Netflix Dataset Report

Objective

This project focuses on cleaning and analyzing a dataset of Netflix titles using Python. The aim is to explore trends in content release, identify genre patterns, highlight top contributors, and assess user-targeted content to better understand Netflix's content strategy between 2008 and 2021.

Dataset Overview

Each row in the dataset represents a Netflix title with attributes such as:

- Type (Movie or TV Show)
- Title
- Director
- Country
- Date Added
- Release Year
- Rating (e.g., TV-MA, PG-13)
- Duration
- Genre Tags (listed_in)

Data Cleaning Process

1. **Removed Duplicate Records**
   Duplicates were identified and eliminated to maintain unique title entries.
2. **Handled Missing Values**
   Rows lacking crucial details (e.g., `director`, `cast`, or `country`) were dropped.
3. **Converted Data Types**
   - `date_added` converted to datetime
   - New columns for `year`, `month`, and `day` were extracted
4. **Split and Structured Genres**
   - 'listed_in' was split to analyze genres separately.

Exploratory Data Analysis (EDA)

1. Content Type Distribution

- **Movies**: 6,126 (~70%)
- **TV Shows**: 2,664 (~30%)

- *Movies dominate the Netflix catalog.*

Pie and bar charts were used to visualize this.

2. Ratings Breakdown

The most common ratings were:

Rating	Count
TV-MA	3,205
TV-14	2,157
TV-PG	861

This indicates a major focus on mature and teen content.

3. Top Contributing Countries

Country	No. of Titles
United States	3,240
India	1,057
United Kingdom	638

These countries lead Netflix content production.

4. Monthly Release Trends

- Most content is released in **July and August**, likely targeting holiday watchers.
- Movie releases consistently outnumber TV shows each month.

5. Yearly Release Trends

- Rapid growth after **2015**, peaking around **2019–2020**.
- Indicates Netflix's content expansion strategy in the late 2010s.

6. Popular Genres

Movies:

- International Movies
- Dramas
- Comedies

TV Shows:

- International TV Shows
- Dramas
- Kids TV

These insights highlight Netflix's focus on culturally diverse and family content.

7. Top Directors

The most frequently featured directors on Netflix include:

- Rajiv Chilaka
- Alastair Fothergill
- Raul Campos & Jan Suter
- Suhas Kadav
- Martin Scorsese
- Steven Spielberg

These names reflect a mix of children's content creators and globally acclaimed directors.

Key Insights

- **Movies are prioritized** over TV shows in terms of quantity.
- **Mature-rated content** (TV-MA, R) forms a significant chunk.
- **USA, India, and UK** dominate the content landscape.
- **Genres** like Drama, Comedy, and International Films are heavily represented.
- **Summer months** see the highest frequency of new releases.