

Netflix Content Analysis Dashboard

Pdf.Report

Analytics by: Sumbul Abdul.Wahid



Data Science internship ____ FutureXcel

Introduction :

This project was conducted to analyze Netflix movies and TV shows data using Python. The main objective was to clean raw data, explore content trends, and build visual dashboards that help understand Netflix's content strategy and distribution patterns.

Data Preparation:

The dataset used in this analysis was cleaned and preprocessed in a previous task. All missing values were handled, data types were standardized, and the dataset was prepared in an analysis-ready format before performing visualization and dashboard analysis.

```
[5] 0s
from google.colab import files
files.download('/content/cleaned.netflix_data_(3)(1).csv')
```

LIBRARIES IMPORT

```
[6] 0s
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

DATA LOAD

```
[7] 0s
df = pd.read_csv('/content/cleaned.netflix_data_(3)(1).csv')
df.head()
```

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	Not Available	United States	2021-09-25	2020.0	PG-13	90 min	Documentaries	As her father nears the end of his life, film...
1	s2	TV Show	Blood & Water	Unknown	Ama Qamata, Kholi Ngema, Gail Mabalane, Thaban...	South Africa	2021-09-24	2021.0	TV-MA	2 Seasons	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	Unknown	2021-09-24	2021.0	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...	To protect his family from a powerful drug lor...
3	s4	TV Show	Jailbirds New Orleans	Unknown	Not Available	Unknown	2021-09-24	2021.0	TV-MA	1 Season	Docuseries, Reality TV	Feuds, flirtations and toilet talk go down amo...
4	s5	TV Show	Kota Factory	Unknown	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	2021-09-24	2021.0	TV-MA	2 Seasons	International TV Shows, Romantic TV Shows, TV ...	In a city of coaching centers known to train l...

```

[8] df.info()
    df.isnull().sum()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8088 entries, 0 to 8087
Data columns (total 12 columns):
 #   Column      Non-Null Count Dtype  
--- 
 0   show_id     8088 non-null   object  
 1   type        8088 non-null   object  
 2   title       8088 non-null   object  
 3   director    8088 non-null   object  
 4   cast         8088 non-null   object  
 5   country     8088 non-null   object  
 6   date_added  8088 non-null   object  
 7   release_year 8088 non-null   float64 
 8   rating      8088 non-null   object  
 9   duration    8088 non-null   object  
 10  listed_in   8088 non-null   object  
 11  description 8088 non-null   object  
dtypes: float64(1), object(11)
memory usage: 758.4+ KB

          0
show_id  0
type     0

          0
label    0
bi_words 0

          0
memory usage: 758.4+ KB
memory usage: 758.4+ KB

```

Data Visualization & Analysis:

Multiple visualizations were created to answer key questions:

What is the ratio of movies to TV shows on Netflix?

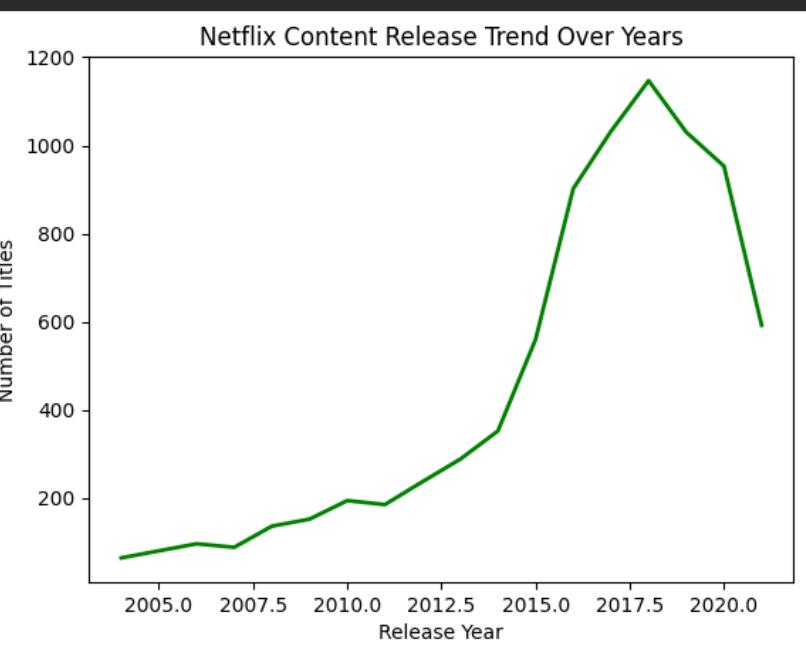
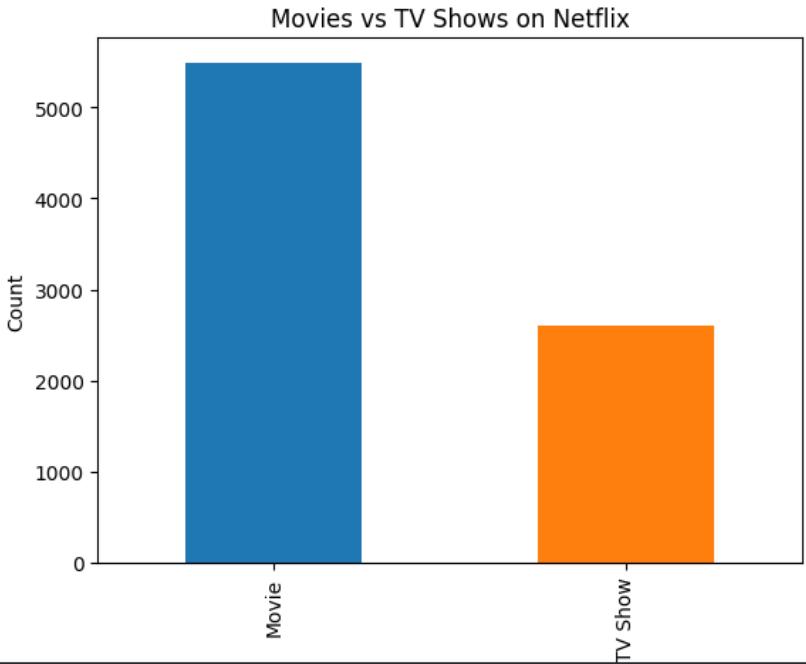
How has Netflix content grown over the years?

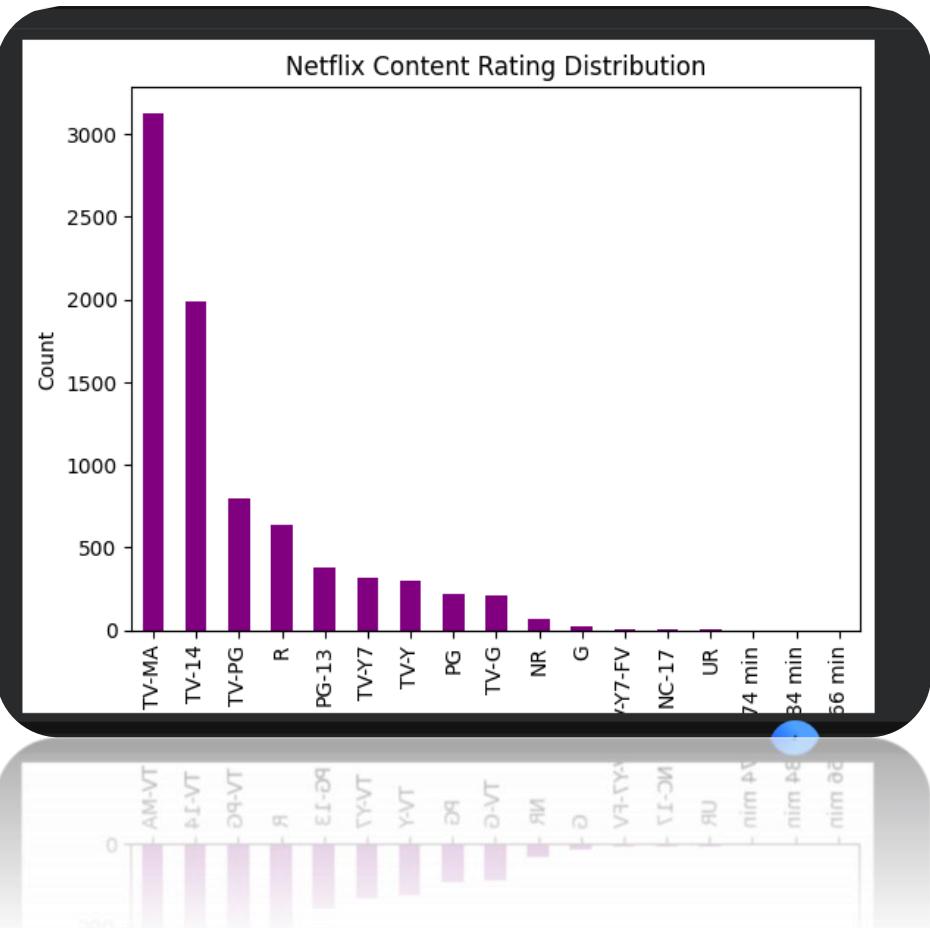
Which countries contribute the most content?

What are the common content ratings?

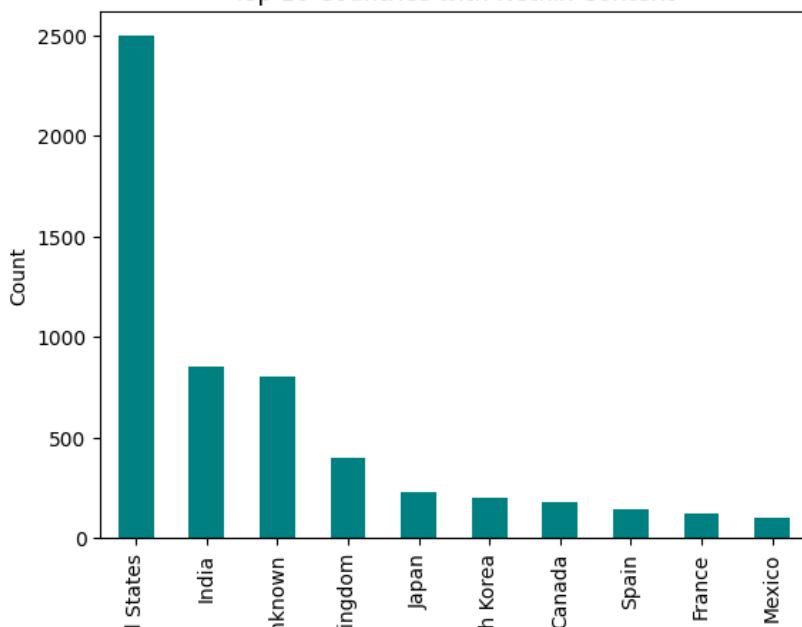
What is the typical duration of movies?

These visualizations helped identify content trends and platform focus.

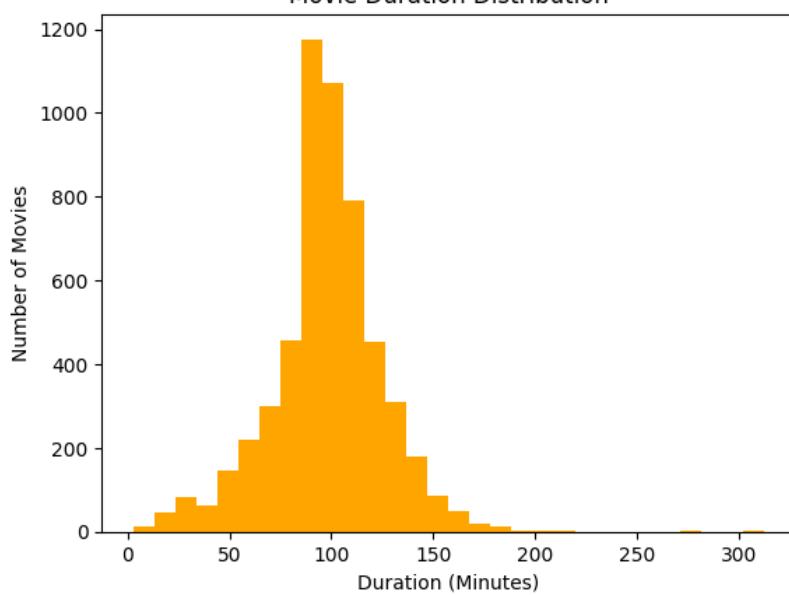


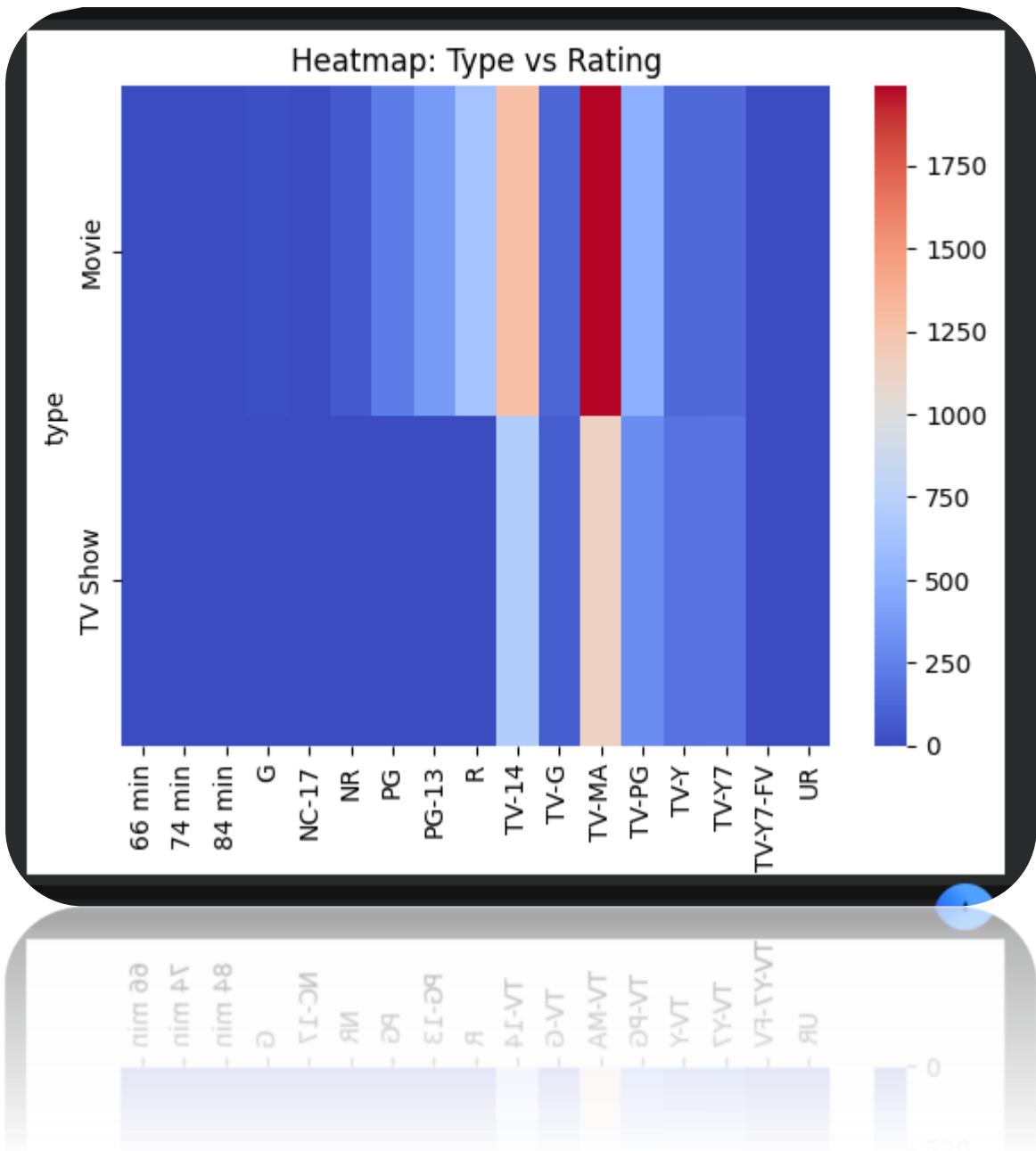


Top 10 Countries with Netflix Content



Movie Duration Distribution





Chatbot Implementation:

A simple rule-based chatbot was developed to allow users to interact with the Netflix dataset. Users can ask basic questions such as total movies, TV shows, top producing country, and most common rating. The chatbot continues interaction until the user exits.

```
9]      return "Goodbye! Chatbot closed."
2m    else:
        return "Sorry, I can answer only basic Netflix data questions."
    # Chatbot loop
    while True:
        user_question = input("Ask a question about Netflix data (type 'bye' to exit): ")
        answer = netflix_chatbot(user_question)
        print(answer)

        if user_question.lower() == "bye":
            break

Ask a question about Netflix data (type 'bye' to exit): how many tv shows are there
Total TV Shows: 2601
Ask a question about Netflix data (type 'bye' to exit): how many movies are there
Total Movies: 5487
Ask a question about Netflix data (type 'bye' to exit): Top countries
Sorry, I can answer only basic Netflix data questions.
Ask a question about Netflix data (type 'bye' to exit): common rating
Most common rating: TV-MA
Ask a question about Netflix data (type 'bye' to exit): top country
Top Country: United States
Ask a question about Netflix data (type 'bye' to exit): bye
Goodbye! Chatbot closed.
```

Conclusion

This project demonstrates practical data science skills including data cleaning, visualization, and basic AI interaction. The insights gained provide a clear understanding of Netflix's content distribution and strategic focus.