DATA ANALYTICS

REINFORCEMENT PROJECT DOCUMENTATION

NETFLIX MOVIES & TV SHOWS

A. DEVIBALA
DADS -JUNE(2025)

19.11.2025

INTRODUCTION :

The entertainment industry has experienced massive growth in digital streaming platforms. Among them, Netflix stands as the global leader, offering thousands of Movies and TV Shows across various genres, languages, and countries.

This project aims to analyze the Netflix Titles Dataset and build an interactive dashboard using Power BI to help users understand the platform's content distribution, growth trends, genre variations, ratings, and global reach.

The dashboard provides insight into:

- How Netflix content is distributed worldwide
- Popular genres and ratings
- Trends in content creation over time
- Comparison between Movies vs TV Shows
- Countries contributing the most content

This documentation provides a complete overview of the project process, data modeling, analysis, findings, and conclusions.

PROBLEM STATEMENT :

Netflix contains thousands of shows, but understanding the patterns behind:

- which countries create the most content,
- which genres are most popular,
- how content has grown over time,
- what kinds of ratings or maturity levels dominate,
- the proportion of Movies vs TV Shows,

is not directly visible from raw data.

So, the challenge is:

How can we convert raw Netflix data into meaningful insights using Power BI?

OBJECTIVES OF THE PROJECT :

The main objectives are:

Business Objectives :

- Analyze Netflix's content library to understand global entertainment patterns.
- Provide insights for content creation decisions, production focus, and market trends.

Technical Objectives :

- Clean and transform Netflix dataset using Power BI.
- Use DAX to create advanced calculations.
- Design an interactive and visually appealing dashboard.
- Present key metrics (KPIs) in a structured format.

To analyze Netflix's content library using data visualization

To identify insights about:

- Total shows, movies, ratings, genres, and directors.
- Countries producing the highest number of movies/shows.
- Genre popularity and audience certification.
- Trends in content release over time.

To create an interactive Power BI dashboard for better understanding and decision-making.

DATASET DESCRIPTION :

Dataset used: "Netflix Titles Dataset" (netflix_titles.CSV)

DATASET COLUMNS :

COLUMN NAME DESCRIPTION :

show_id	Unique identifier
type	Movie or TV Show
title	Name of the show
director	Director of the content
cast	Actors featured
country	Country where show was produced
date_added	Date Netflix added the title
release_year	Year of original release
rating	Maturity rating (TV-MA, R, PG-13, etc.)
duration	Runtime (Movies) / Number of Seasons (TV Shows)
listed_in	Genres
description	Summary

Dataset Size :

- **Total Records:** 8800+
- **Movies:** 6131
- **TV Shows:** 2680+
- **Total Ratings:** 19
- **Total Genres:** 515 unique classifications

TOOLS & TECHNOLOGIES USED

Software

- **Power BI Desktop**
- **Power Query** (for data transformation)

Techniques

- Data Cleaning
- DAX (Data Analysis Expressions)
- Data Modeling
- Dashboard Design
- Visual Analytics

DATA CLEANING AND TRANSFORMATION :

Data was cleaned using **Power Query Editor** in Power BI.

Steps Taken:

1. **Removed Duplicate Rows**
   Some records were repeated, so duplicates were removed.
2. **Handled Null Values**
   Missing values in *director, cast, country* were replaced with "Not Available".
3. **Split Columns**

- Genre column ("listed_in") contained multiple genres → kept as is for summary analysis.
- Country column had multiple values → used *first country* for mapping.

4. **Formatted Columns**
    - Converted date_added to Date format.
    - Duration column split into:
        - "Duration Minutes" for Movies
        - "Number of Seasons" for TV Shows
5. **Removed Unnecessary Spaces and Symbols**

DATA MODELING

This project uses a **single-table model** since Netflix dataset is a flat file. DAX measures were created for:

- Total Movies
- Total TV Shows
- Unique Directors
- Genre Count
- Ratings Count
- Year-wise Show Count

No relationships were needed because it is a single dataset.

DAX MEASURES USED :

Total Movies Added using:

```
Total Movies = CALCULATE(COUNTROWS('netflix_titles'),
'netflix_titles'[type]="Movie")
```

DASHBOARD EXPLANATION :

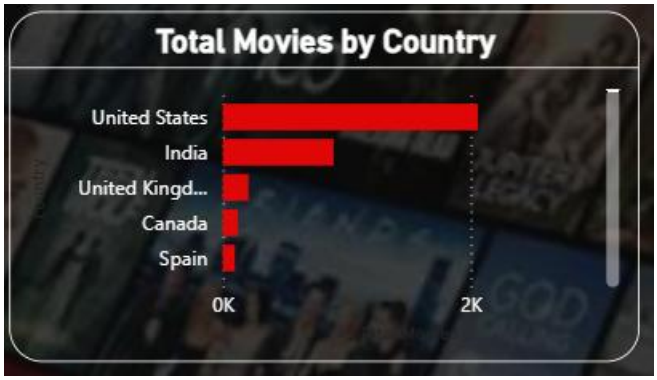
KPI SECTION

This section provides high-level numbers:

- **Total Shows:** 8800
- **Total Ratings:** 19
- **Unique Genres:** 515
- **Total Directors:** 4527
- **Total Movies:** 6131

These KPIs immediately give a snapshot of Netflix's entire catalog.

1.Total Movies by Country (Bar Chart)

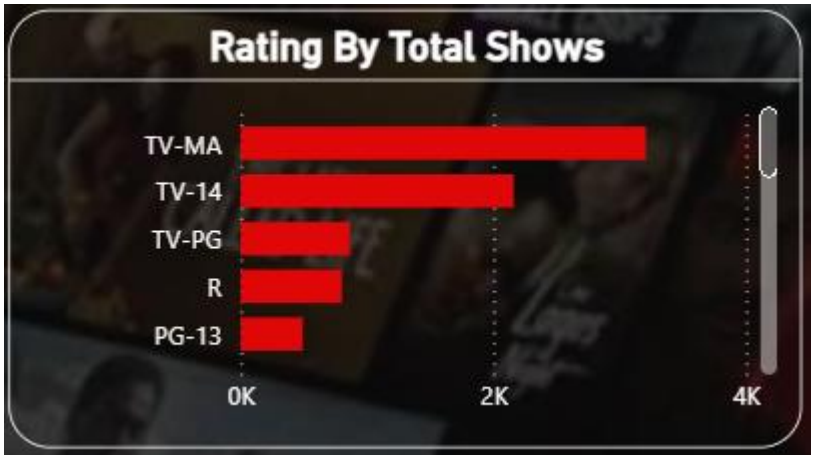


Insights:

- **United States** has the highest movie production.
- **India** is the second highest, showing huge film output.
- Countries like U.K., Canada, Spain follow next.

This indicates Netflix has strong partnerships with these countries.

2. Rating by Total Shows (Bar Chart)



Insights:

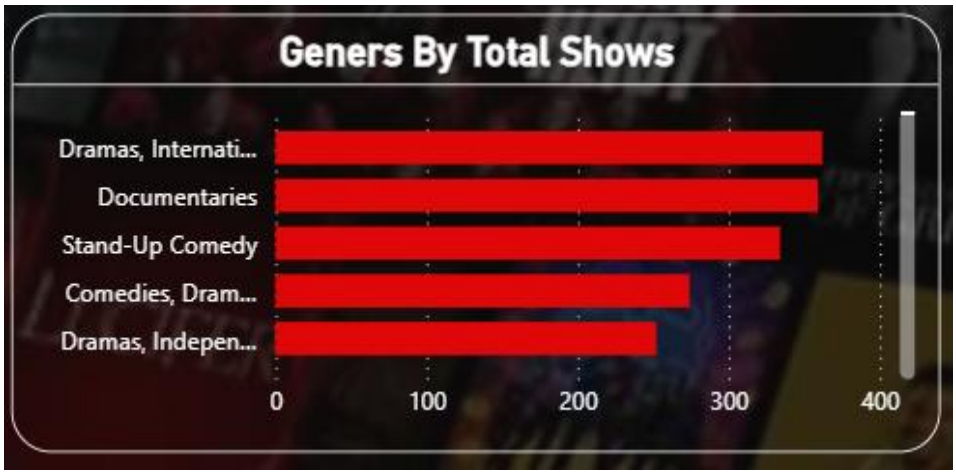
Shows are categorized based on maturity ratings:

- **TV-MA** is the highest → Mature audience content dominates.
- **TV-14** is next → Suitable for teens + adults.
- Lower ratings like PG, R, PG-13 appear lesser.

This shows Netflix largely produces adult-oriented content.

---

3. Genres by Total Shows (Bar Chart)



Insights:

Top genres include:

- Documentaries
- International Dramas

- Stand-Up Comedy
- Comedies

This suggests Netflix invests heavily in:

- Real-life content
- Global cultural storytelling
- Comedy entertainment

4. Total Shows by Country (Map Visualization)

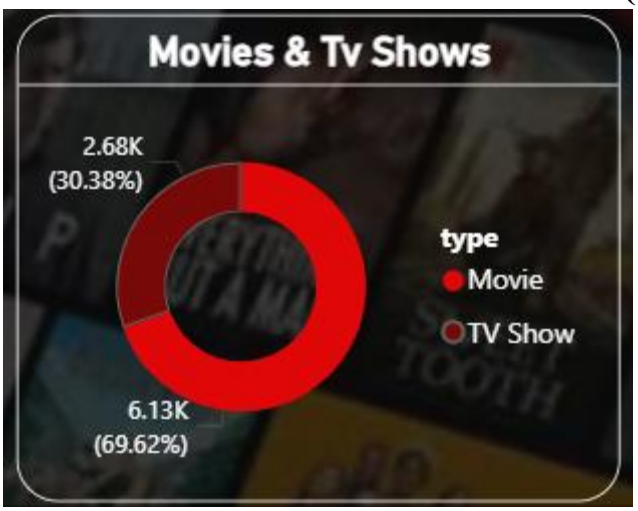
Insights:

A world map shows:

- Dark red regions = high content production
- Light colors = low content

Helps visually identify content-rich countries.

5. Movies vs TV Shows Distribution (Donut Chart)



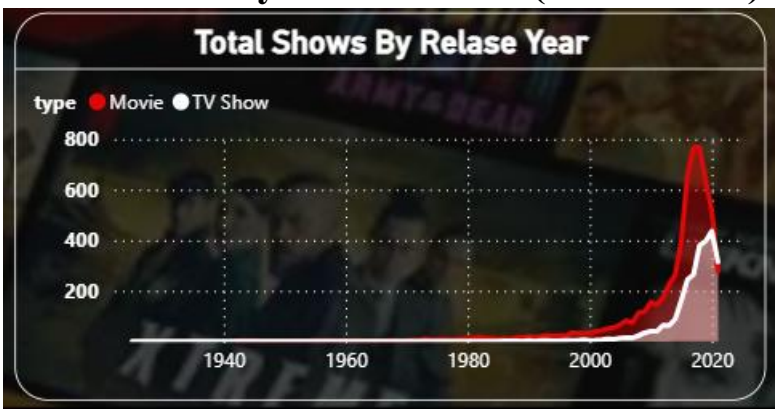
Insights:

Distribution:

- **Movies:** 69.62%
- **TV Shows:** 30.38%

Netflix library is dominated by movies but still maintains a significant TV series catalog.

6. Total Shows by Release Year (Line Chart)



Insights:

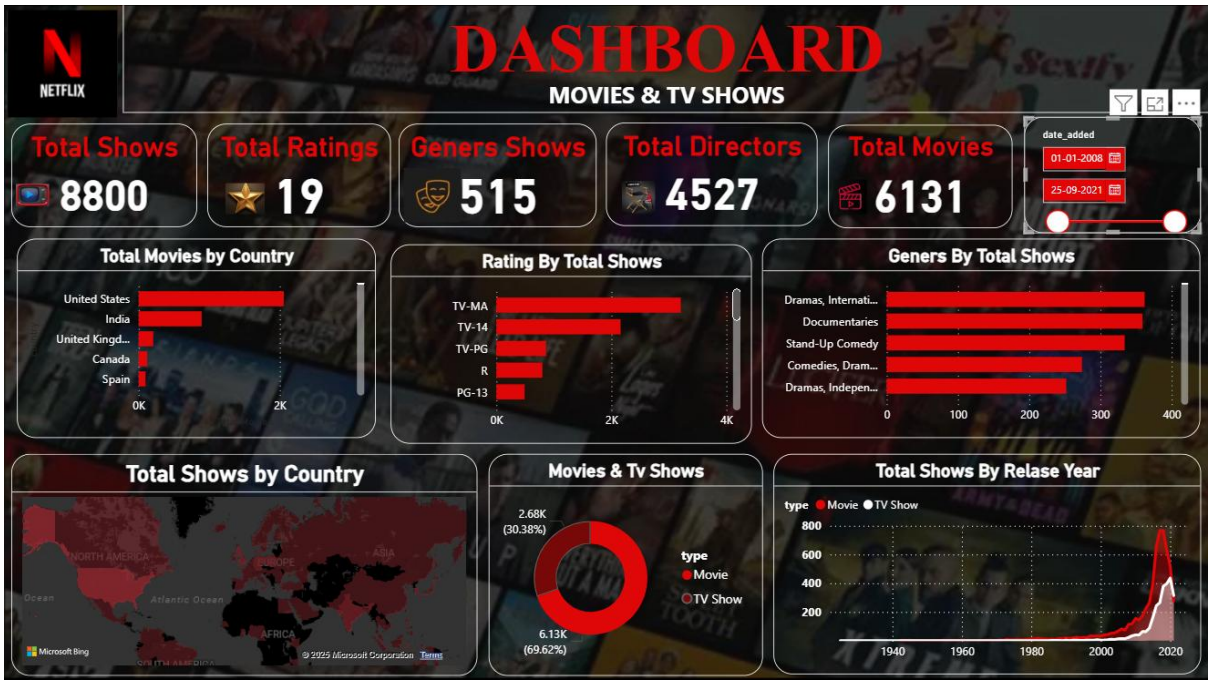
Shows:

- Almost no content before 1960
- Slow growth from 1960–2000
- **Massive explosion after 2015**

- Peak between **2018–2020**

This indicates Netflix's rapid rise as a content-production platform.

DASHBOARD OVERVIEW :



KEY INSIGHTS (VERY DETAILED) :

1. USA leads global Netflix content production

The United States contributes more content than any other country.

2. India's strong content presence

India is among the top contributors globally, especially for movies.

3. Netflix focuses on adult-oriented content

Most shows are rated:

- TV-MA
- TV-14

Suggesting content targeting mature audience groups.

4. Documentary content is very strong

Documentaries are among top genres, showing demand for real-life stories and learning content.

5. International content growing fast

Genres like *International Dramas* show Netflix is focused on global content diversity.

6. Major growth after 2015

Netflix increased production and acquisition aggressively since 2015.

7. Movies dominate

Movies account for 70% of all content.

CONCLUSION :

The Netflix Dashboard gives a complete analytical view of the content library. This project helped uncover:

- Worldwide content production patterns
- Rating-wise and genre-wise show distribution
- Year-wise trends showing Netflix’s explosive growth
- Movie vs TV Show dominance
- Strong global diversification in Netflix's content

This dashboard is useful for:

- Data analysts
- Entertainment industry researchers
- Media companies
- Academic case studies
- Business decision-making

It successfully transforms raw data into meaningful insights.

FUTURE IMPROVEMENTS :

In future versions, the following enhancements can be added:

✓ **Integrate IMDb or Rotten Tomatoes score**

To analyze audience preference and content quality.

✓ **Sentiment analysis on descriptions**

To classify content based on theme or mood.

✓ **Predictive Analytics**

To forecast future Netflix content trends.

✓ **Genre clustering**

To identify hidden genre groups.

✓ **Actor-based analysis**

To find top-performing actors across Netflix.