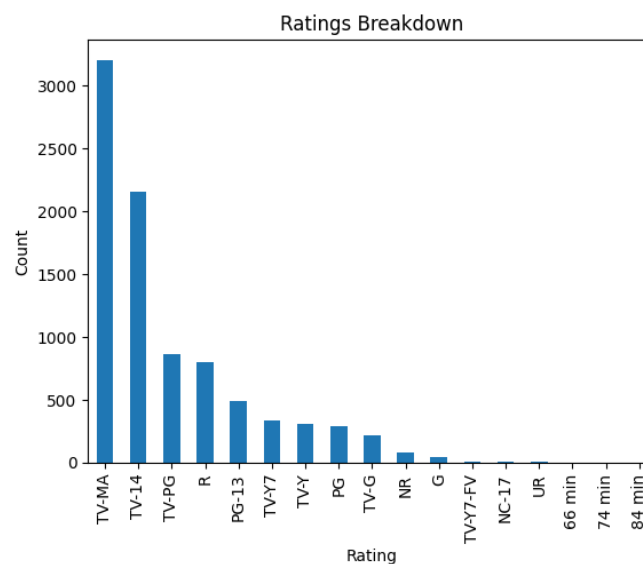


Figure 4: Distribution of ratings.

**Insight:** TV-MA and TV-14 dominate, showing Netflix's focus on mature audiences.

### 4.5 Top Genres

**Insight:** Drama and Comedy dominate the catalog, followed by Action and Thriller.

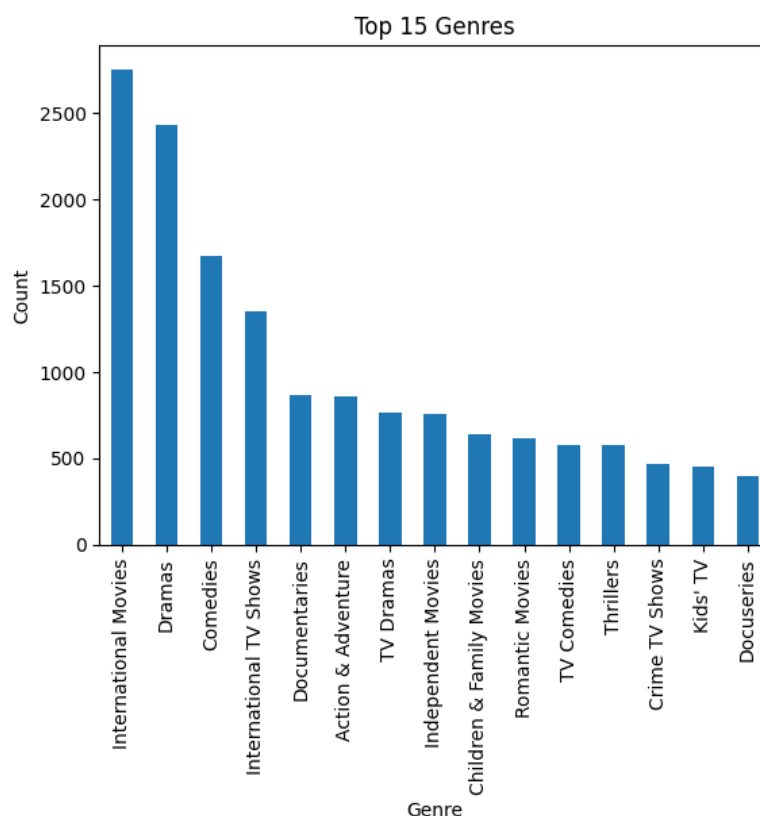


Figure 5: Top genres available on Netflix.

## 5 Conclusion

- Netflix is dominated by **Movies ( 70% of titles)**.
- **The US and India** lead in content production.
- Content additions **peaked in 2019**.
- **Drama and Comedy** are the most popular genres.
- Mature ratings (**TV-MA, TV-14**) are most frequent.

## 6 References

- Netflix Dataset (Kaggle): <https://www.kaggle.com/datasets/shivamb/netflix-shows>
- Google BigQuery Documentation: <https://cloud.google.com/bigquery/docs>