

Netflix Data Analysis

A comprehensive analysis of Netflix's content library, revealing patterns in production, genres, geography, and ratings that shape the streaming landscape.

Your Name: Gourab Saha

Guide Name: Ms. NAINA DEVI

Streamlit App: <https://gourab-saha-netflix-data-analysis-6txkafs72jyduyksjsbjzo.streamlit.app/>



Agenda

Explore the journey of analyzing Netflix data, from initial concept to an interactive dashboard.

01

Introduction to the Project

03

Project Workflow & Implementation

05

Applications

07

Future Scope

02

Tools and Technologies Used

04

Output and Key Insights

06

Problems Faced & Solutions

08

Project Link



Introduction to the Project

This project delves into Netflix's vast content library to uncover compelling trends and insights.

Exploratory Data Analysis (EDA)

Performing comprehensive EDA on Netflix data to understand content trends and distribution patterns.

Interactive Dashboard

Developing an interactive Streamlit application for dynamic exploration of genre distribution, ratings, and release years.

Hidden Trends

The primary goal is to reveal underlying patterns in Netflix's diverse global content catalog.

Tools and Technologies Used

A robust set of tools powered this data analysis and application development.



Python

The core programming language for all data manipulation and analysis.



IDE

Jupyter Notebook and VS Code facilitated efficient coding and iterative analysis.



Data Libraries

Pandas, NumPy for data handling; Matplotlib, Seaborn, Plotly for advanced visualizations.



GitHub

Essential for version control, collaborative development, and project hosting.



Streamlit

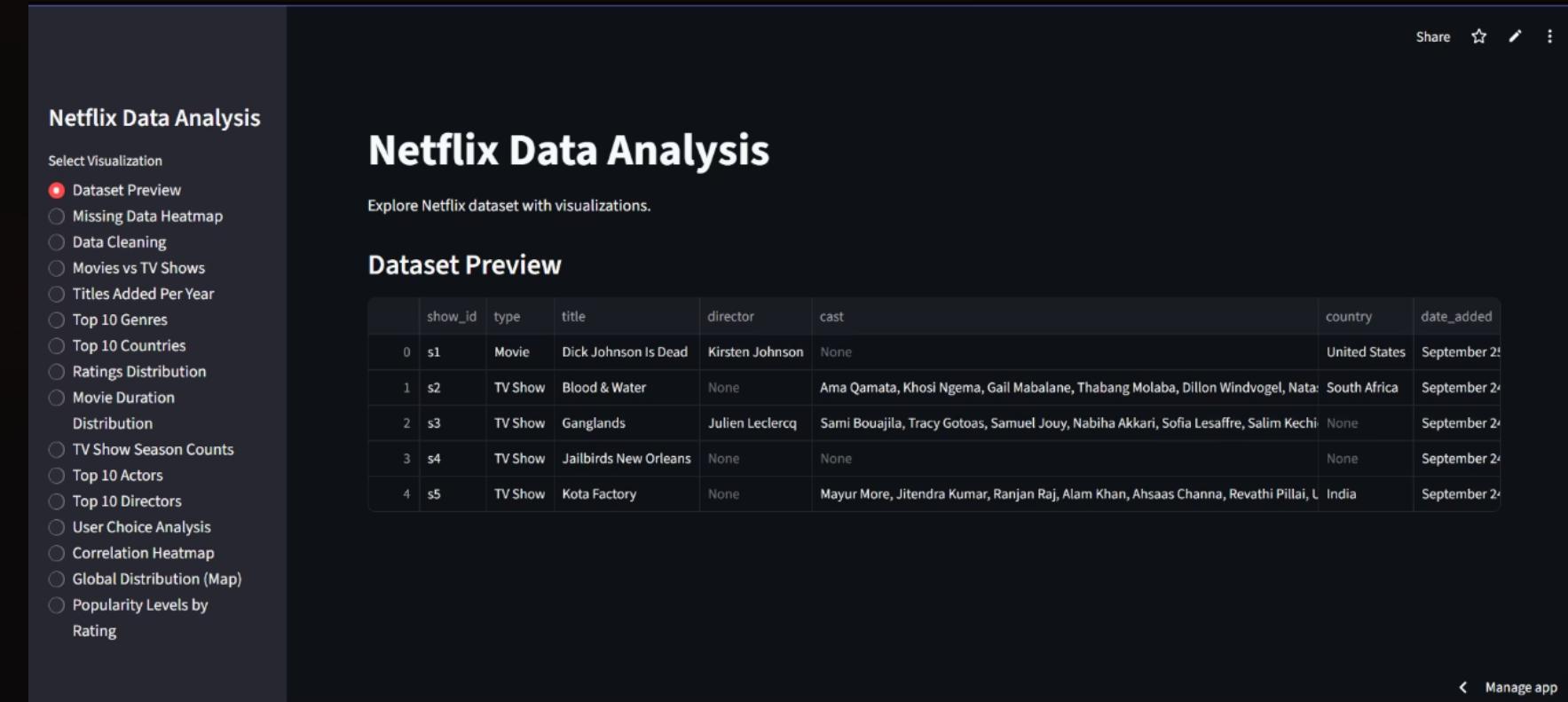
For building the intuitive and interactive web application dashboard.

Project Working: Implementation & Output

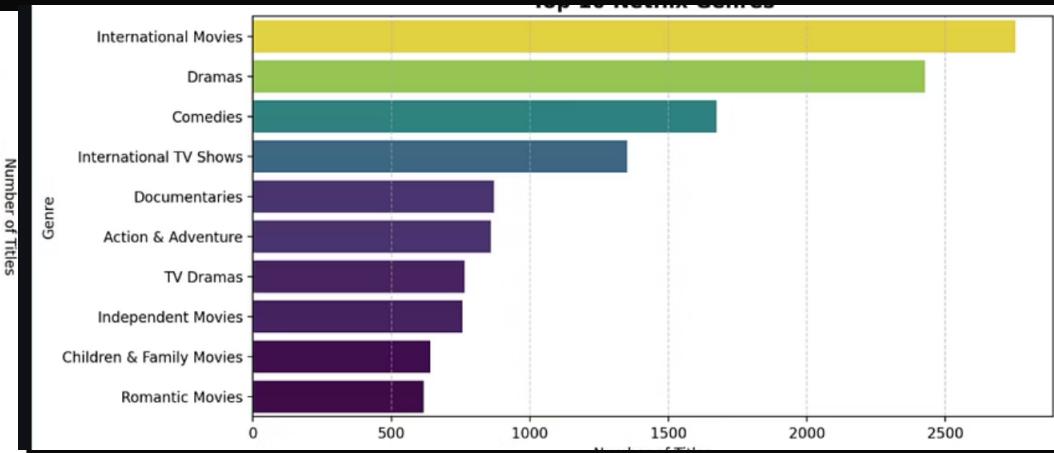
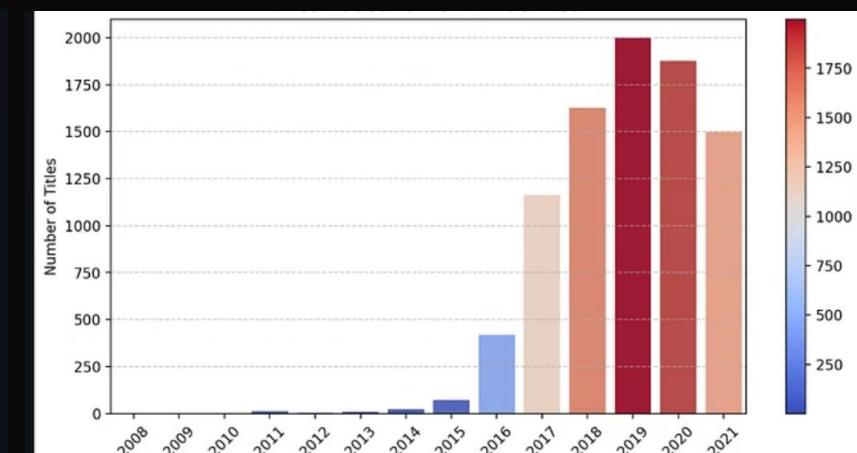
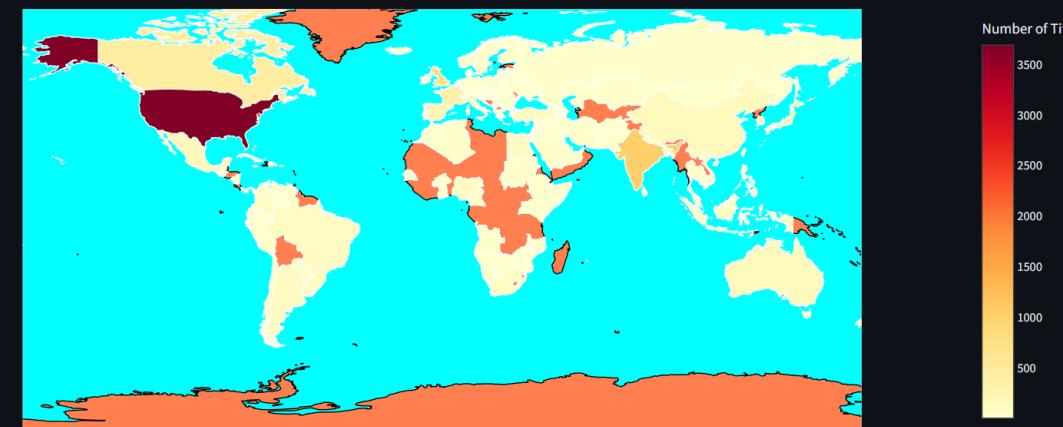
From raw data to an interactive interface, here's how the project came to life.

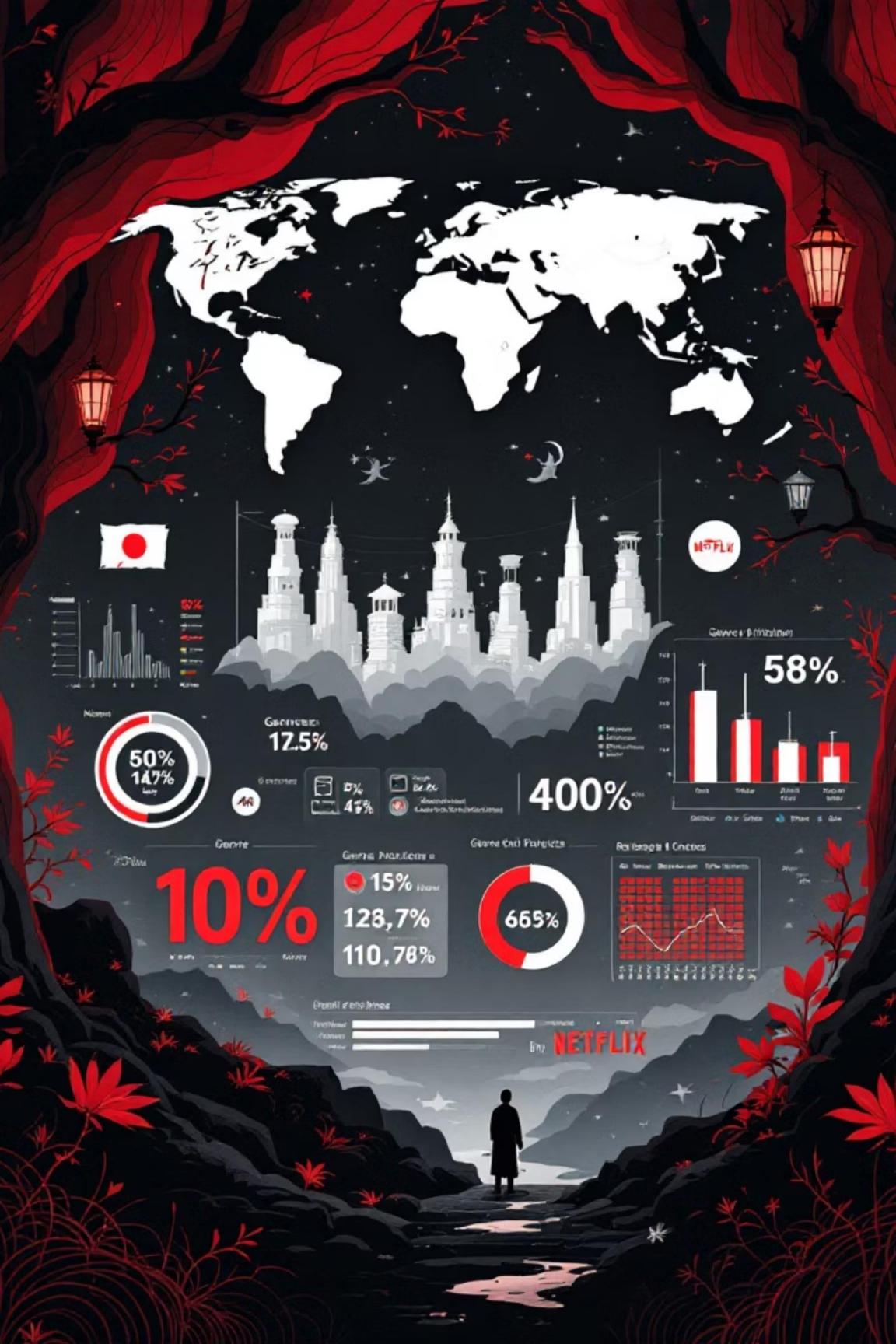
Implementation Steps:

1. Data Collection and Cleaning
2. Data Exploration and Visualization
3. Building Interactive Dashboards with Streamlit
4. Deployment using Streamlit Cloud



Below are glimpses of the interactive dashboard, showcasing key data points.





Applications

This project serves a variety of practical and educational purposes.



Content Strategy

Assists media companies in making data-driven decisions for content acquisition and production.



Data Enthusiasts

Provides a hands-on platform for data enthusiasts and analysts to study Netflix's content landscape.



Recommendation Prototyping

Acts as a foundational step for developing and testing content recommendation systems.



Educational Project

Demonstrates practical applications of EDA, data visualization, and web application deployment.

Problems Faced & Solutions

Every project encounters hurdles; here's how we overcame them.

Handling missing and duplicate data

Utilized Pandas methods like `dropna()` and `drop_duplicates()` for efficient data cleaning.

Genre and country column had multiple values per row

Developed custom functions to split strings and accurately count occurrences for precise analysis.

Streamlit deployment issues

Adjusted the application structure and meticulously configured the `requirements.txt` file.

Large dataset visualization lag

Implemented data sampling techniques and Streamlit's caching mechanisms for optimized performance.



Future Scope

The journey doesn't end here; exciting enhancements await.



Machine Learning

Integrate ML algorithms to provide personalized content recommendations.

User Personalization

Implement user login and personalized dashboard features based on viewing history.

Real-Time Updates

Incorporate APIs for real-time trend updates, reflecting the most current content data.

Multi-Page Dashboard

Expand the dashboard into multiple pages for more advanced and segmented data exploration.



Project Link

Explore the live application and source code for a deeper dive.

Streamlit App: <https://gourab-saha-netflix-data-analysis-6txkafs72jyduyksjsbjzo.streamlit.app/>

GitHub Repository:

https://github.com/Gourabsaha18/Gourab_Saha_Netflix-Data-Analysis

Demo Video Link: <https://youtu.be/KmMTYa2eO-A>