NETFLIX Data SQL Analysis

SUBTITLE HERE

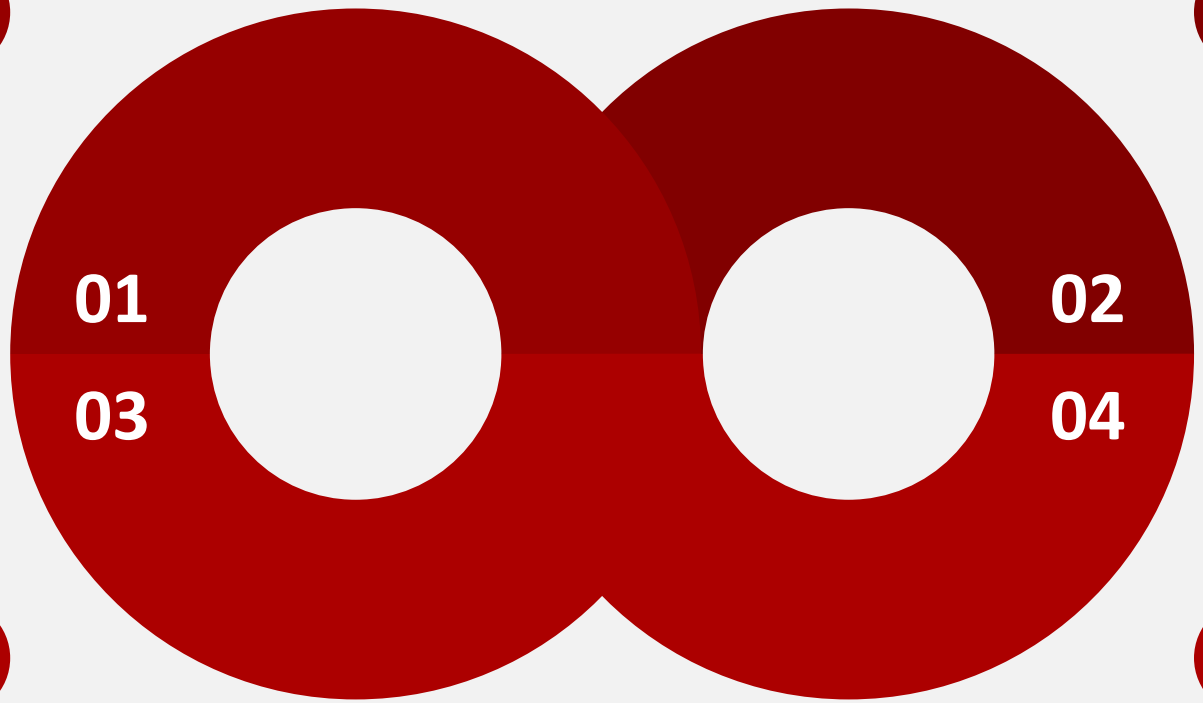
Project Overview

01 SQL-based analysis of Netflix dataset

02 Focus on Movies, TV Shows, Genres, Ratings, Countries, Actors

03 Answered 15 real-world business questions

04 Useful for streaming platforms, analysts & strategists



Dataset Used

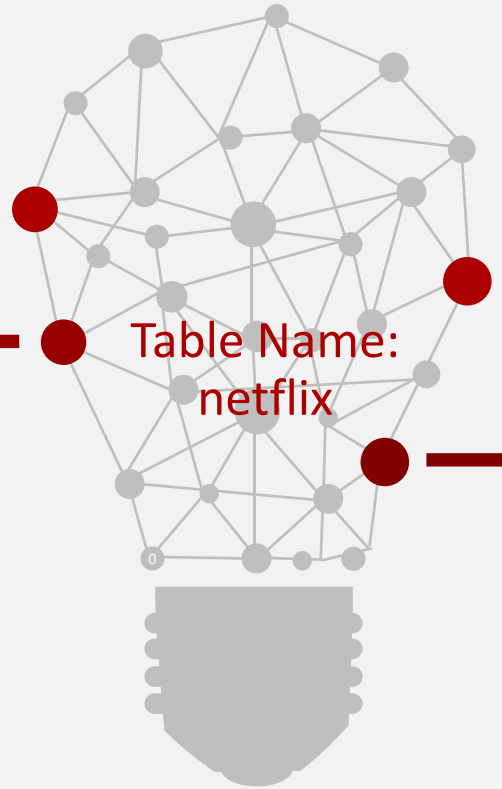
Key Columns:

1

show_id, type, title, director, casts

2

country, date_added, release_year



3

rating, duration, listed_in, description

4

Source: Single Excel dataset → imported into PostgreSQL

01 02 03

Project Objectives

Compare Movies vs TV Shows count

Analyze country-wise distribution

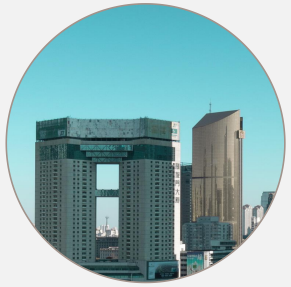
Track yearly release trends (India focus)

Identify top actors & directors

Study ratings & audience segmentation

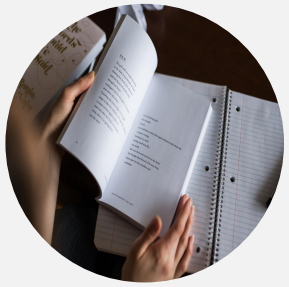
Detect violent/sensitive content patterns

Business Problems Addressed (Part 1)



2025

How many Movies vs TV Shows exist?

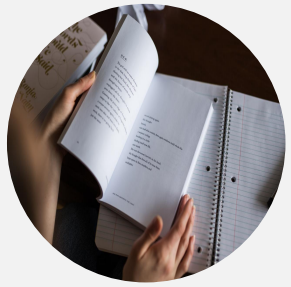


2025

What's the most common rating?

2025

Which year had the highest releases?



2025

Top 5 countries producing Netflix content?

2025

Which are the longest movies?

Business Problems Addressed (Part 2)

01

- Content added in last 5 years
- Which directors have the most shows?

02

- TV shows with more than 5 seasons
- Which genres dominate Netflix?

03

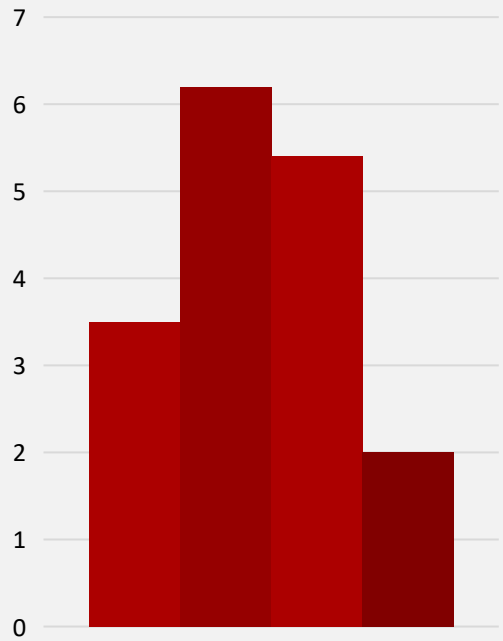
- Netflix content from India (trends by year)
- Actor insights (Salman Khan + Top 10 Indian actors)

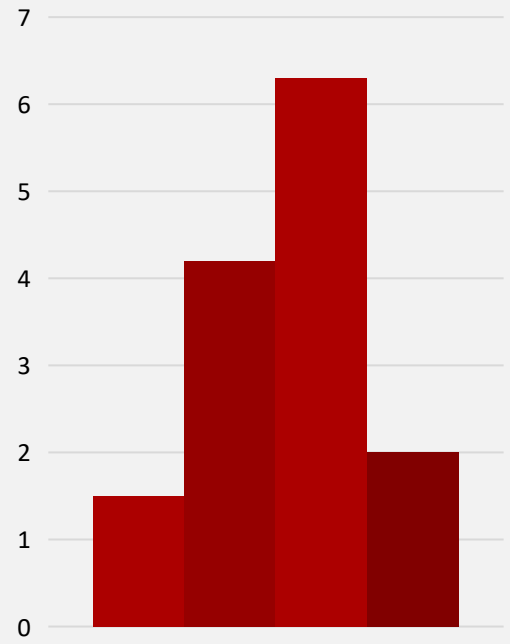
04

- Content flagged as Good vs Bad (violence)

Methodology

Steps followed:



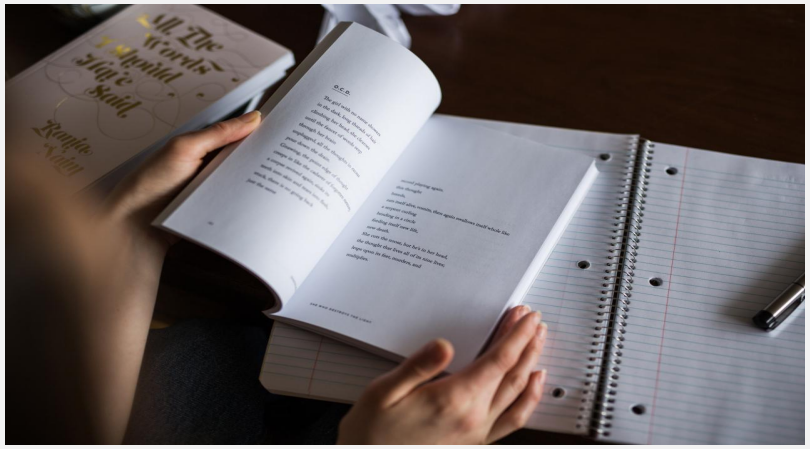


1. Data Preparation – Import Excel → PostgreSQL, assign data types
2. Basic Exploration – SELECT, DISTINCT, WHERE
3. Advanced Analysis – Aggregations, Date filters, Text search
4. Array Handling – UNNEST, STRING_TO_ARRAY for actors/genres

Key Insights

1. Movies dominate over TV Shows
2. Most common rating = TV-MA (mature audience)
3. Top producing countries = USA, India
4. Longest movie identified by runtime field
5. Recent years = steady content growth
5. Top Indian actors frequently featured
6. Salman Khan appeared in multiple movies (last 10 years)
7. Genres: Drama, Comedy, International Movies dominate
8. Categorization: Good vs Bad (Violent content detection)

Sample Queries



```sql
--15. Categorize the content based on the presence of the keywords 'kill' and 'violence' in
--content as 'Good'. Count how many items fall into each category.
SELECT category,TYPE,
COUNT(*) AS content_count
FROM (SELECT *,
        CASE
            WHEN description ILIKE '%kill%' OR description ILIKE '%violence%' THEN 'Bad'
            ELSE 'Good'
        END AS category
    FROM netflix
) AS categorized_content
GROUP BY 1,2
ORDER BY 2;
```

Output Messages Notifications

Showing rows: 1 to 4 Page No: 1

category text	type character varying (10)	content_count bigint
Bad	Movie	110
Good	Movie	2681
Bad	TV Show	20
Good	TV Show	905

Successfully run. Total query runtime: 2

Skills Gained

SQL wrangling with PostgreSQL

Using ILIKE, UNNEST, SPLIT_PART

Data aggregation & ranking

Querying business-focused problems

Entertainment trend analysis

Why This Project Matters

02 Can extend to Power BI/Tableau dashboards

01 Simulates real streaming analytics

03 Perfect for GitHub/LinkedIn portfolio

03 Proof of SQL skills on real dataset

03 Strong interview showcase project

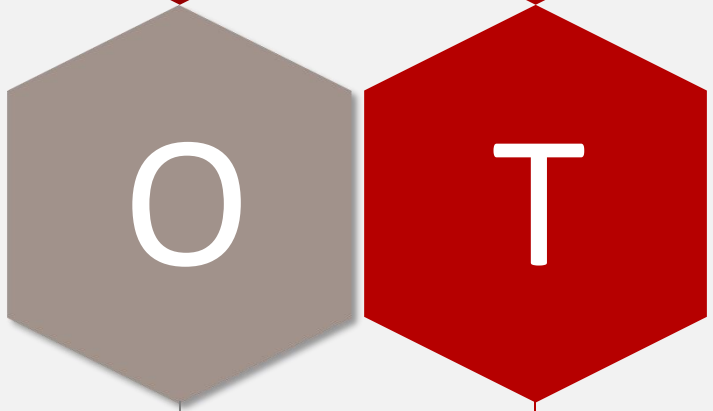
Conclusion

Analyzed Netflix dataset with 15+ SQL queries

Movies vs TV Shows

Actors & Directors

Countries & Genres



Ratings & Trends

Sensitive content categorization

Thank You!

Every great presentation is complete with a great audience
— and that’s you!