

Netflix Data Analysis – Final Report

Introduction

This project performs a comprehensive **Exploratory Data Analysis (EDA)** on the official Netflix **Movies and TV Shows dataset**.

Using **Python**, **Pandas**, **Matplotlib**, and **Seaborn**, we uncover trends in content growth, popular genres, production countries, and movie durations.

The primary goal is to transform raw data into clear insights about **Netflix's global content strategy and evolution over time**.

Objectives

1. Analyze how many titles were added to Netflix per year.
2. Identify the top countries producing Netflix content.
3. Explore the most frequent and popular genres.
4. Examine the distribution of movie durations.
5. Determine which months saw the highest number of new titles.
6. Compare the ratio of movies to TV shows.
7. Investigate the relationship between **release year** and **genre** using a heatmap.

Dataset Overview

Source

Netflix Movies and TV Shows Dataset

Available on Kaggle

Key Information

- **Records:** 8,807
- **Main columns:**
 - type – Movie or TV Show
 - title – Title name
 - country – Country of origin
 - date_added – Date added to Netflix
 - release_year – Year of production

- duration – Duration in minutes or number of seasons
- listed_in – Genres

Data Cleaning Process

- Missing values (NaN) were replaced with "Unknown" or filled using the mode (for categorical data).
- Date strings were converted to datetime format using pd.to_datetime.
- Movie durations were converted from "90 min" strings to numeric format (float).
- Non-movie duration values like “2 Seasons” were filtered out.
- Columns with invalid or missing critical information were dropped.

Exploratory Data Analysis Results

Content Growth by Year

Insight:

Netflix's content library expanded rapidly between **2016 and 2020**.

Movie releases peaked in 2018–2019, while TV shows grew steadily after 2016.

Top 10 Countries with Most Productions on Netflix

Overview

This bar chart displays the **top 10 countries** that have produced the highest number of titles available on Netflix's platform.

Each bar represents the **total count of movies and TV shows** originating from that country.

Insights

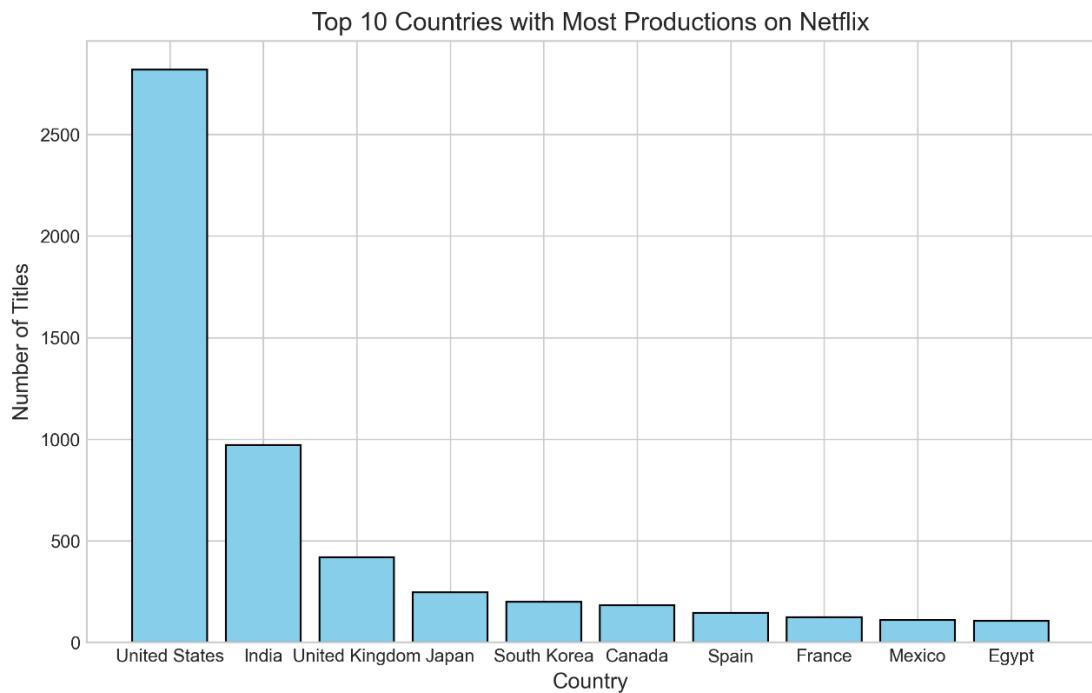
1. **The United States** dominates Netflix's content catalog, contributing **over 2,700 titles**, which is more than twice as many as any other country.
2. **India** ranks second, showcasing Netflix's deep investment in the South Asian market and the growing demand for Bollywood and regional Indian productions.
3. **The United Kingdom** and **Japan** follow closely, reflecting the strength of English-language and East Asian entertainment.
4. **South Korea**'s increasing presence aligns with the global popularity of K-dramas and Korean cinema in recent years.
5. Other countries such as **Canada, Spain, France, Mexico, and Egypt** also maintain a steady contribution — showing Netflix's efforts toward global content diversity.

Interpretation

- The data highlights Netflix's **U.S.-centric origins**, but also its **strategic global expansion**.
- The company's push into **Asian and European markets** is evident, aligning with its goal of becoming a truly international platform.
- This trend reflects the **rise of regional storytelling** — Netflix is increasingly investing in non-English productions to appeal to global audiences.

Key Takeaway

Netflix's catalog is still dominated by U.S. content, but international markets — especially **India, Japan, and South Korea** — are rapidly growing and reshaping the platform's cultural balance.



Top 10 Most Popular Genres on Netflix

Overview

This horizontal bar chart visualizes the **top 10 most popular genres** available on Netflix. Each bar represents the total number of titles (movies or TV shows) that fall under a specific genre or a combination of genres.

Insights

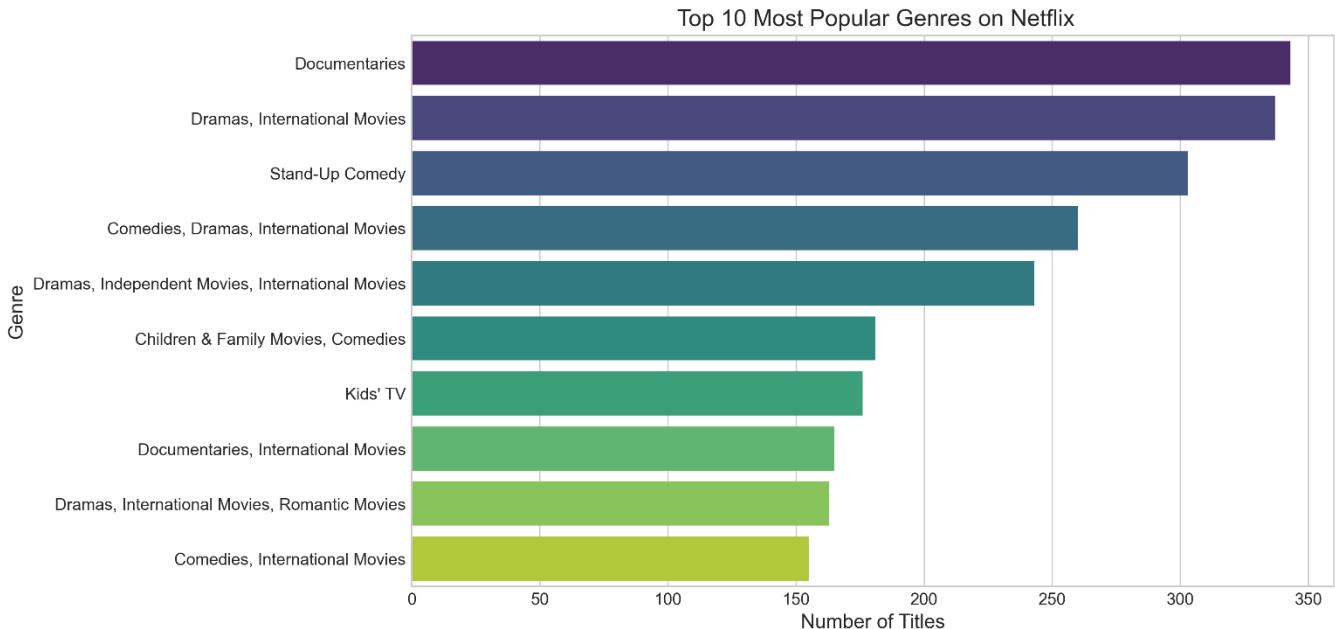
1. **Documentaries** rank first, showing Netflix's strong focus on real-world stories, social issues, and biographies.
2. **Dramas and International Movies** come next, confirming the audience's high demand for emotionally engaging and diverse storytelling.
3. **Stand-Up Comedy** has a major presence, reflecting Netflix's dominance in comedy specials and global stand-up shows.
4. The mix of genres such as *Comedies + Dramas + International Movies* shows Netflix's strategy of **cross-genre storytelling** — combining humor, emotion, and culture.
5. **Children & Family Movies** and **Kids' TV** highlight Netflix's investment in younger audiences.
6. The presence of *Romantic*, *Independent*, and *Documentary* categories indicates Netflix's aim to cater to a wide range of tastes.

Interpretation

- Netflix's strongest content pillars are **Documentaries, Drama, and Comedy**.
- The combination of **international and cross-genre titles** illustrates Netflix's global approach — producing content that resonates with viewers from different cultures.
- This pattern also suggests Netflix is not relying only on Hollywood-style entertainment but expanding into **localized genres**.

Key Takeaway

Netflix's genre diversity is one of its biggest competitive advantages — blending global storytelling, educational content, and entertainment across cultures.



Distribution of Movie Lengths on Netflix

Overview

This histogram visualizes the distribution of movie durations on Netflix.

Each bar represents the number of movies within a specific time range (in minutes), helping to identify common runtime patterns.

Insights

- The majority of Netflix movies have durations between 80 and 120 minutes, forming a clear central peak.**
- Very few movies are shorter than 60 minutes or longer than 150 minutes, indicating that feature-length films dominate Netflix's catalog.**
- The distribution shows a right-skewed pattern, meaning there are a small number of long films (over 150 minutes) but most titles cluster around the 1.5-hour mark.**
- This pattern aligns with traditional cinema standards, as 90-minute films are typically considered ideal for maintaining audience engagement.**

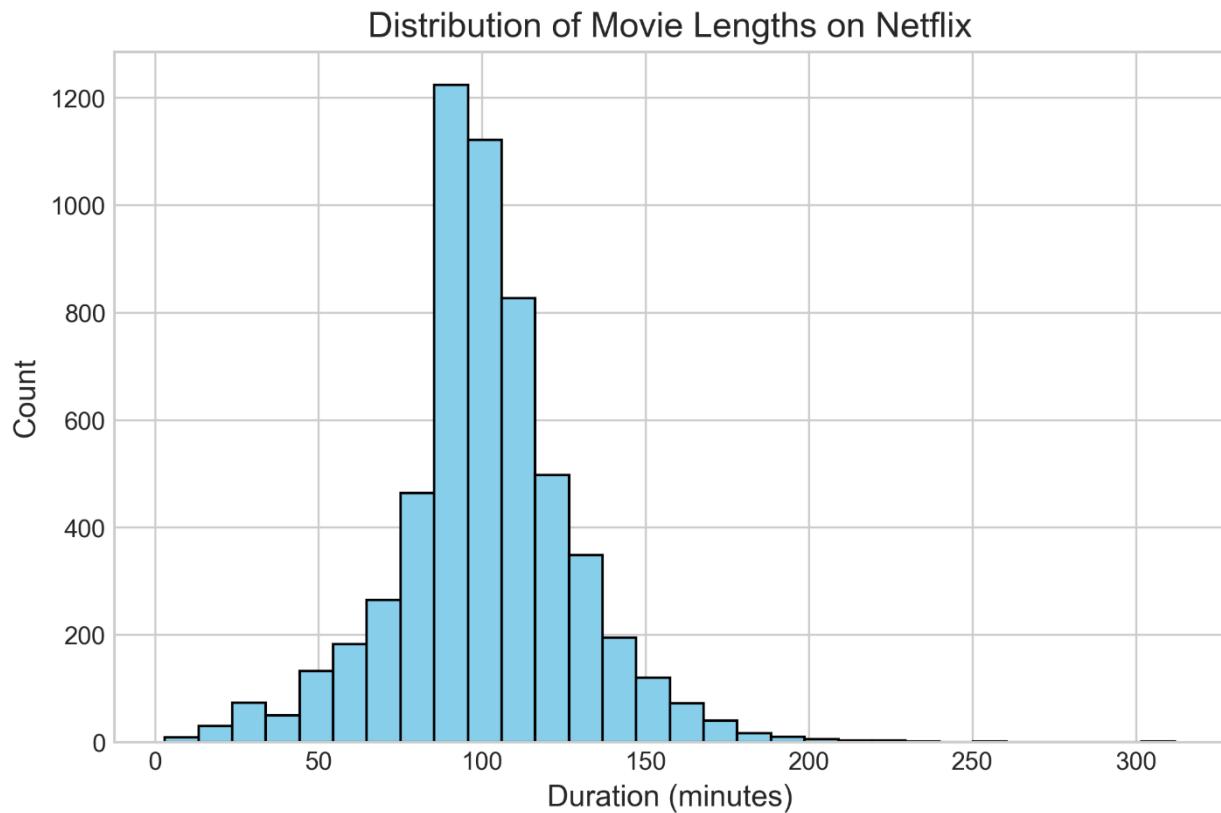
Interpretation

- The concentration around 90–100 minutes suggests Netflix prioritizes films that balance storytelling depth with watchability.**
- Longer movies (>150 minutes) are relatively rare, likely reserved for epic dramas or special releases.**

- The shape of the curve indicates Netflix's catalog favors concise, binge-friendly content, ideal for digital streaming audiences.

Key Takeaway

Most Netflix movies are designed to be feature-length, typically around 90–100 minutes, balancing audience attention span and cinematic quality.



Number of Titles Added per Month on Netflix

Overview

This bar chart illustrates the monthly distribution of new titles added to Netflix. Each bar represents the number of movies and TV shows released or added to the platform in a given month.

Insights

1. The beginning and end of the year — particularly January, December, and October — show the highest number of new additions.
2. February and May record the lowest activity, suggesting seasonal slowdowns in content release.

3. The gradual decline from January to mid-year indicates a strategic release pattern, likely tied to audience engagement cycles and marketing campaigns.
4. The overall difference between months isn't extreme, implying Netflix maintains a steady content flow throughout the year.

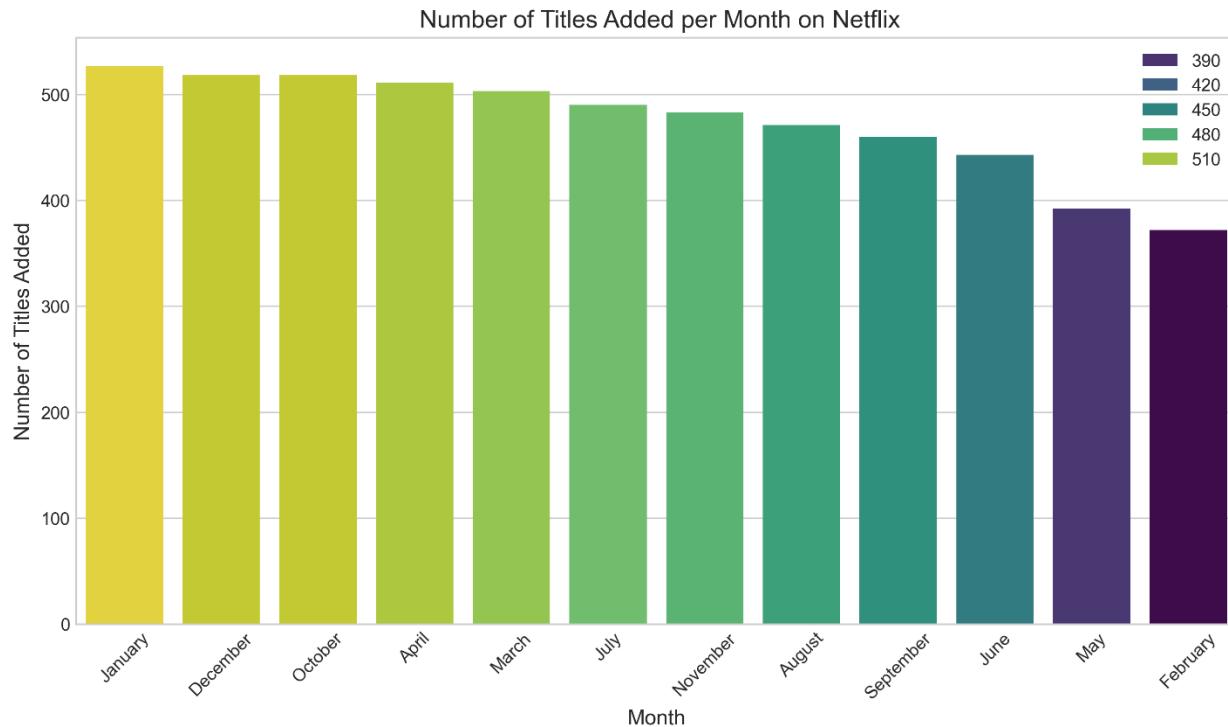
Interpretation

- Netflix tends to release major titles during holiday and winter months, when streaming demand peaks.
- Lower additions in mid-year months may coincide with production cycles or fewer global events.
- The steady trend across all months reflects Netflix's commitment to maintaining user engagement year-round, avoiding long content gaps.

Key Takeaway

Netflix strategically increases content additions during holiday-heavy months (like January and December)

to maximize viewer engagement — while maintaining a consistent stream of new releases throughout the year.



Distribution of Content Types on Netflix

Overview

This pie chart shows the **distribution of content types** available on Netflix — specifically comparing **Movies** and **TV Shows**.

Insights

1. **Movies dominate** the platform, making up approximately **70%** of all titles.
2. **TV Shows** represent around **30%** of the catalog, highlighting Netflix's gradual but steady expansion into serialized content.
3. The chart clearly indicates Netflix's **movie-first library**, even though many of its original productions (like *Stranger Things* or *The Witcher*) are shows.

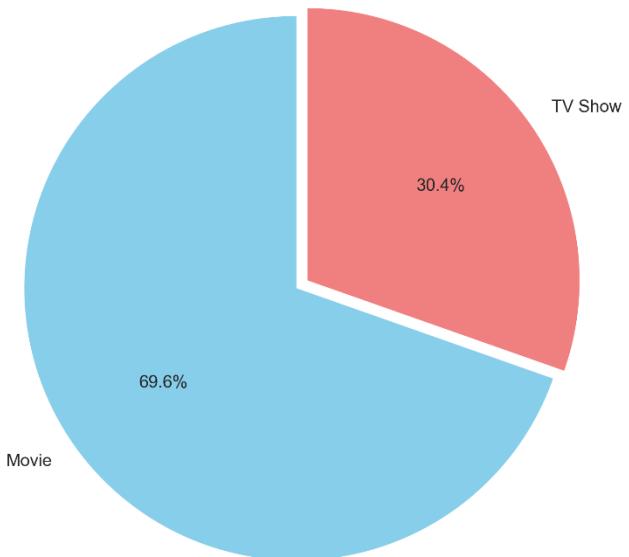
Interpretation

- The dominance of movies suggests that **Netflix's foundation was built on film licensing and global movie distribution**.
- However, the 30% share of TV shows demonstrates a **strategic diversification** toward longer, binge-worthy series — a trend that keeps viewers subscribed for extended periods.
- This balance between **short-form (movies)** and **long-form (series)** content helps Netflix target both casual viewers and dedicated binge-watchers.

Key Takeaway

Approximately **70% Movies** and **30% TV Shows** —
Netflix continues to lead with films while steadily expanding its original series catalog to strengthen viewer retention.

Distribution of Content Types on Netflix



Relationship Between Release Year and Genre on Netflix

Overview

This heatmap visualizes the **relationship between movie release years and genres** available on Netflix. Each cell represents the **number of titles** released in a given year for a specific genre — with warmer colors (red/orange) indicating higher counts.

Insights

1. There is a **significant concentration of titles from 2010 onward**, showing Netflix's rapid growth during the era of digital streaming.
2. **Documentaries, Dramas, and International Movies** are the most consistently represented genres across recent years.
3. A clear **rise in diversity** of genres can be observed post-2015 — aligning with Netflix's global expansion strategy.
4. Older decades (before 2000) show very few titles, which confirms that **Netflix's catalog primarily focuses on modern content** rather than archival or classic films.

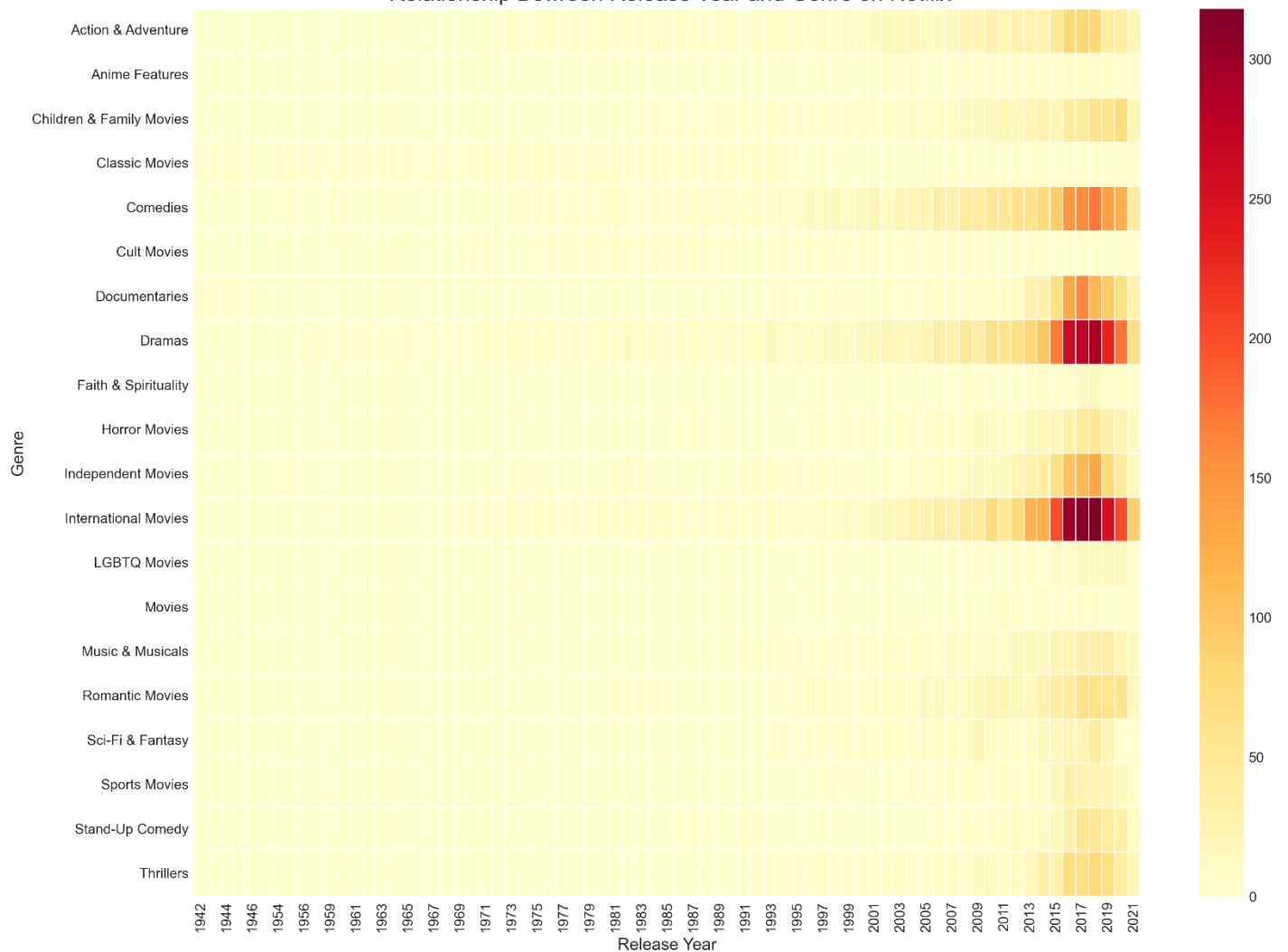
Interpretation

- The strong presence of **Documentaries** reflects Netflix's investment in **non-fiction and social storytelling** since 2016.
- The rapid increase in **International Movies** corresponds to the company's entry into new markets such as India, Korea, and Europe.
- The heatmap reveals how Netflix has transitioned from a **U.S.-centric platform to a global content provider**, adapting to diverse cultural demands.

Key Takeaway

Most Netflix content has been released **after 2010**, with dominant genres such as **Documentaries, Dramas, and International Movies** driving the platform's global growth.

Relationship Between Release Year and Genre on Netflix



Summary of Findings

Category	Key Insight
Content Type	~70% Movies, ~30% TV Shows
Top Producers	USA, India, UK, Japan, South Korea
Growth Period	Explosive increase between 2016–2019
Average Duration	90–100 minutes typical movie length
Popular Genres	Documentaries, Dramas, Comedies
Peak Months	January and December
Globalization Trend	Surge in international titles post-2015

Tools & Technologies

Tool	Purpose
Python	Main programming language
Pandas	Data cleaning and manipulation
NumPy	Numerical processing
Matplotlib	Visualization and charting
Seaborn	Advanced, aesthetic plots

Conclusions

1. Netflix experienced **rapid global growth** between 2016–2019.
2. The platform's catalog is **movie-heavy**, but the share of **TV shows** continues to rise.
3. **USA and India** dominate production, followed by several European and Asian countries.
4. Most movies fall within a **90–100 minute** duration range.
5. Netflix's content strategy is clearly expanding toward **international and diverse genres**.

Author

Navid Bordbar
Data Analyst

Email: novid.bordbar.ds@gmail.com
GitHub: github.com/navidbordbar-data