# Netflix Movies and TV Shows Analysis Report

## Project Overview

This project involves analyzing the Netflix dataset to gain insights into the distribution of Movies and TV Shows available on Netflix. We performed data cleaning, descriptive statistics, and visualizations to understand the release patterns, ratings, genre distribution, and more. Key questions include:

- What is the average release year for content on Netflix?
- What is the distribution of movies vs. TV shows?
- How do ratings and genres vary on the platform?
- Which countries have the most content on Netflix?
- What is the best time to release content?

---

## Data Overview

The dataset used for this analysis contains information about Netflix's content catalog, including details like the title, director, cast, country of origin, release year, rating, and genres.

### Columns in the Dataset

- `show_id`: Unique identifier for each show.
- `type`: Type of content (Movie or TV Show).
- `title`: Title of the content.
- `director`: Director of the content.
- `cast`: Cast involved in the content.
- `country`: Country where the content was produced.
- `date_added`: Date when the content was added to Netflix.
- `release_year`: Year when the content was released.
- `rating`: Age rating for the content.
- `genre`: The genres of the content (renamed from `listed_in`).

---

## Data Cleaning and Preparation

1. **Missing Values**:
   - We handled missing values by filling in:
     - Missing ratings with `TV-MA`.
     - Missing countries with `United States`.
2. **Column Renaming**:
   - The `listed_in` column was renamed to `genre` for clarity.

3. **Datetime Conversion**:
   - The `date_added` column was converted to datetime format, and missing values were handled.

---

# Analysis and Visualizations

## 1. Distribution of Release Years

We examined the distribution of content over the years and calculated the mean release year. The mean release year is **2013.93**.

**Visualization:**

- A horizontal bar plot was created to show the number of releases per year with a red dashed line indicating the mean number of releases per year.

## 2. Movies vs. TV Shows

We analyzed the breakdown of content types (Movies vs. TV Shows).

**Results:**

- **Movies** make up the majority of content, followed by **TV Shows**.

**Visualization:**

- A bar chart was created to compare the number of Movies vs. TV Shows on Netflix.

## 3. Content Types Percentage

A pie chart was generated to visualize the percentage distribution of Movies vs. TV Shows.

## 4. Rating Distribution

We explored how content is distributed across different Netflix ratings (e.g., TV-MA, PG, etc.).

**Visualization:**

- A count plot was generated to show the distribution of content by rating.

## 5. Top 5 Countries with the Most Content

We identified the top 5 countries that contribute the most content to Netflix.

**Visualization:**

- A pie chart showing the top 5 countries (by content count).

### 6. Country-wise Distribution of Movies and TV Shows

We separated content by type (Movies and TV Shows) and examined the top 5 countries contributing to Netflix.

**Visualization:**

- A bar graph showing the number of Movies and TV Shows for each of the top 5 countries.

### 7. Top 10 Genres

We identified the top 10 most frequent genres across all content.

**Visualization:**

- A bar plot showing the top 10 genres.

### 8. Best Time to Release Content (Heatmap)

We explored the best times (months and years) for releasing content on Netflix, using a heatmap.

**Visualization:**

- A heatmap was generated, showing the distribution of content added to Netflix over different months and years.

---

# Conclusions

- **Content Growth**: The number of releases on Netflix has grown steadily over the years, with a spike in recent years.
- **Movies vs. TV Shows**: Movies significantly outnumber TV Shows on Netflix.
- **Popular Genres**: The most popular genres on Netflix include "Dramas", "Comedies", and "Documentaries".
- **Top Countries**: The United States is the largest contributor of content on Netflix, followed by India and the United Kingdom.
- **Best Time to Release Content**: Based on the heatmap, the most common months for releasing content on Netflix are **October** and **December**.

---

# Future Recommendations

- **Content Strategy**: Netflix can focus on increasing its TV Show library to balance the ratio of Movies to TV Shows.
- **Genre Diversification**: There is an opportunity to add more niche genres to cater to diverse audiences.
- **Content Release Planning**: Based on the heatmap, Netflix should consider strategic release timings around high-demand months like October and December.

**Technologies Used**

- **Pandas**: For data manipulation and cleaning.
- **Matplotlib** and **Seaborn**: For data visualizations.
- **NumPy**: For numeric operations and calculations.