Netflix Data Analysis: Report

Introduction:

This report provides a high-level analysis of Netflix's content catalog, sourced from the latest dataset (~8,800 titles). The objective is to uncover platform trends and inform on content variety, country representation, and prominent genres, actors, and directors.

Data Overview:

- **Total Titles:** 8,807

- **Dataset Columns:** show_id, type, title, director, cast, country, date_added, release_year, rating, duration, listed_in, description

- **Missing Data:**
  Director – 2,634 missing | Cast – 825 missing | Country – 831 missing
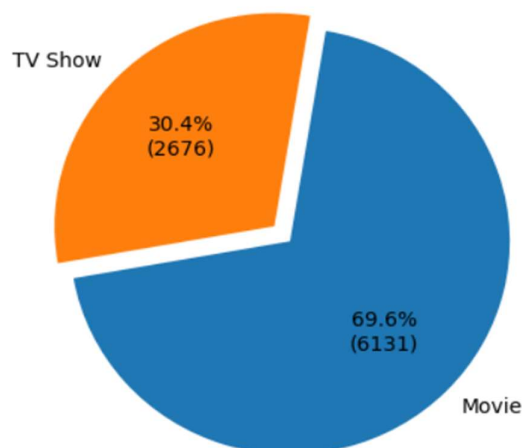
Data Cleaning:

- Imputed missing values with "Not Specified" or "Unknown" as appropriate

- Parsed and extracted year/month from date_added

- Duplicates removed

Key Insights:

1. Content Breakdown:
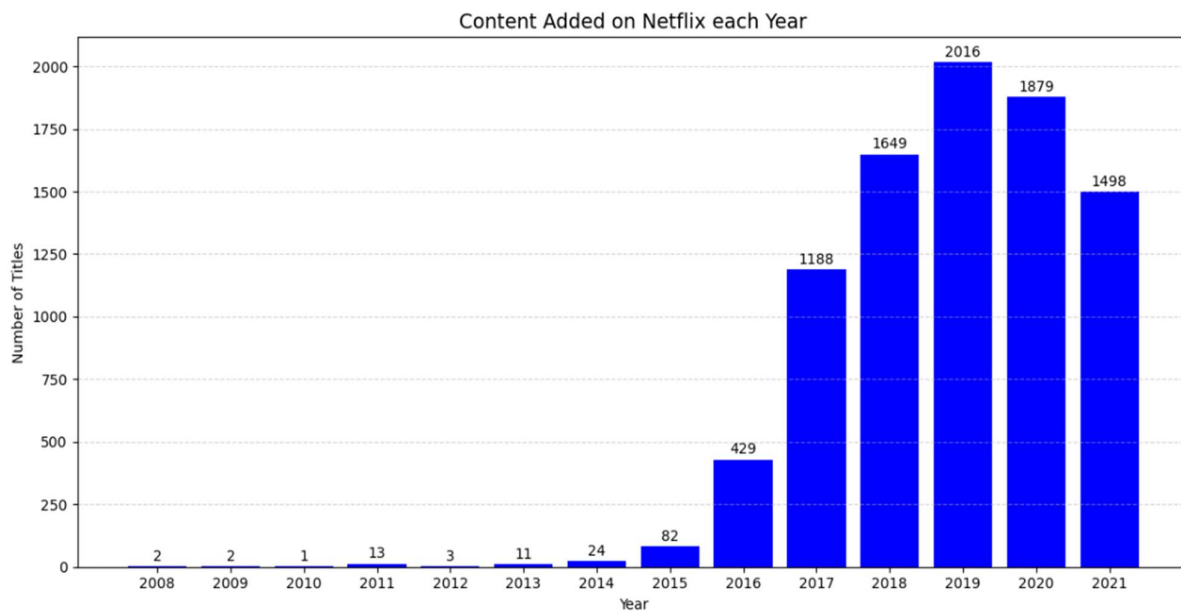
- **Movies:** 6131 (69.6%)
- **TV Shows:** 2676 (30.4%)

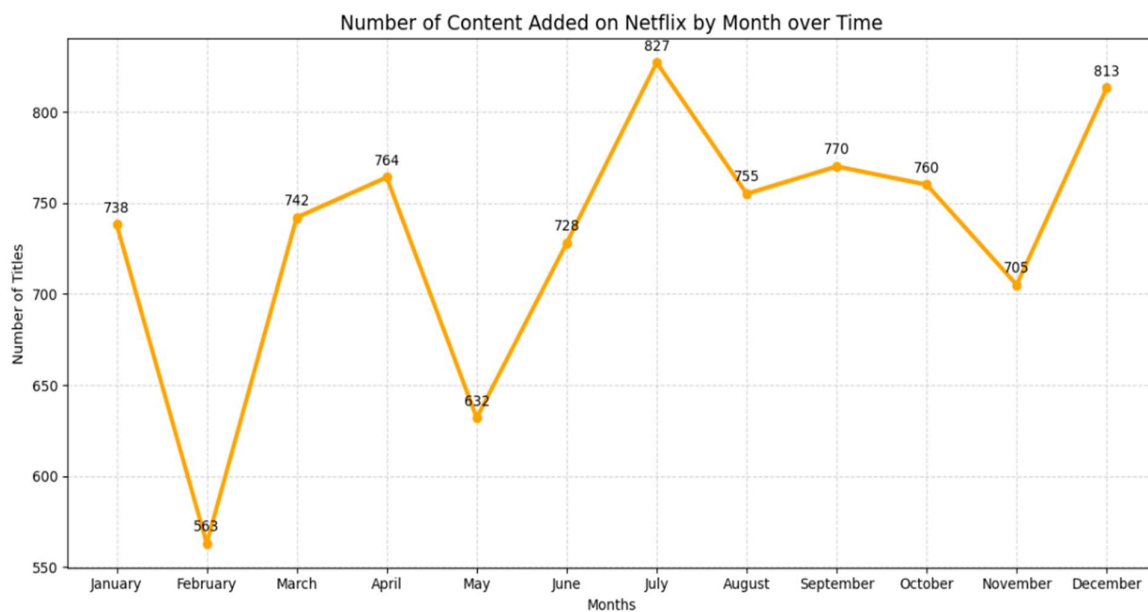Distribution of Movies v/s TV Shows on Netflix

2. Trend Over Years:

- **Oldest Content:**
  Pioneers: First Women Filmmakers (1925, TV Show)

- **Most Recent Content (2021):**
  Blood & Water (South Africa), *Ganglands*, *Jailbirds New Orleans*, *Kota Factory* (India), *Midnight Mass*

- **Peak Content Addition Year:**
  Recent years, especially 2019–2021, show a spike in content addition.
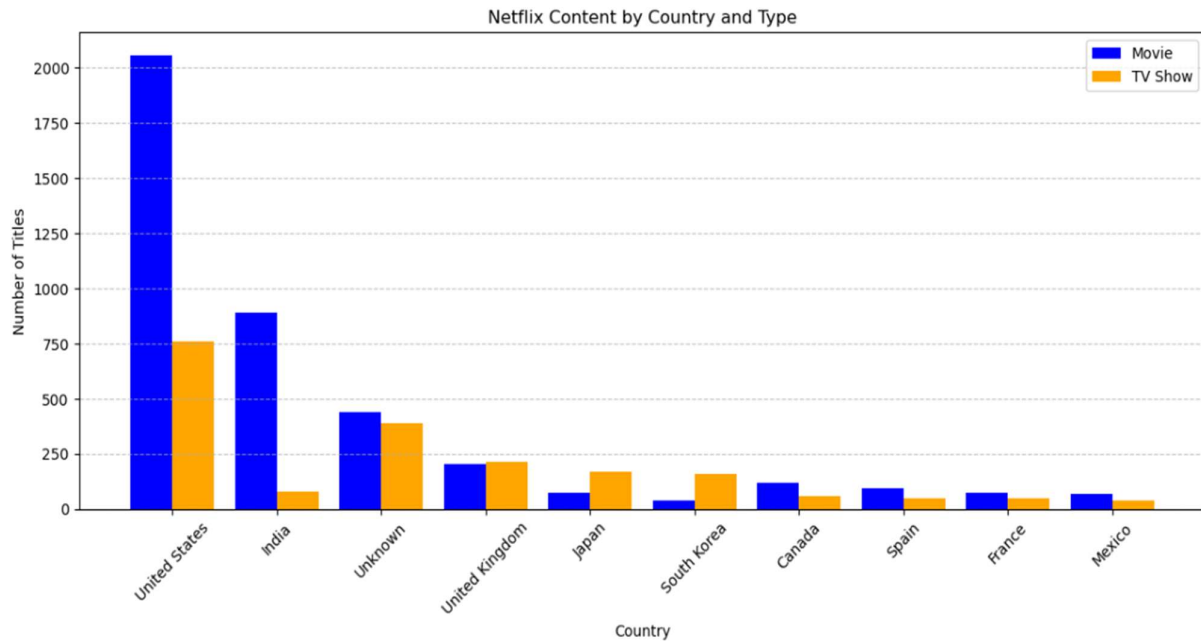

- **Most Common Addition Month:**
  Generally, July or December have peaks for content additions.
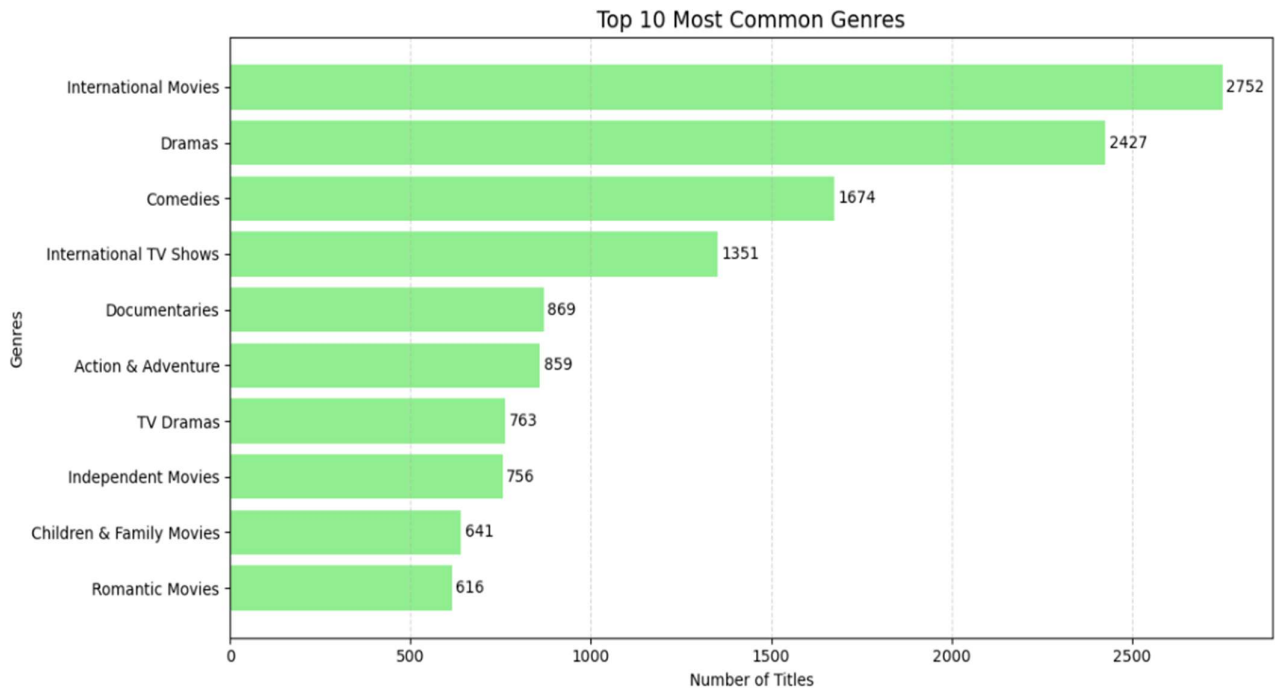
3. Geography:

- **Top-Contributing Countries:**

Netflix Content by Country and Type
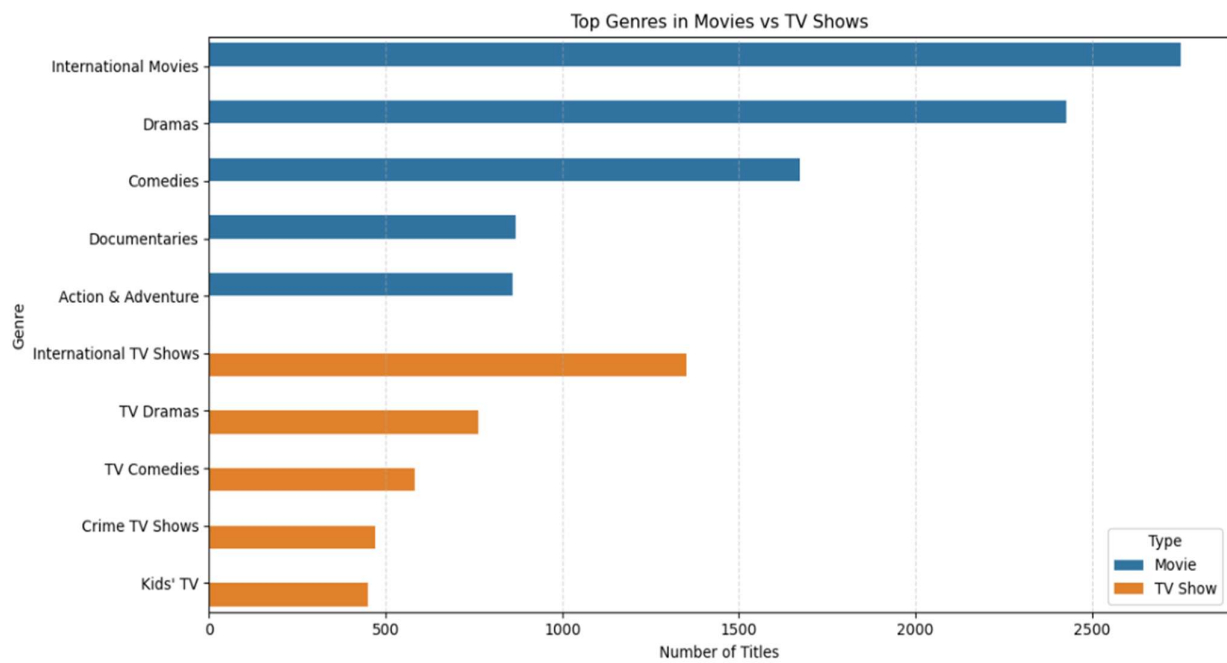
- **Indian Content:**

  - Second largest national share, skewed toward movies

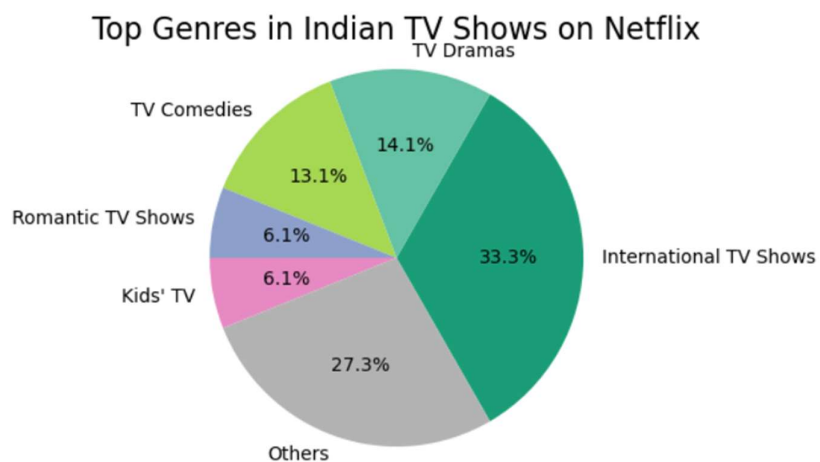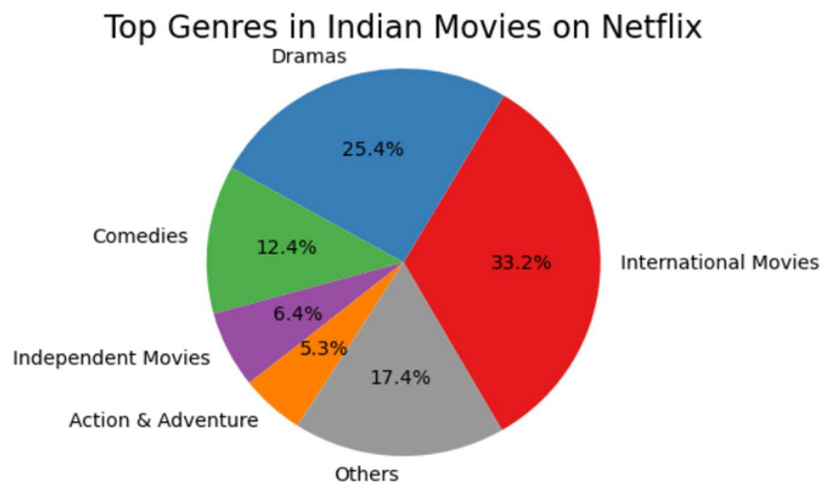 - Predominant genres: Dramas, Comedies, Action

4. Genres and Ratings:

- **Most Frequent Genres (Overall):**
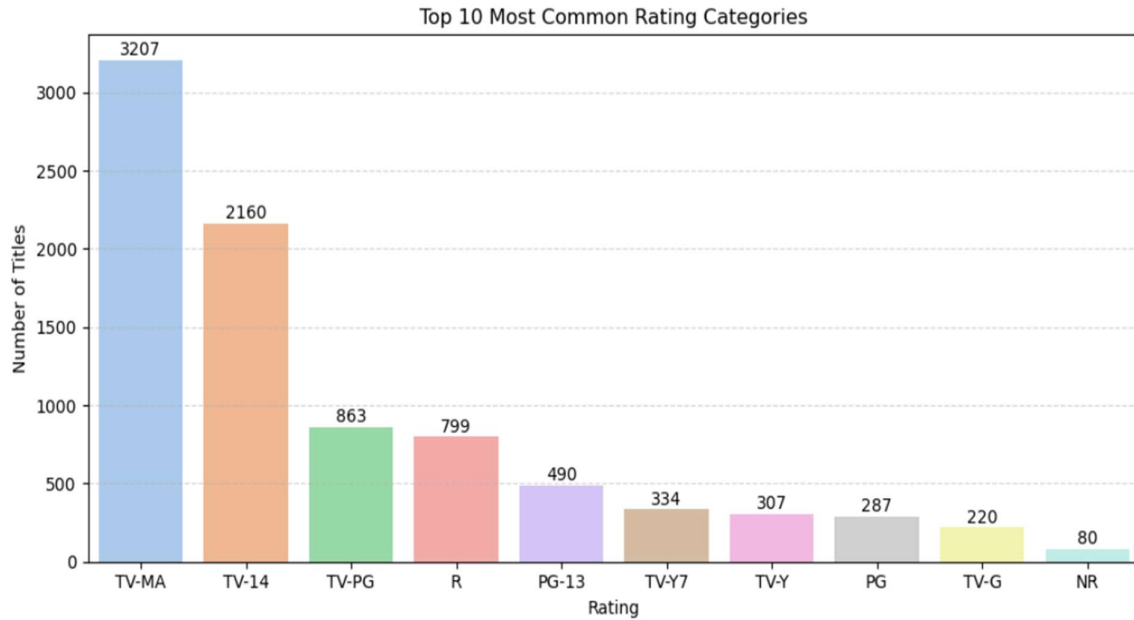

Top 10 Most Common Genres

- **Movies vs. TV Shows – Top Genres:**

Top Genres in Movies vs TV Shows

- **Indian Titles – Top Genres (TV Shows & Movies):**

Top Genres in Indian Movies on Netflix

Top Genres in Indian TV Shows on Netflix

- **Most Common Ratings:**

5. Notable People:

- **Top 5 Actors:**

  1. Anupam Kher (43 titles)
  2. Shah Rukh Khan (35 titles)
  3. Julie Tejwani (33 titles)
  4. Naseeruddin Shah (32 titles)
  5. Takahiro Sakurai (32 titles)

- **Top 5 Directors:**

  1. Rajiv Chilaka (22 titles)
  2. Jan Suter (21 titles)
  3. Raúl Campos (19 titles)
  4. Suhas Kadav (16 titles)
  5. Marcus Raboy (16 titles)

6. Data Quality Notes:

- **Unspecified Data:**
  - Country: 831 titles
  - Cast: 825 titles

Conclusions:

- The Netflix catalog is dominated by movies, with a strong secondary presence of TV shows.

- The United States and India lead in content production.

- Most entries are contemporary (post-2010), but Netflix has a multi-decade range.

- Genre diversity is high; mature audience content is prevalent.

- Bollywood talent and production are highly represented.

- Several entries still lack detailed country, director, or cast data.

Prepared by:

Prathamesh P. Ekbote