BỘ GIÁO DỤC VÀ ĐÀO TẠO

TRƯỜNG ĐẠI HỌC TÂN TẠO

Netflix Data Analysis Report

Tên: Nguyễn Duy Kha
MSSV: 2202082
Khoa: Khoa học máy tính
Giảng viên hướng dẫn: TS Nguyễn Xuân Hà

Table of contents

I. Introduction

1.1 Rationale for Topic Selection

In the context of increasingly digital content consumption, Netflix has become one of the world's largest content platforms. Analyzing the characteristics of Netflix's content helps us better understand production trends, distribution strategies, and viewer preferences in the digital media age.

1.2 Research Objectives

This report aims to analyze and visualize Netflix's content data after 2000 based on several specific dimensions:

1. Analyze the distribution of content by country of production.

2. Compare the quantity between two types of content: Movies and TV Shows.

3. Analyze the release trend of content over time.

4. Analyze the genre distribution of Movies and TV Shows.

1.3 Scope and Limitations

- Focus only on Netflix content released after 2000.

- Analyze the following fields: type, release_year, country, listed_in, rating.

- Exclude personal details (director, actors), ratings, or viewership data.

II. Theoretical Overview

2.1 Importance of Data Visualization

Data visualization transforms dry data tables into vivid charts and graphs, supporting faster and more effective decision-making.

2.2 Netflix Dataset

The dataset used in this report was sourced from Kaggle (Netflix Movies and TV Shows Dataset), containing over 8,800 records with various descriptive attributes such as title, release year, country, genre, and target audience.

III. Research Methodology

3.1 Data Processing Workflow

- Filter data after 2000.

- Handle missing values.

- Split and explode multi-valued fields such as country and genre.
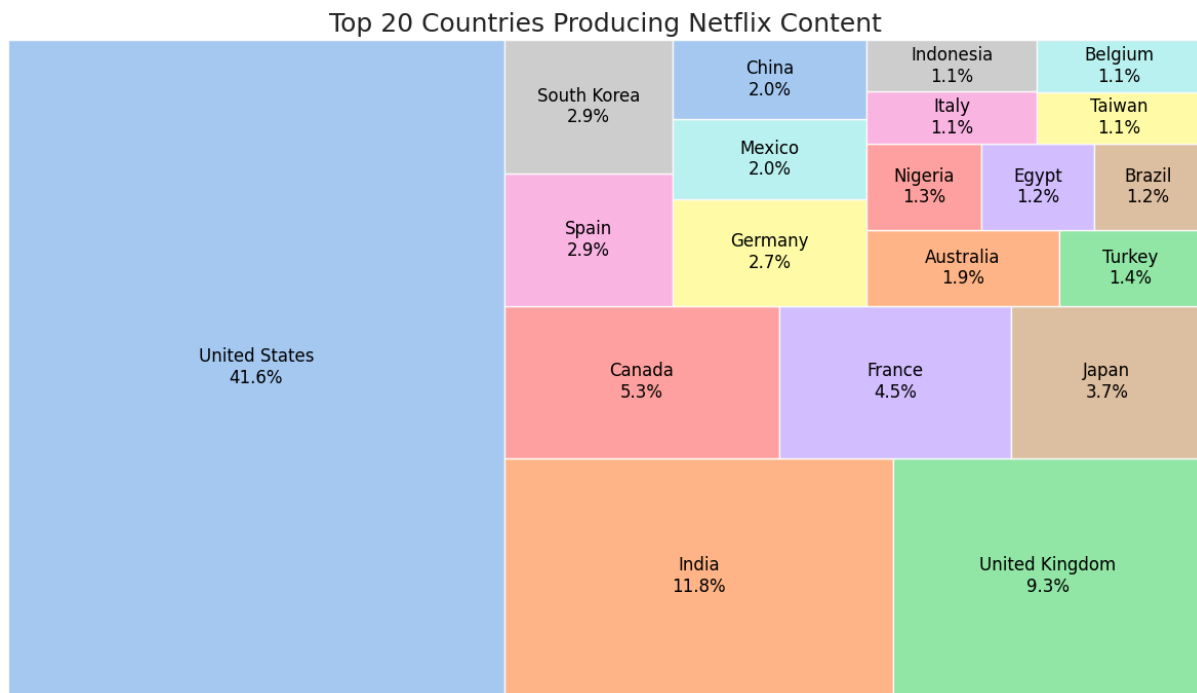
- Group and aggregate data.

3.2 Tools Used

- Python: pandas, matplotlib, seaborn, squarify

- Google Colab

3.3 Types of Visualizations Applied

- Line chart

- Bar chart (horizontal and vertical)

- Treemap

IV. Data Analysis

4.1 Content Distribution by Country of Production

Top 20 Countries Producing Netflix Content

The treemap shows that the United States holds the largest share among the top 20 countries. India, the United Kingdom, Canada, and South Korea are also key content providers. The spread reflects Netflix's globalization strategy in content offerings.

Conclusion: Netflix content is predominantly from the U.S., but countries like India, the U.K., and Canada are playing increasingly important roles in the global content library.

4.2 Comparison of Movies and TV Shows

Comparison of Number of Movies and TV Shows on Netflix
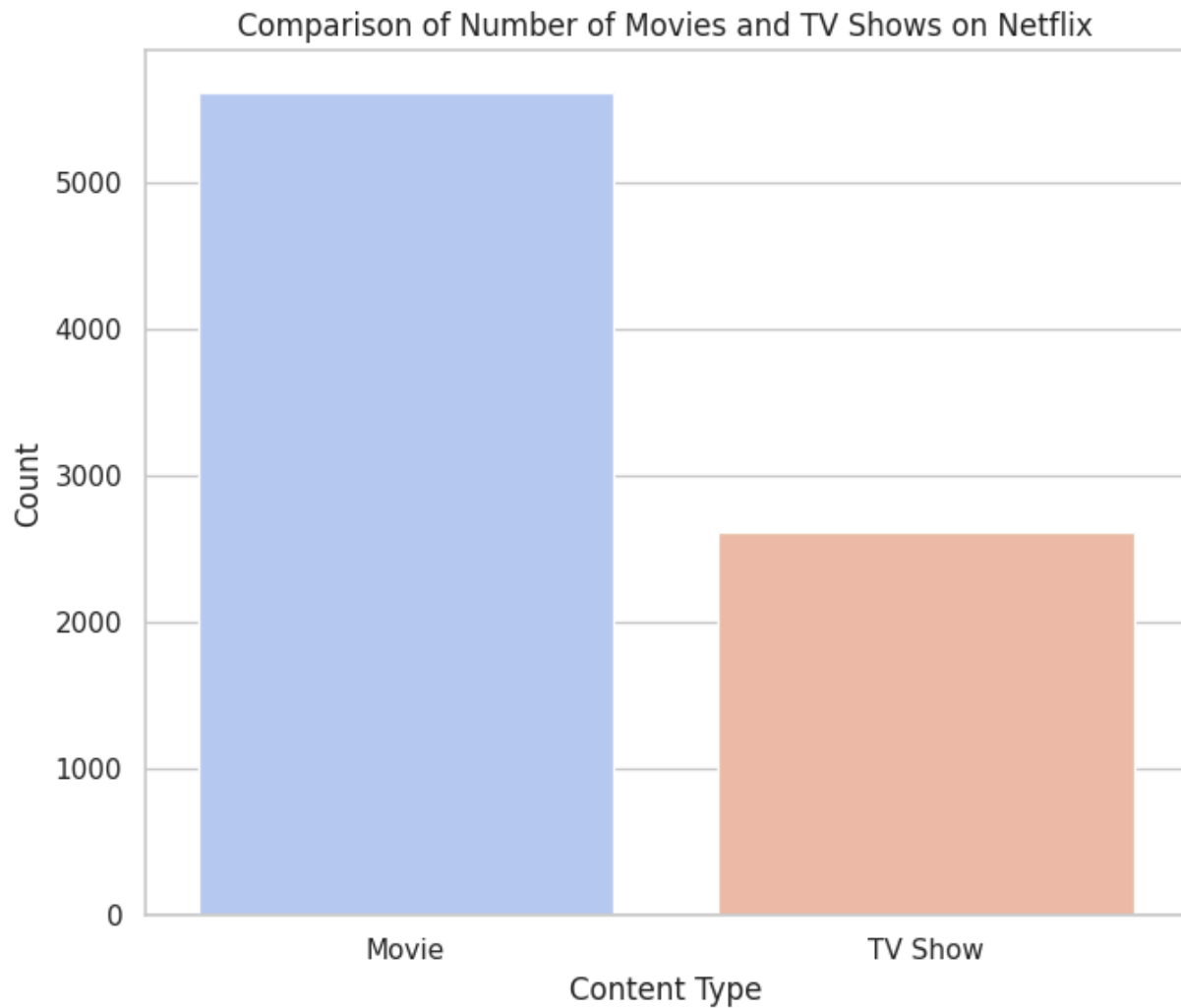
The bar chart indicates that Movies far outnumber TV Shows. However, TV Shows have been consistently maintained, reflecting a dual-investment strategy.

Conclusion: Netflix prioritizes releasing Movies, but still maintains a balanced strategy by investing in TV Shows to meet long-term viewer demand.

4.3 Release Trend Over Time

The number of movies and TV shows released from 2000 to 2021.
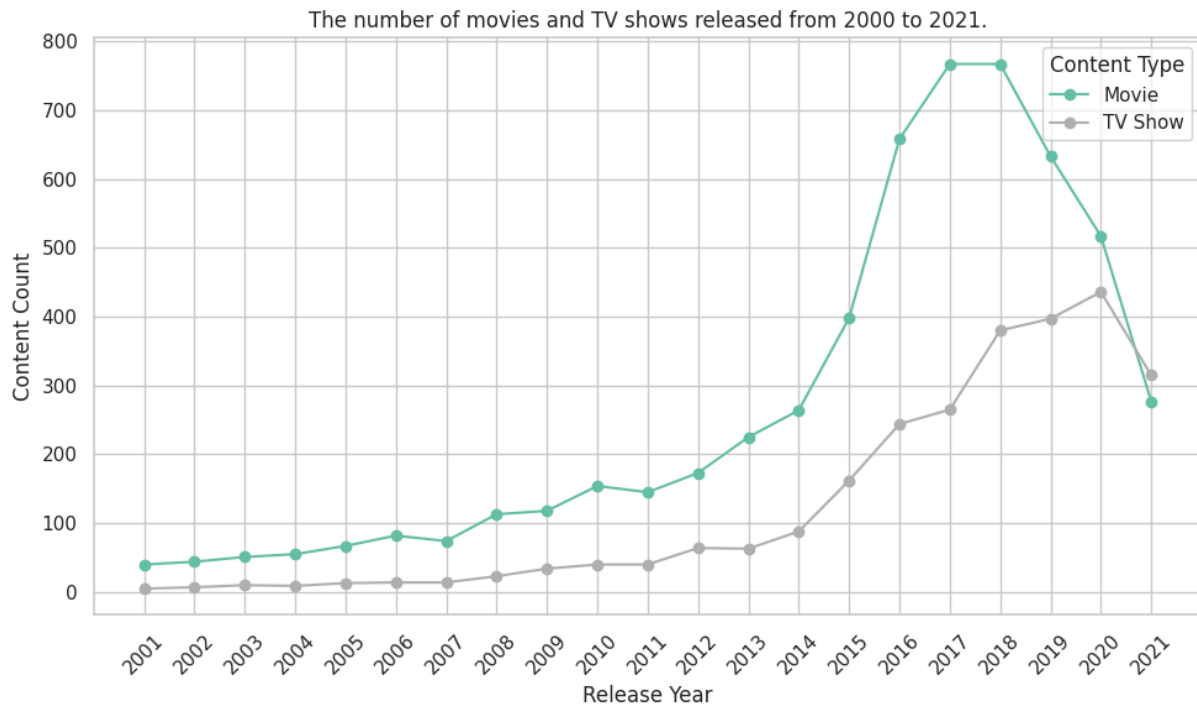
From 2015 onward, the number of Movies and TV Shows surged. The period from 2018 to 2020 marked a peak, reflecting a content production boom.

Conclusion: The strong growth post-2015 highlights Netflix's global market expansion and aggressive content production strategy.

4.4 Genre Distribution of Movies

Top 15 Most Common Genres in Movies on Netflix
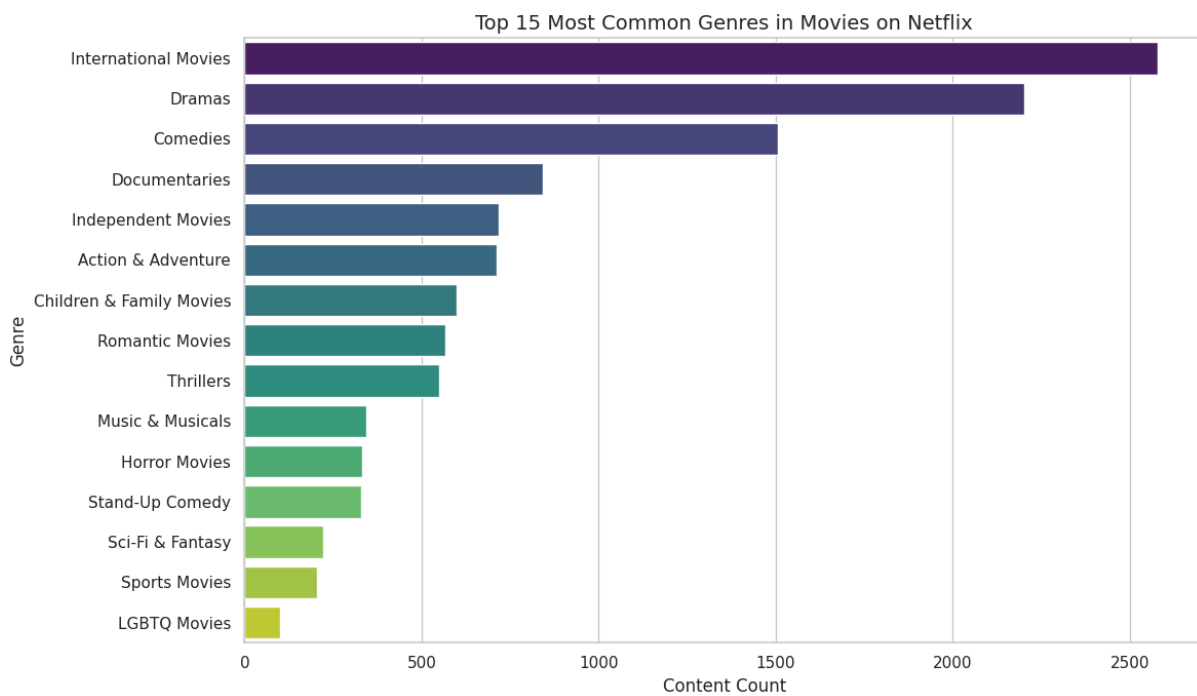
The top 15 most popular Movie genres are Drama, Documentary, Comedy, and Action, reflecting a diverse audience.

Conclusion: The most invested genres are mainstream and accessible, indicating a commercially driven content strategy.

4.5 Genre Distribution of TV Shows

Top 15 Most Common Genres in TV Shows on Netflix
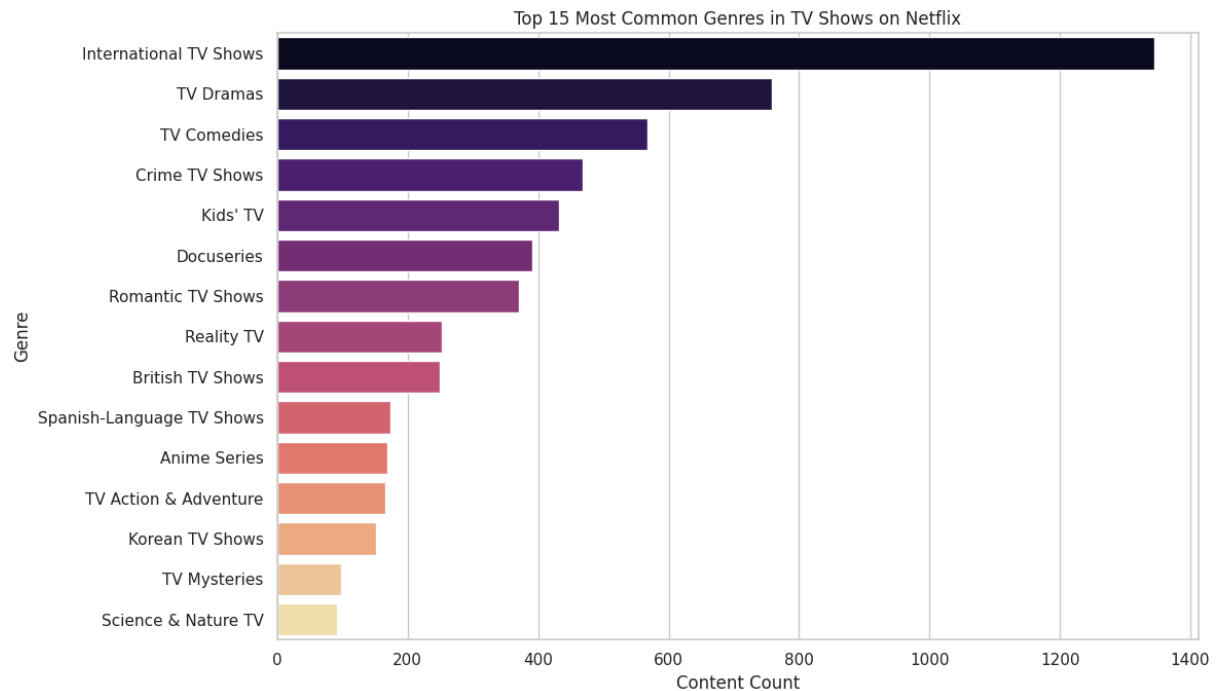
Popular TV Show genres include International TV Shows, Crime, Kids, and Reality. This trend caters to a wide range of international audiences.

Conclusion: Netflix's television content strategy clearly emphasizes global diversity and genre variety to reach a broad audience.

V. Conclusion and Recommendations

5.1 Conclusion

- Netflix's content grew rapidly after 2015.

- Movies outnumber TV Shows, but TV Shows maintain a consistent presence.

- The U.S., India, U.K., and Canada lead in content production.

- Drama and Documentary are dominant genres.

5.2 Recommendations

- Invest more in children's content and science fiction.

- Expand content production in Asian and African countries.

- Diversify ratings to better serve family audiences.