

EDA ON NETFLIX MOVIES ON MOVIES AND TV SHOWS DATASET

- Presented By: **Shreya Singh**
- Guided By: Naina Devi
- College: Techno India University



AGENDA:

- Understand and analyze the Netflix movies and TV shows dataset.
- Perform data cleaning and preprocessing.
- Visualize important trends and insights.
- Solve encountered problems and discuss solutions.
- Explore potential applications and future development.



NETFLIX



INTRODUCTION:

This project explores the Netflix Movies and TV Shows dataset using Exploratory Data Analysis (EDA). The dataset contains details about various titles, including their type, cast, country, release date, and genre. By analyzing this data, we can uncover interesting trends and patterns in what's available on Netflix and better understand its content.

OBJECTIVE:

- The objective of this EDA is to examine the Netflix Movies and TV Shows dataset to find key patterns, trends, and insights about the content available on the platform
 - To identify content trends and viewer preferences on Netflix to help inform decisions about what types of movies and shows to recommend or produce.



LIBRARIES AND TECHNOLOGIES USED:

Programming Language: Python

Libraries Used:

- **Numpy** for data manipulation
- **Matplotlib**
- **Sckit learn** for predictive modelling
- **Joblib** for model serialization

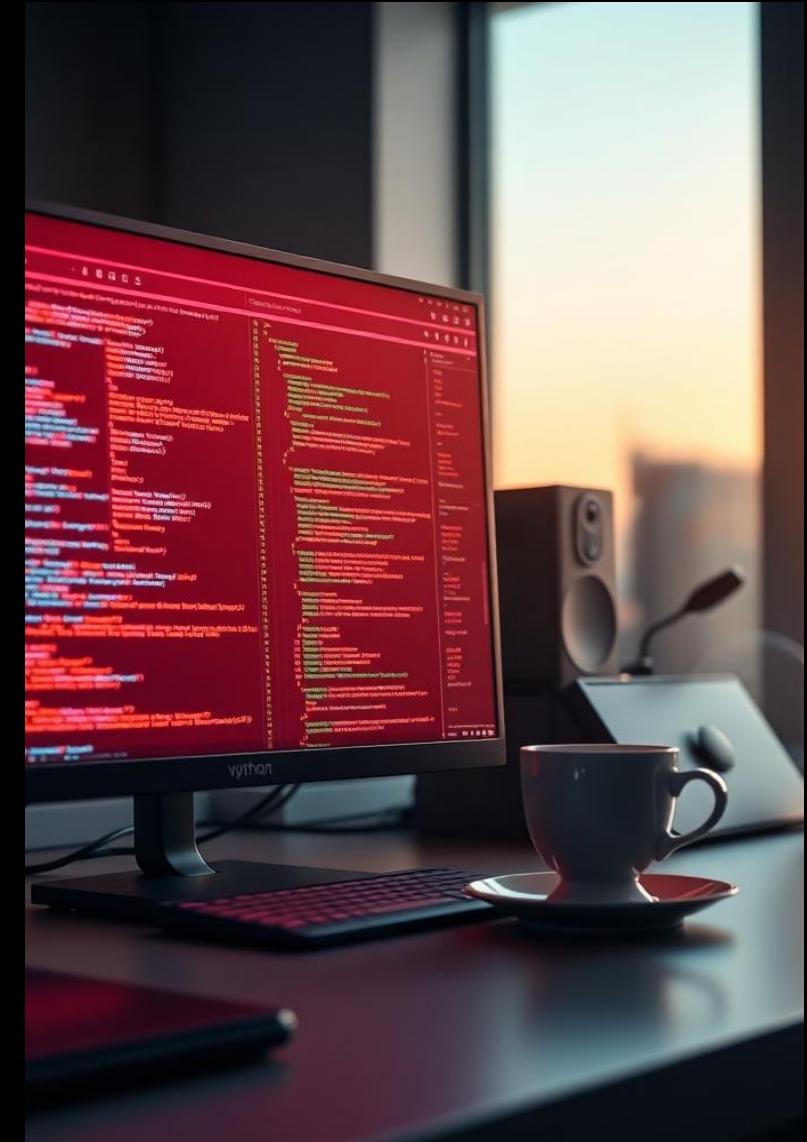
• **Deployment and Interface:**

- **Streamlit** for rapid development

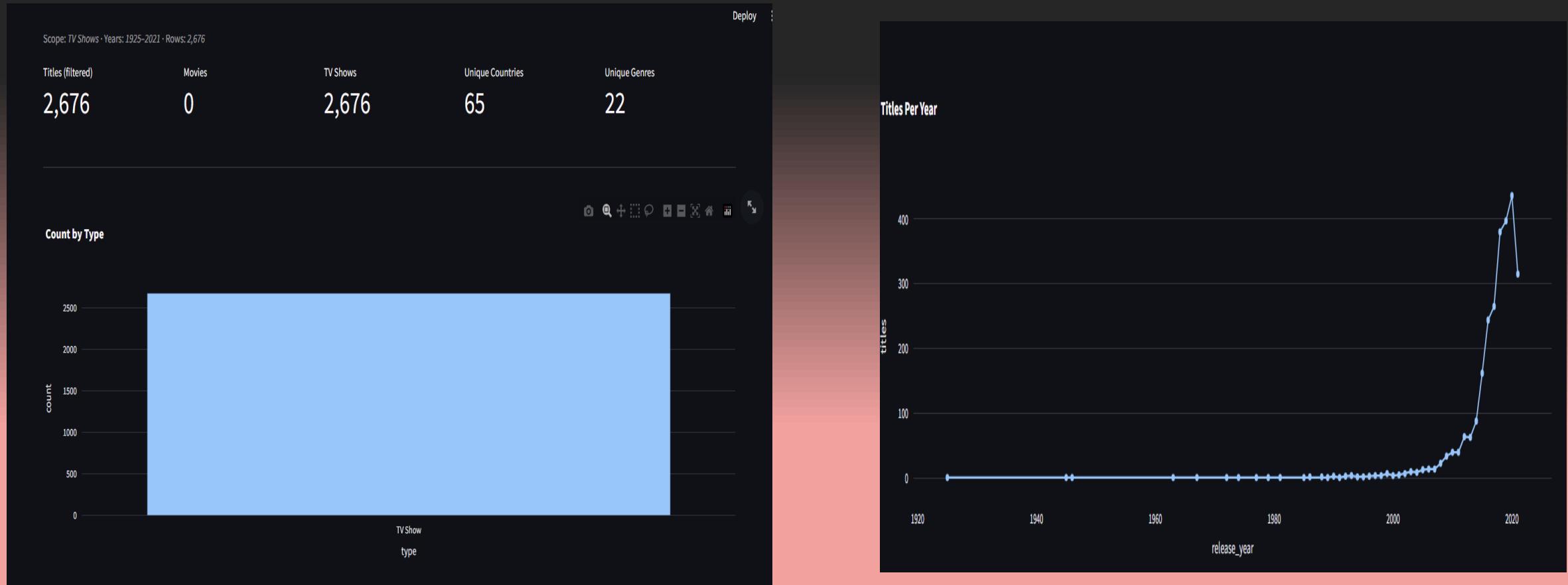
- **Render** for cloud deployment

• **Dataset Source:**

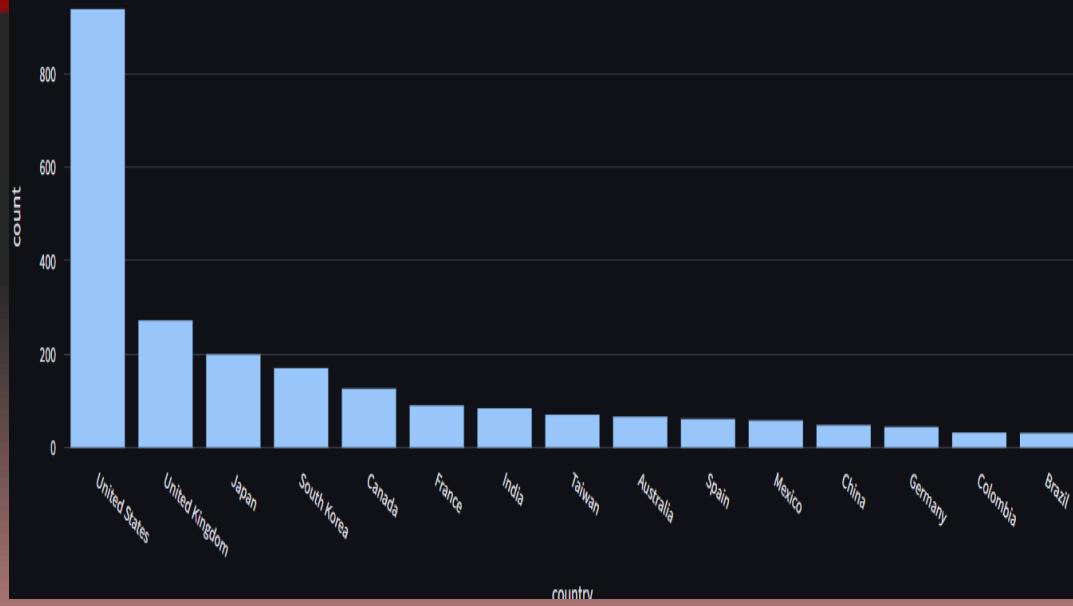
- Kaggle
- Mymoviedb.csv provides rich features for Netflix shows EDA
- Development in **Jupyter Notebook** and version control with **GitHub**



PROJECT WORKING



Top Countries (by mentions)



Titles (filtered)

2,676

Movies

0

TV Shows

2,676

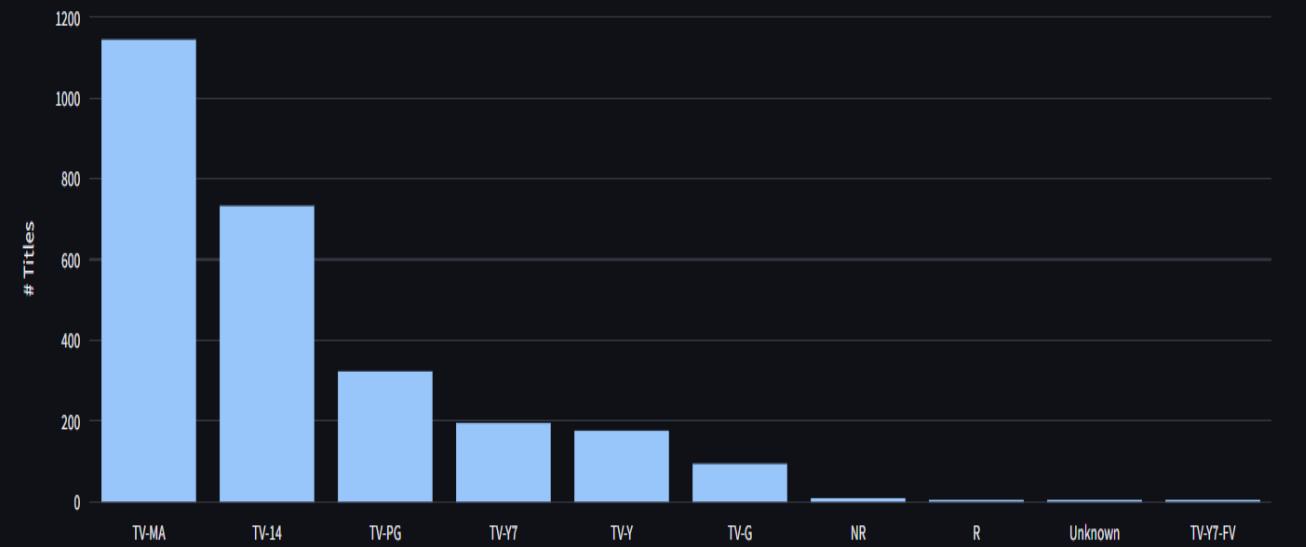
Unique Countries

65

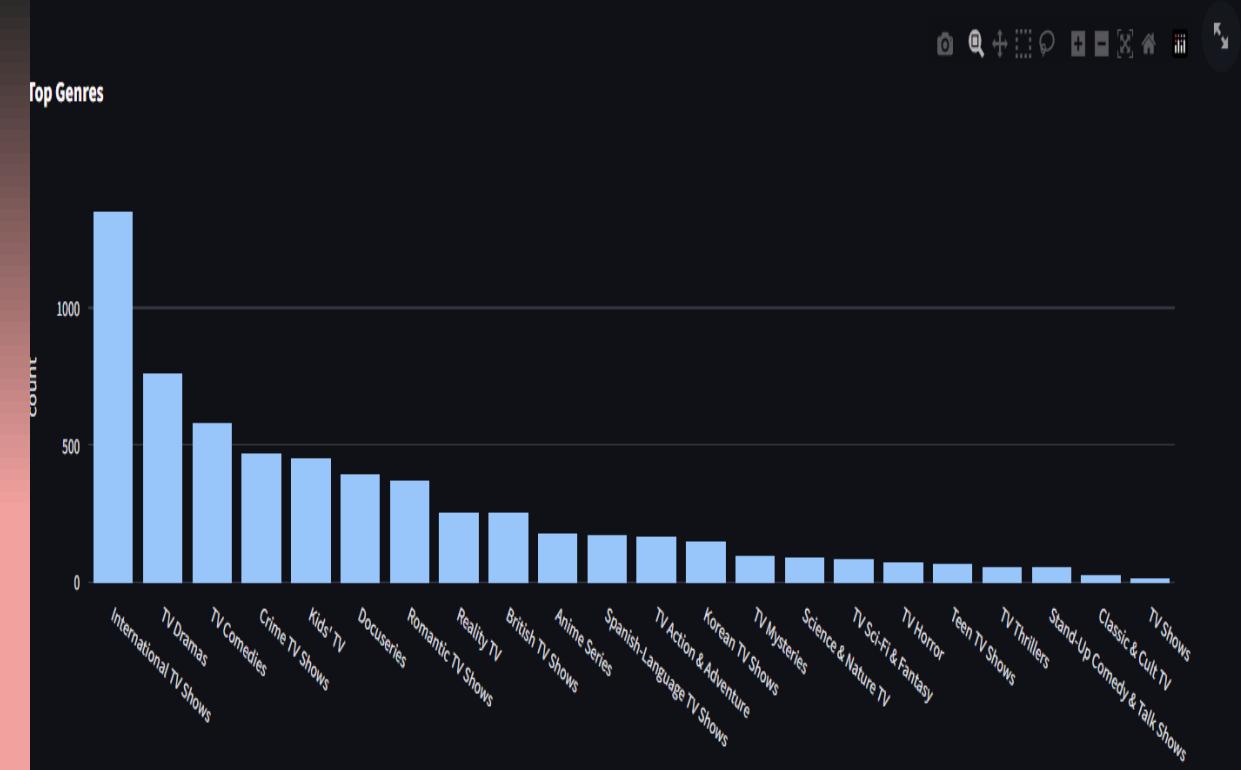
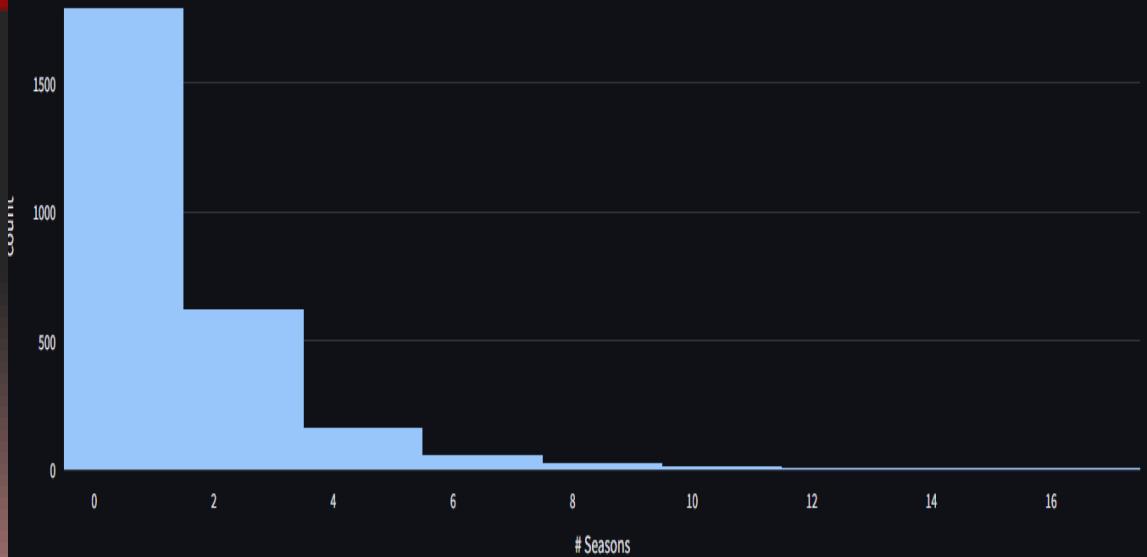
Unique Genres

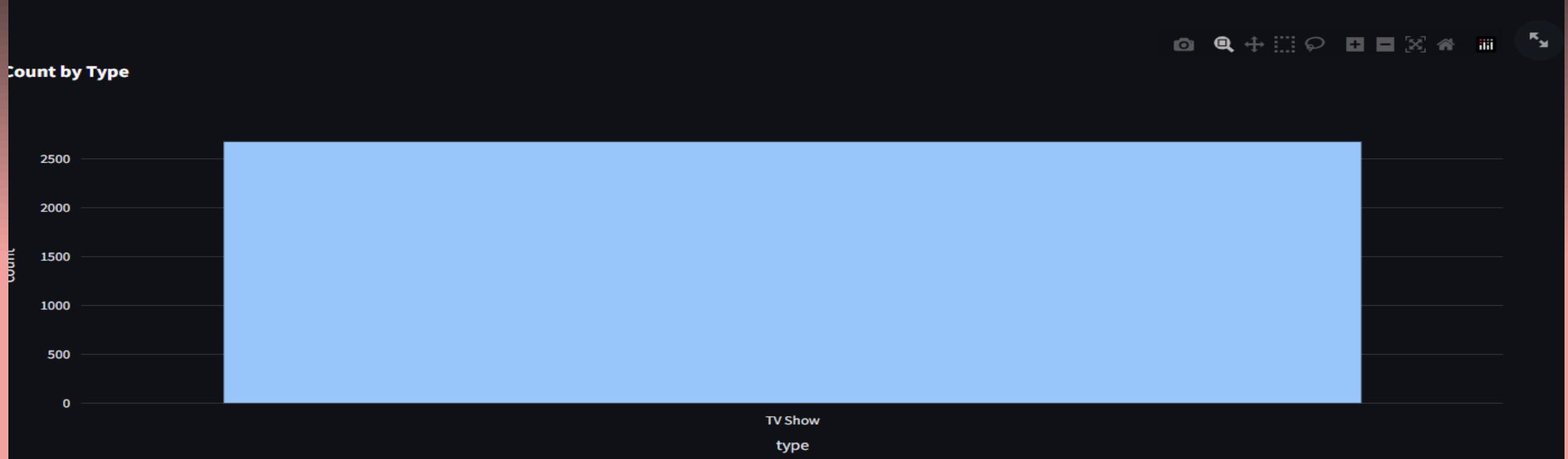
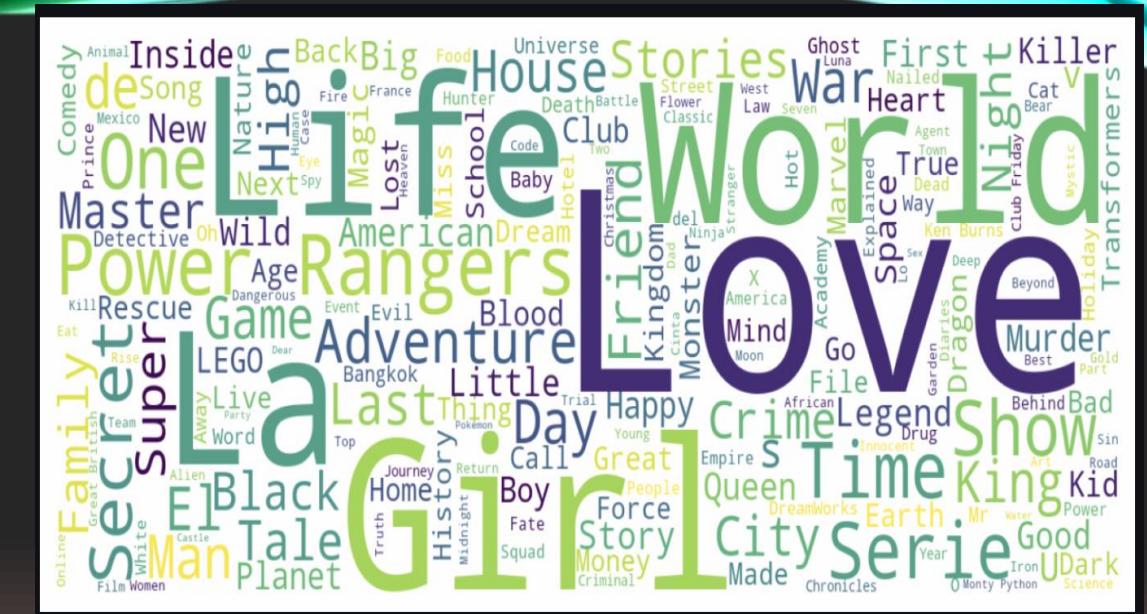
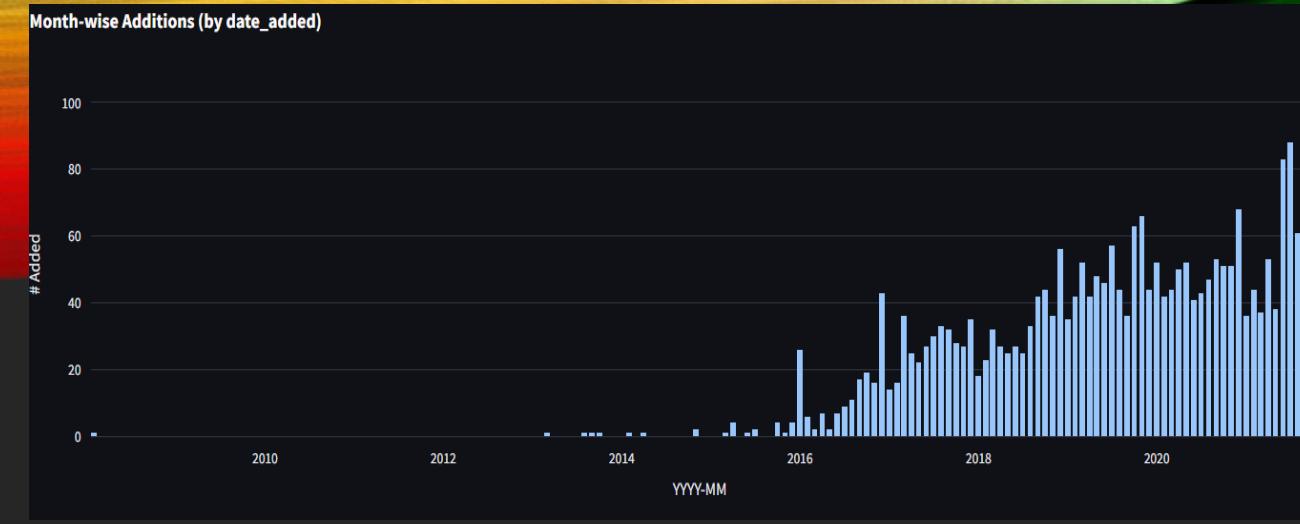
22

Rating Distribution



TV Show Seasons Histogram





APPLICATIONS:

- Content analytics for streaming services
- Personalized recommendation systems based on user trends and genre preferences
- Market research for content acquisition strategy
- Sentiment and pattern analysis for improved user engagement and usage statistics



PROBLEMS FACED & SOLUTIONS:

Challenges

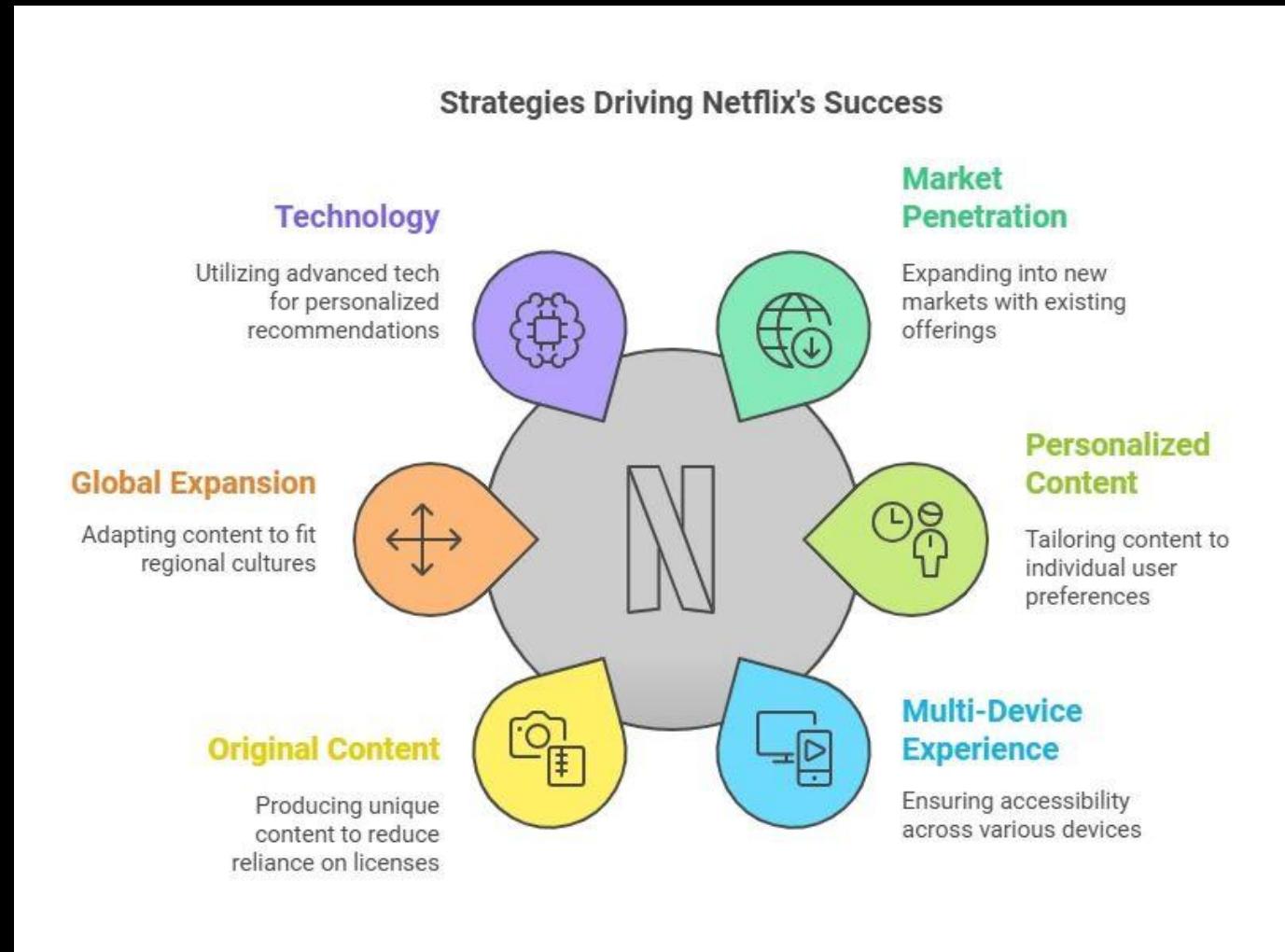
- Some data columns were in incorrect or mixed formats (like dates and time durations).
- Needed a way to make the project more engaging and user-friendly.
- It was hard to decide which visualizations best represent the data.
- Large dataset with many attributes

Solutions

- Reformatted those columns into the right data types to make analysis smoother.
- Developed an interactive Streamlit dashboard where users can explore data by selecting different options.
- Tried different chart types like bar graphs, pie charts, and word clouds, then chose the ones that made the insights clear and easy to understand.
- Preprocessing and efficient filtering using Pandas and Scikit-learn

FUTURE SCOPE:

- Use machine learning to predict popular genres and content types
- Advanced clustering and recommendation systems with deep learning
- Integrate user interaction data for personalized content suggestions
- Expand analysis with real-time streaming data and sentiment analysis





THANK YOU