***Netflix Data Analysis: A Beginner Project***

*By Varun Kumar*

**Tools Used:**

* Python (Pandas, NumPy, Matplotlib)
* VS Code
* Kaggle (Dataset Source)

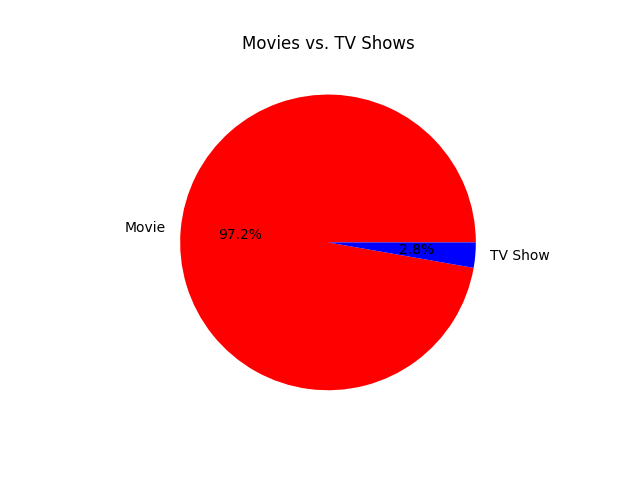
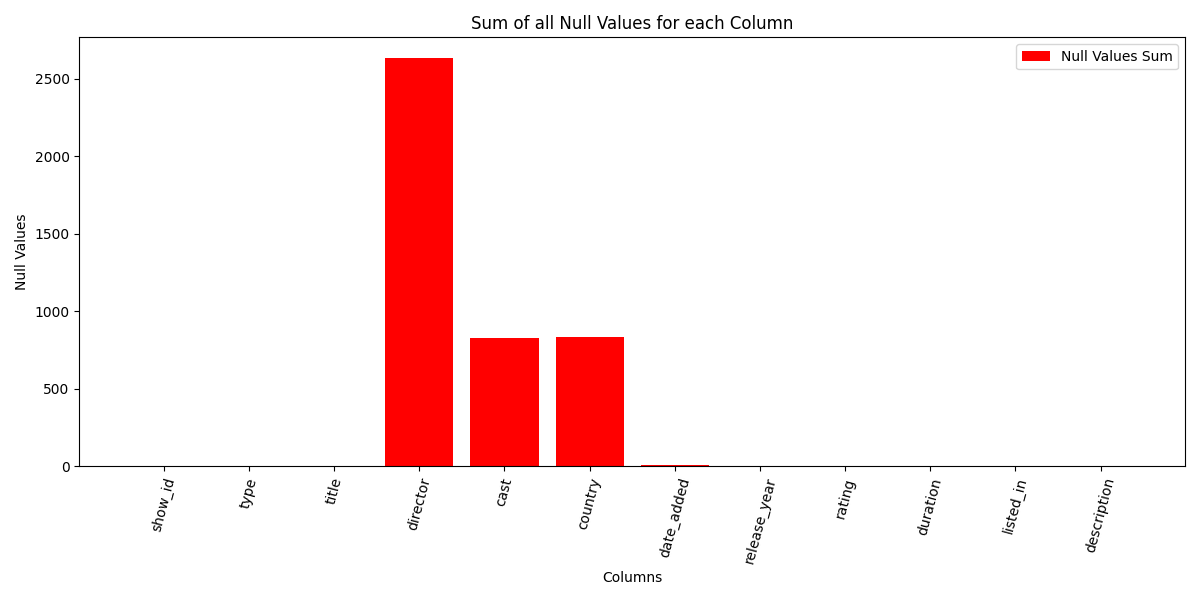
**Dataset Source:**

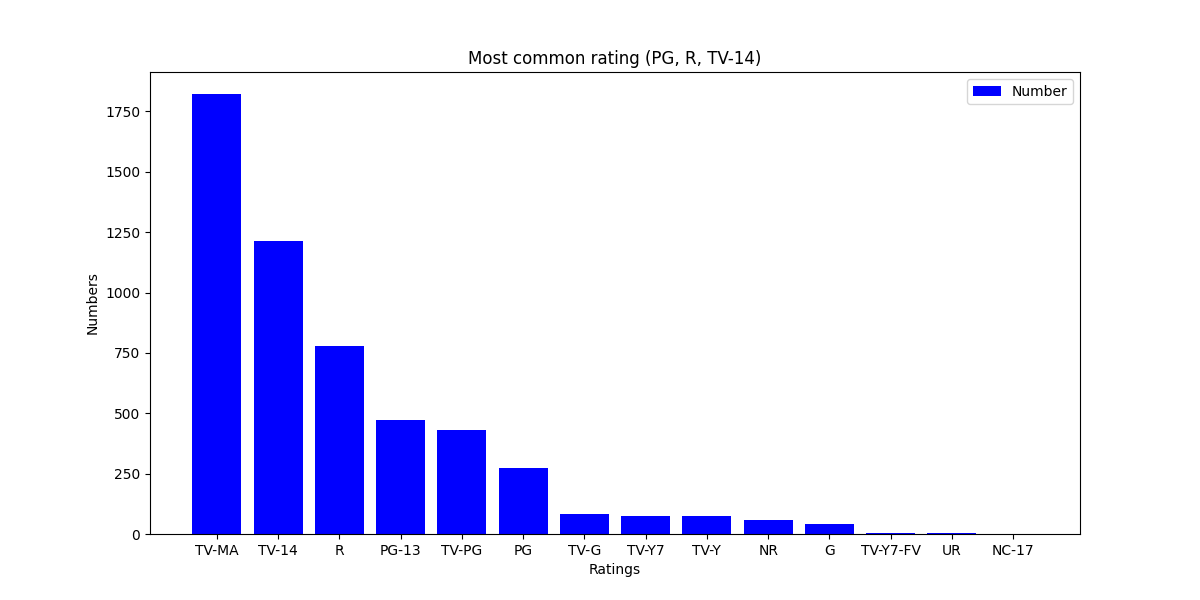
<https://www.kaggle.com/datasets/shivamb/netflix-shows>

**Questions Analyzed:**

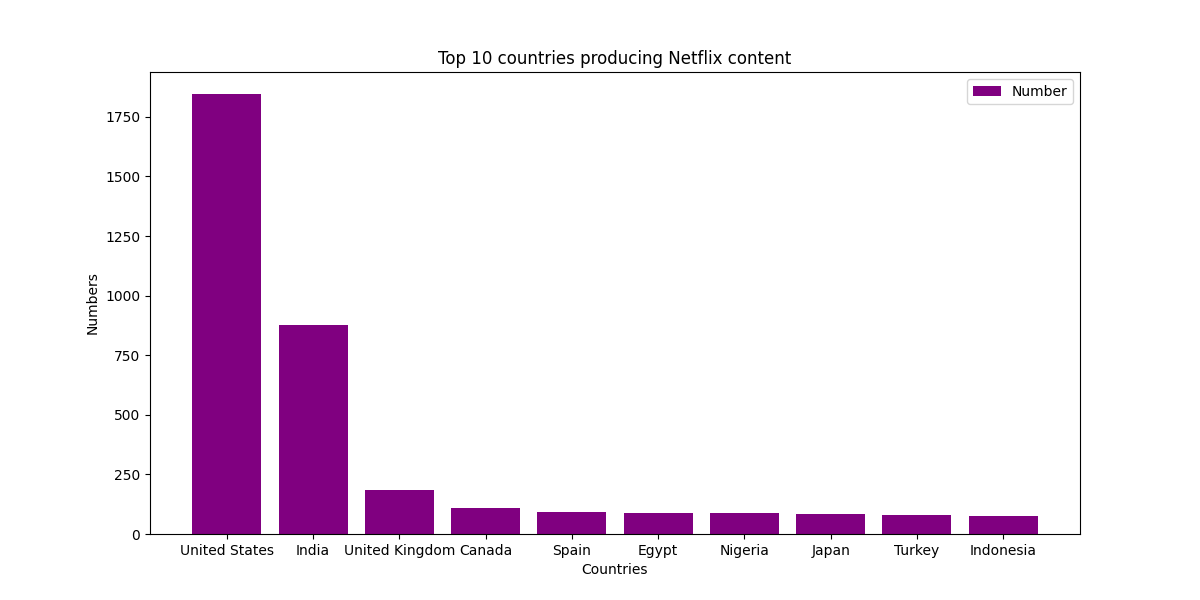
* How many shows are Movies vs. TV Shows?
* What’s the most common rating?
* Which year had the highest number of releases?
* Top 10 countries producing Netflix content
* Average duration of movies

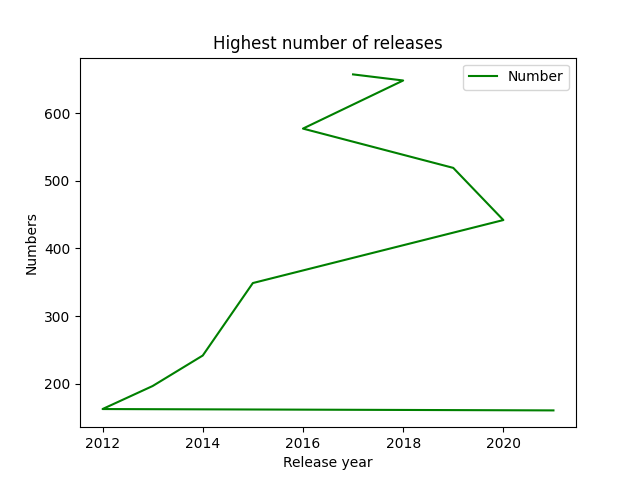
**Graphs Created:**

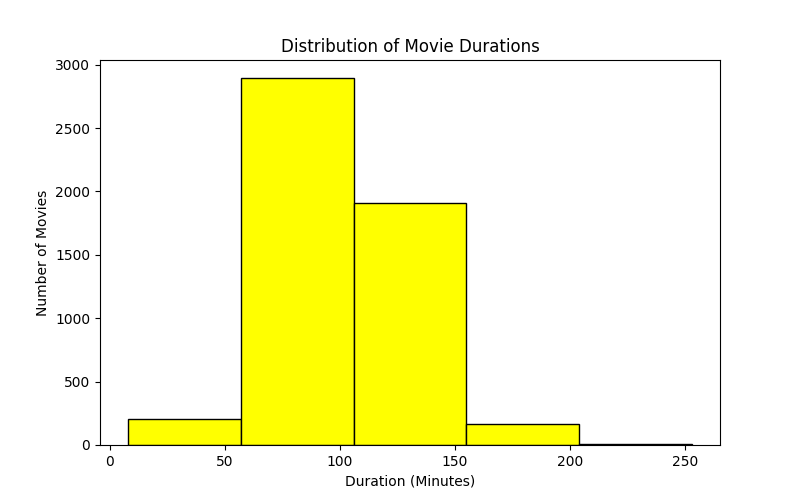
* **Pie Chart:** Number of Movies vs TV Shows
* **Bar Chart 1 :** Null values per column
* **Bar Chart 2 :** Most Common Ratings



* **Bar Chart 3 :** Top 10 Countries producing Netflix Content



* **Line Plot:** Year-wise release trend
* **Hist Plot:** Distribution of Movie Durations



**Observations:**

* Movies are more frequent than TV Shows
* ‘TV-MA’ is the most common rating
* 2019 saw the highest number of releases
* Most content originates from the USA, India, and UK
* Average movie duration is around **102.7** **mins**

**Conclusion:**  
 This project helped me understand basic data cleaning, exploratory data analysis (EDA), and visualizations using Python libraries. It also highlighted trends in Netflix’s content distribution.

**Future Improvements:**

* Include actor/genre-wise analysis
* Analyze trends by month and day
* Use interactive dashboards with Tableau or Power BI