Netflix Content Analytics Dashboard

Interactive Descriptive Analytics Using Data Visualisation

Course Code: DS451
Course Title: Data Visualization

Project Report

Submitted By

Hamza Motiwala (2022380)
Umer Sami (2022640)

Contents

List of Figures

1. Introduction

Digital streaming platforms such as Netflix generate vast amounts of content-related data that reflect trends in production, regional focus, genre popularity, and temporal growth. Analysing this data using static tools limits insight discovery and does not support exploratory analysis.

Data visualization enables analysts to transform raw datasets into meaningful visual representations that highlight patterns, relationships, and trends. Interactive dashboards further enhance this process by allowing users to dynamically filter and explore data from multiple perspectives.

This project presents an interactive Netflix Content Analytics Dashboard designed to perform descriptive analytics on Netflix's content catalogue. The dashboard integrates multiple visualisation techniques to support exploratory analysis related to time, geography, genre, ratings, and content type.

2. Project Objectives

The primary objective of this project is to design and implement an interactive dashboard that enables descriptive analysis of Netflix content data.

Specific objectives include:

- Analysing Netflix content trends across time and geography
- Comparing content types, genres, and ratings
- Visualising hierarchical and relational structures in the data
- Enabling real-time filtering and exploration
- Applying data visualisation principles to support insight generation

3. Dataset Description and Preprocessing

The dataset used in this project is the Netflix Titles Dataset, which contains metadata about movies and television shows available on Netflix. Key attributes include title name, content type, country of origin, listed genres, ratings, release year, date added to Netflix, and duration.

Data preprocessing involved handling missing values, extracting primary country and genre fields, converting dates into usable temporal attributes, and deriving additional features such as year added and duration metrics. These steps were essential to ensure consistency and analytical accuracy across all visualisations.

4. System Architecture and Workflow

The system follows a modular data-to-visualisation pipeline. Raw CSV data is processed using Python scripts to clean and engineer features. The processed dataset is then loaded into a Streamlit-based web application.

User inputs such as dropdowns and sliders dynamically control filtering logic. Plotly is used to generate interactive visualisations that update instantly in response to filter changes. The dashboard is rendered as a web-based interface accessible locally or via deployment.

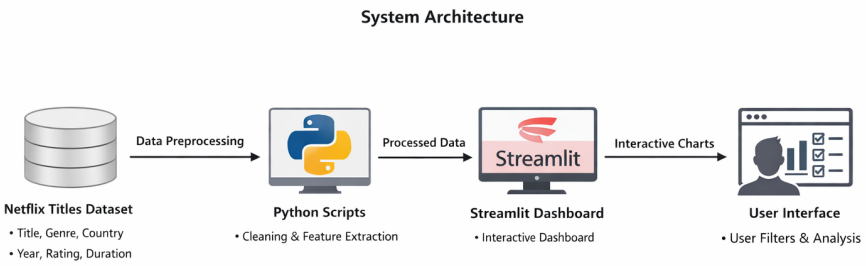

Figure 1: High-Level System Architecture

5. Dashboard Design Philosophy

The dashboard design prioritises clarity, usability, and analytical flow. Filters are placed in a sidebar to allow users to control data selection without cluttering the main visual area. Visualisations are grouped into logical sections based on analytical purpose.

A dark theme inspired by Netflix branding is used to reduce visual fatigue and highlight important data points. Smooth transitions and animations improve user experience while maintaining analytical focus.

6. Key Metric Overview

The dashboard begins with a set of key metric cards that provide a high-level summary of the filtered dataset. These metrics include the total number of titles, breakdown by content type, titles added in the selected year, and a featured title highlight.

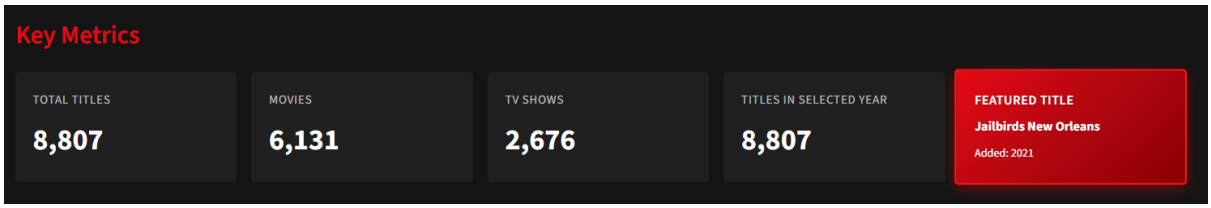

Figure 2: Key Metric Summary Cards

These metrics allow users to quickly understand the scale and composition of the selected data before exploring detailed visualisations.

7. Content Flow and Hierarchical Analysis

7.1 Genre to Country Content Flow

The Sankey diagram visualises relationships between content genres and producing countries. The width of each flow represents the number of titles, allowing strong associations to be identified immediately.

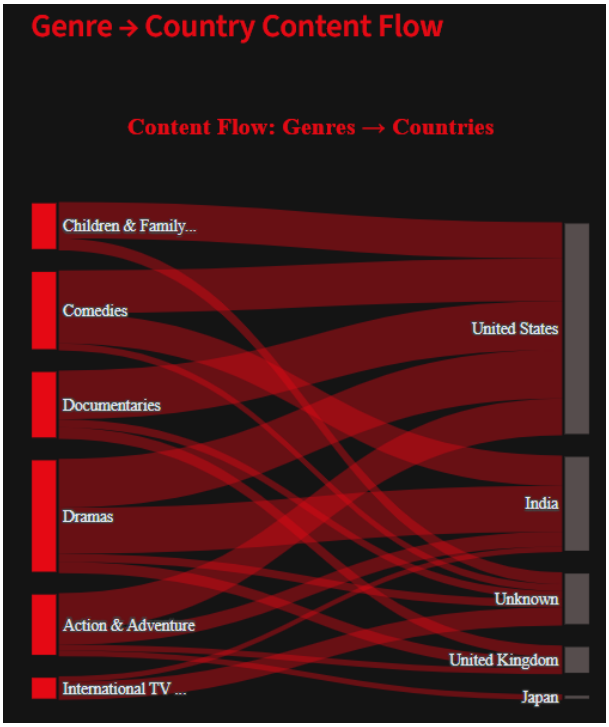

Figure 3: Genre to Country Content Flow

7.2 Hierarchical Content Structure

A treemap is used to represent hierarchical relationships between countries, genres, and content types. Larger rectangles indicate dominant categories, while colour intensity reinforces content volume.

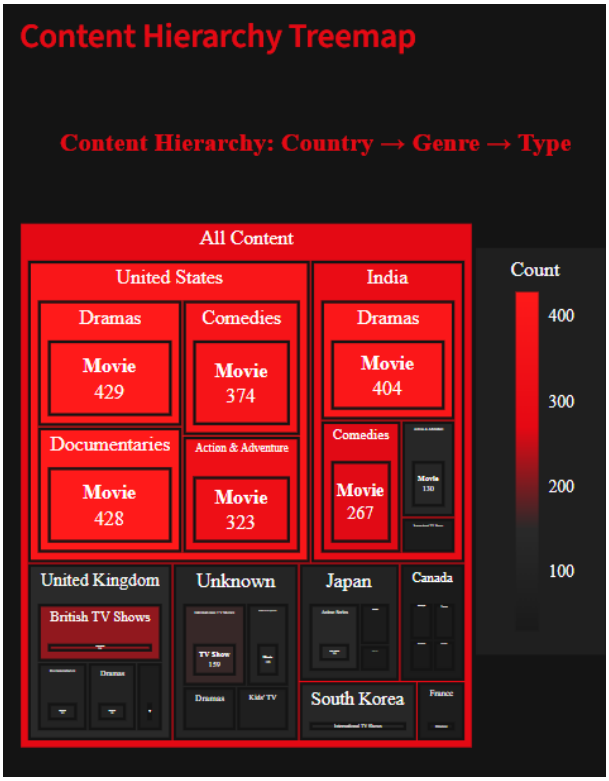

Figure 4: Country to Genre to Type Hierarchy

7.3 Global Content Distribution

A choropleth map displays the global distribution of Netflix content. Countries are coloured based on total title count, providing geographic context to production patterns.

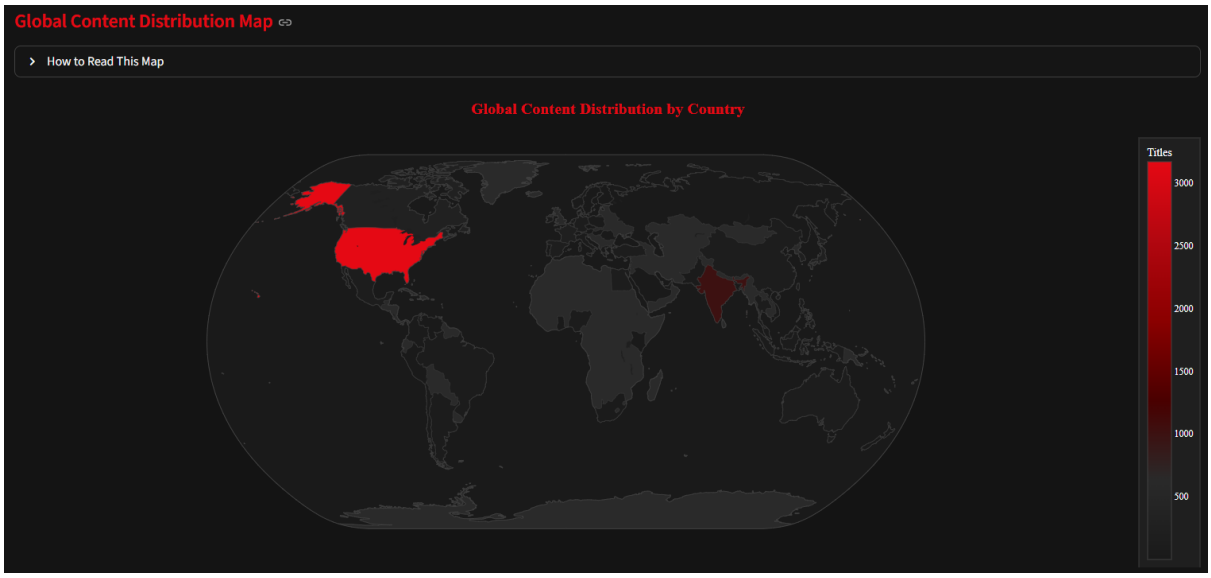

Figure 5: Global Distribution of Netflix Content

8. Temporal Trends and Evolution

8.1 Yearly Content Addition

A line chart illustrates how Netflix content additions have evolved over time. Peaks indicate periods of rapid expansion, while declines suggest strategic shifts.

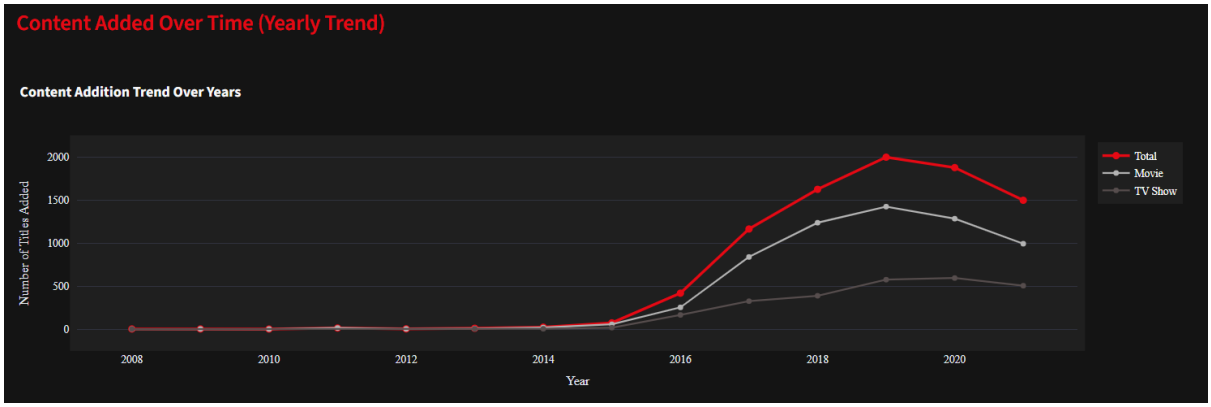

Figure 6: Content Added Over Time (Yearly)

8.2 Monthly Content Patterns

A bar chart shows the number of titles added per month, revealing seasonal trends in content releases.

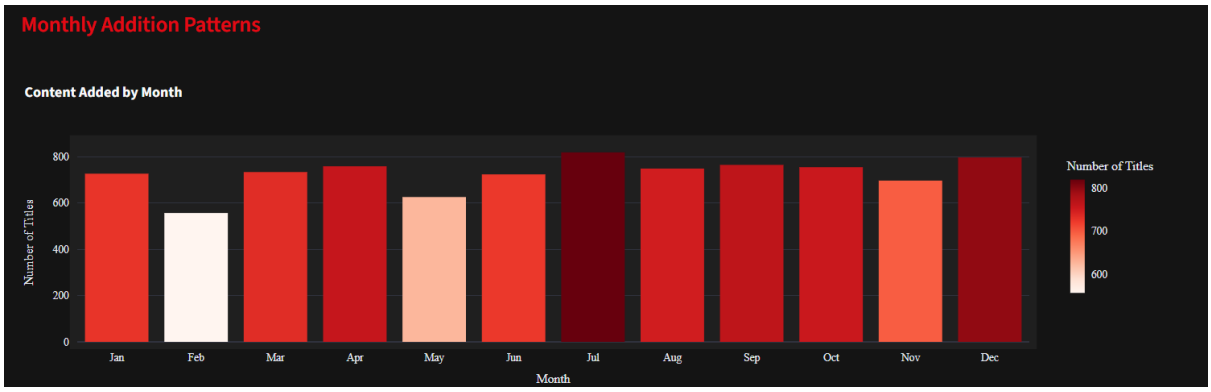

Figure 7: Monthly Content Addition Patterns

9. Comparative Analysis

9.1 Movies vs TV Shows

A pie chart compares the proportion of movies and TV shows in the dataset, providing insight into Netflix's content strategy.

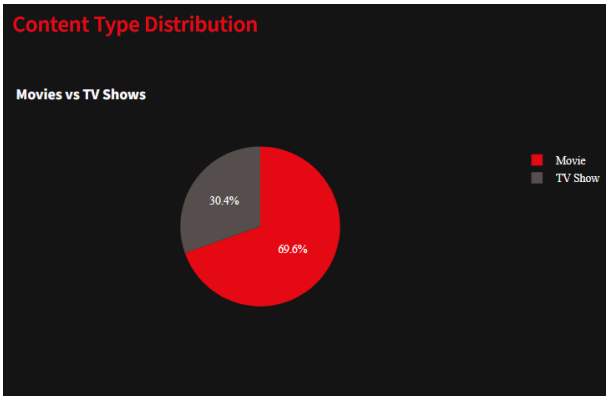

Figure 8: Movies vs TV Shows Distribution

9.2 Top Producing Countries

A horizontal bar chart ranks the top countries by content production.

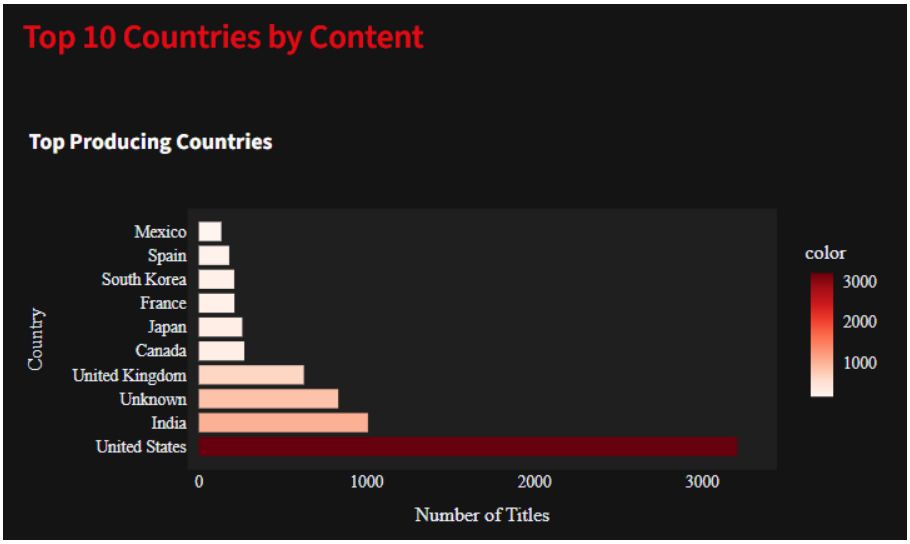

Figure 9: Top Producing Countries

9.3 Top Genres

A bar chart displays the most common genres in the Netflix catalogue.

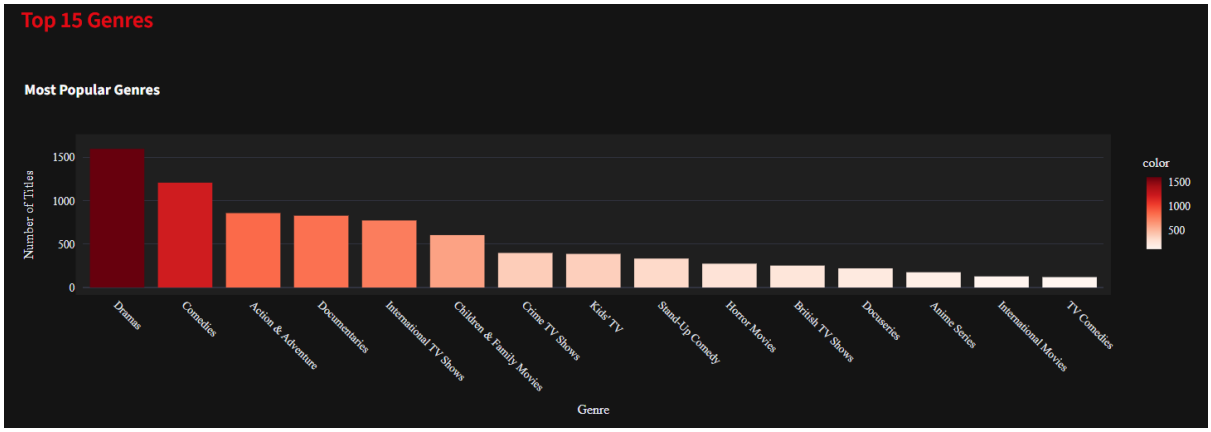

Figure 10: Top Genres by Content Volume

9.4 Content Rating Distribution

A rating-based bar chart highlights the distribution of content maturity levels.

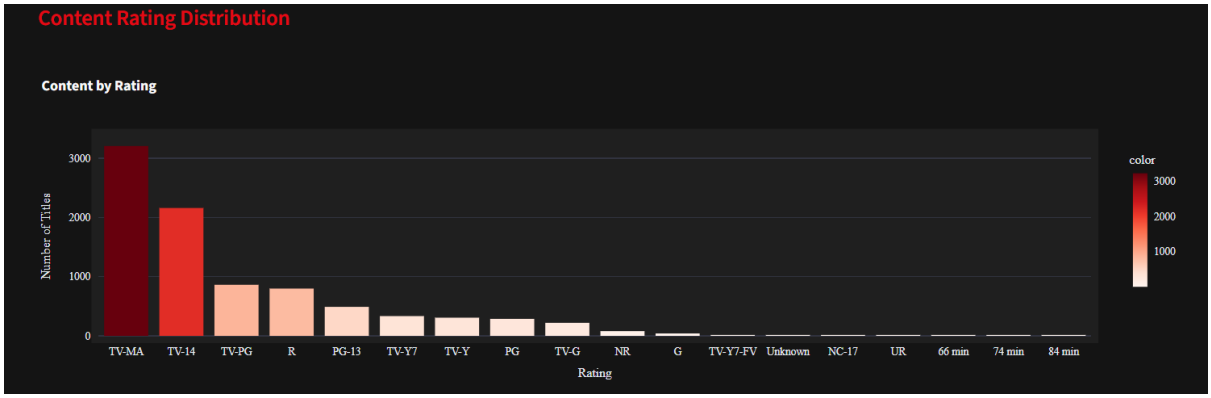

Figure 11: Content Rating Distribution

10. Pattern Analysis

10.1 Genre Evolution Over Time

A multi-line chart visualises how the popularity of top genres has evolved over time.

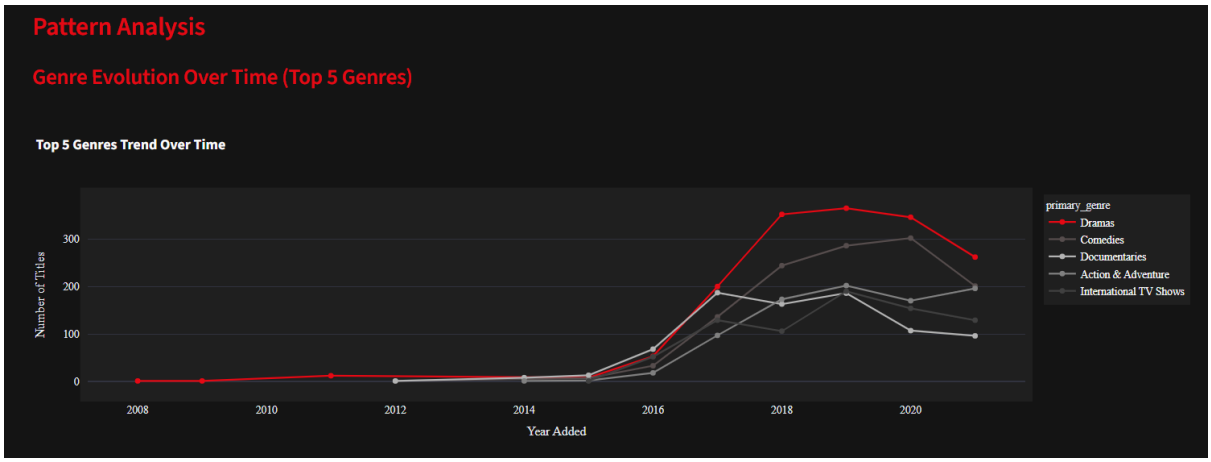

Figure 12: Genre Evolution Over Time

10.2 Content Type by Country

A grouped bar chart compares movie and TV show production across top countries.

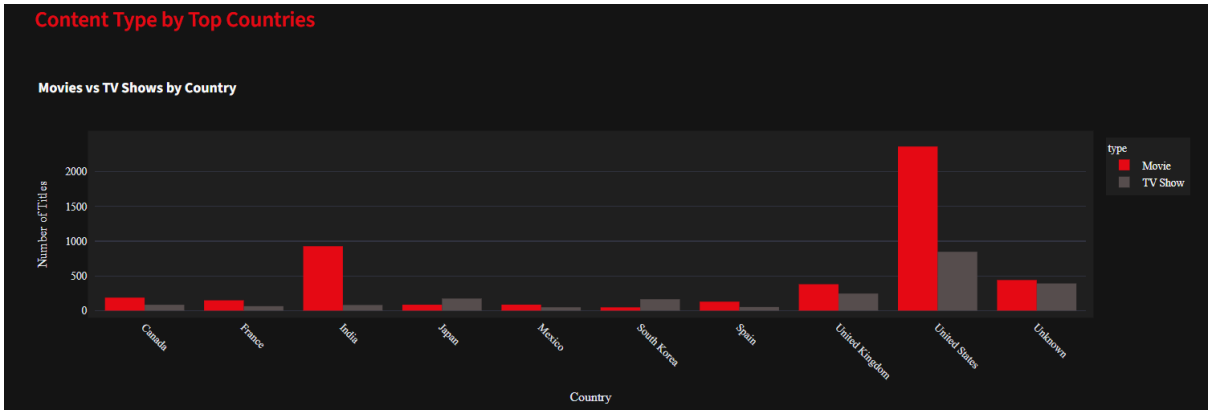

Figure 13: Movies vs TV Shows by Country

10.3 Movie Duration Distribution

A histogram displays the distribution of movie durations, revealing typical content lengths.

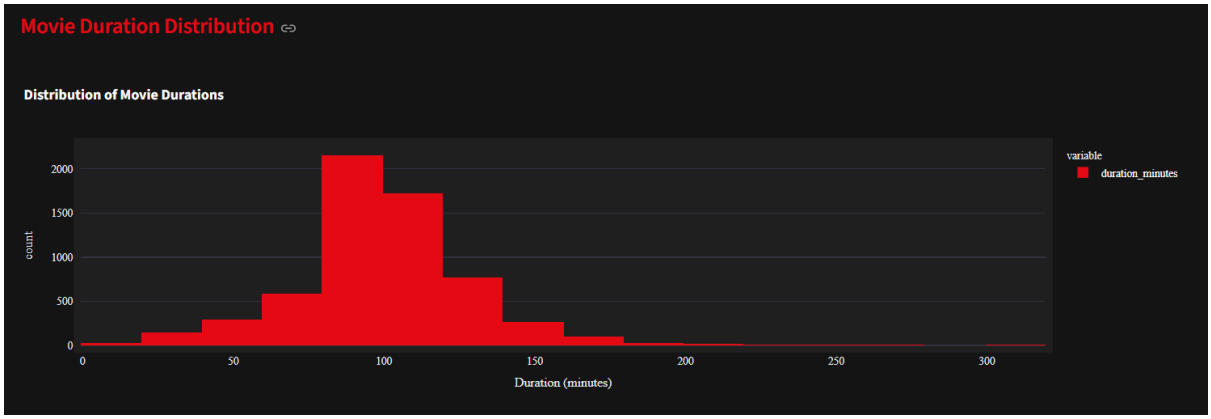

Figure 14: Movie Duration Distribution

11. Interactive Data Table

The dashboard includes an interactive data table that displays the filtered dataset. Users can scroll, sort, and inspect individual records to validate insights derived from visualisations.

title	type	Country	Genre	rating	Release Year	Year Added	duration
Dick Johnson Is Dead	Movie	United States	Documentaries	PG-13	2020	2021	90 min
Blood & Water	TV Show	South Africa	International TV Shows	TV-MA	2021	2021	2 Seasons
Ganglands	TV Show	Unknown	Crime TV Shows	TV-MA	2021	2021	1 Season
Jailbirds New Orleans	TV Show	Unknown	Docuseries	TV-MA	2021	2021	1 Season
Kota Factory	TV Show	India	International TV Shows	TV-MA	2021	2021	2 Seasons
Midnight Mass	TV Show	Unknown	TV Dramas	TV-MA	2021	2021	1 Season
My Little Pony: A New Generation	Movie	Unknown	Children & Family Movies	PG	2021	2021	91 min
Sankofa	Movie	United States	Dramas	TV-MA	1993	2021	125 min
The Great British Baking Show	TV Show	United Kingdom	British TV Shows	TV-14	2021	2021	9 Seasons
The Starling	Movie	United States	Comedies	PG-13	2021	2021	104 min

Figure 15: Filtered Dataset View

12. Key Insights and Observations

Analysis reveals strong geographic concentration of content, dominance of certain genres, and significant temporal growth trends. Movies constitute a larger portion of the catalogue, while TV shows show sustained growth in recent years.

13. Limitations and Future Improvements

The project focuses on descriptive analytics and does not incorporate user engagement metrics. Future improvements may include predictive modelling, sentiment analysis, and enhanced animations.

14. Conclusion

This project demonstrates the power of interactive data visualisation in analysing complex content datasets. By combining descriptive analytics with a well-structured dashboard, meaningful insights can be efficiently extracted.

15. GitHub Repository and Deployment

GitHub Repository: https://github.com/moti987/NETFLIX-CONTENT-ANALYTICS-DASHBOARD

16. Project Walkthrough and Live Dashboard

16.1 Walkthrough Video

Video Link: https://youtu.be/your-walkthrough-video-link

16.2 Live Dashboard

Live Dashboard Link: https://netflix-content-analytics-dashboard.streamlit.app