PROJECT REPORT

COUSE CODE : CSE317

COURSE TITLE : SOFTWARE PROJECT V

SUBMITTED TO :

FARIHA JAHAN

LECTURER, DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

DAFFODIL INTERNATIONAL UNIVERSITY(DIU)

SUBMITTED BY :

SHAZID NAWAZ (191-15-12929)

M SHAHRIAR ISHTIQUE (191-15-12938)

RADUAN AHAMAD (191-15-12943)

FARAZ AHMED (191-15-12948)

SAKIB ROKONI (191-15-12961)

GITHUB LINK : Project Link (Group 4)

Title: TV Shows and Movies listed on Netflix

Abstract: This dataset consists of tv shows and movies available on Netflix as of 2019. The dataset is collected from Fixable which is a third-party Netflix search engine.

In 2018, they released an interesting report which shows that the number of TV shows on Netflix has nearly tripled since 2010. The streaming service's number of movies has decreased by more than 2,000 titles since 2010, while its number of TV shows has nearly tripled. It will be interesting to explore what all other insights can be obtained from the same dataset.

Integrating this dataset with other external datasets such as IMDB ratings, rotten tomatoes can also provide many interesting findings

Introduction: Data visualization is the graphical representation of information and data. By using visual elements like charts, graphs, and maps, data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data.

In the world of Big Data, data visualization tools and technologies are essential to analyze massive amounts of information and make data-driven decisions.

We used TV Shows and Movies listed on Netflix dataset for this project. We performed the following tasks from this dataset:

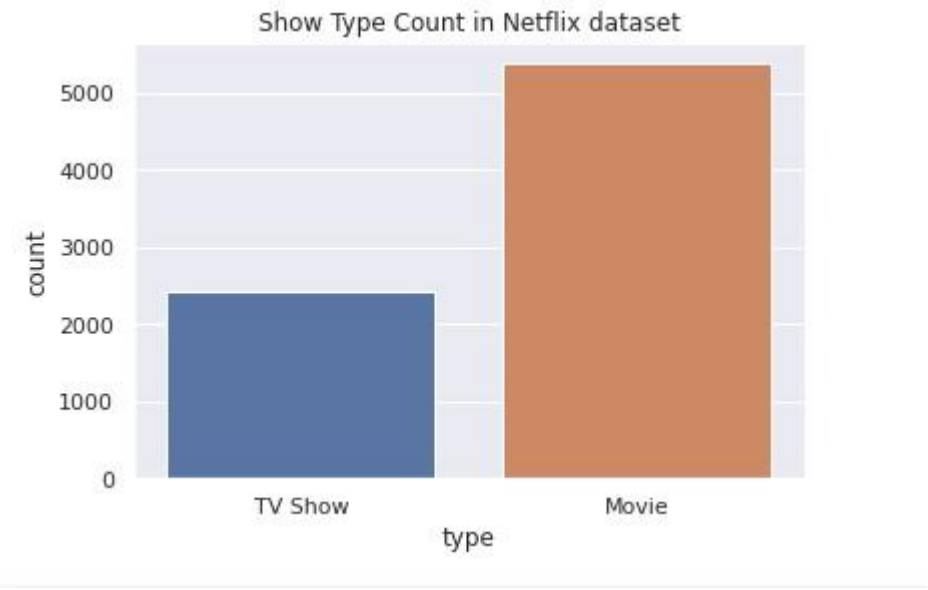
1. Understanding what content is available in different countries
2. Is Netflix has increasingly focusing on TV rather than movies in recent years.

Methodology:

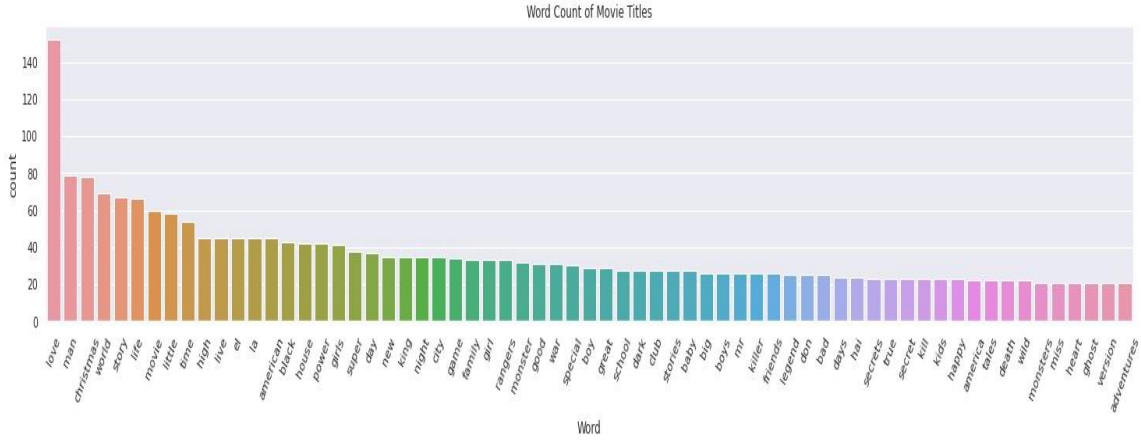
Dataset: The dataset has 7000 rows and 12 columns. Columns are show_id ,type , title , director , date_added , release_year ,country , listed_in , description , cast .

Colaboratory: Colaboratory or "Colab" for brief , could even be a product from Google Research. Colab allows anybody to write down and to execute arbitrary python code through the browser, and is particularly compatible to machine learning, data analysis and education. More technically, Colab could even be a hosted Jupyter notebook service that needs no setup to use, while providing free access to computing resources including GPUs. Google have released Colaboratory: a web IDE for python, to enable Machine Learning with storage on the cloud — this internal tool had a reasonably quiet public release in late 2017, and is on the brink of form an enormous difference within the earth of machine learning, AI and data science work.

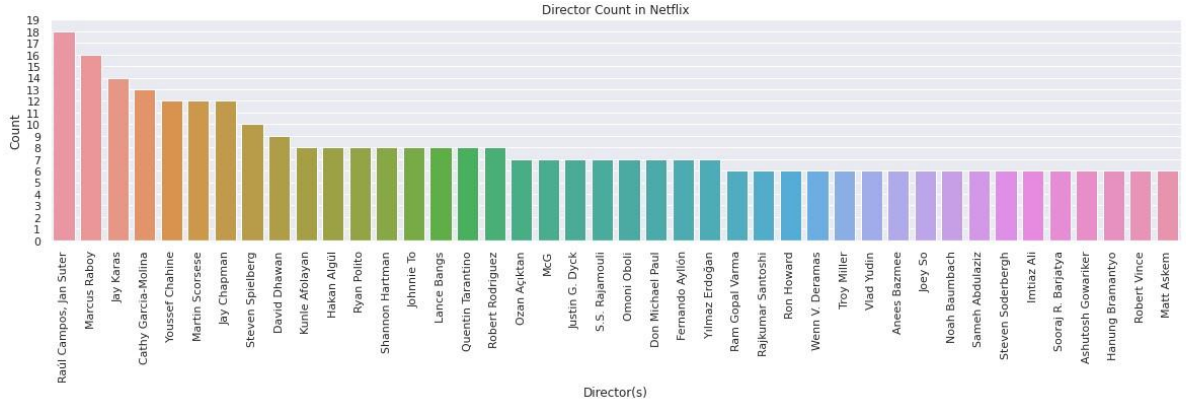
Results & Discussion:



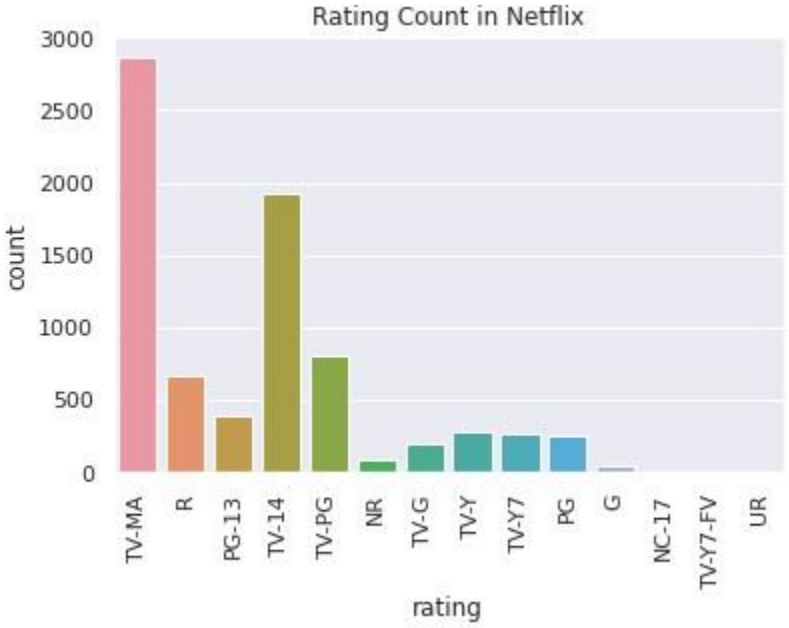
In this stacked column chart on 'Show type Count on Netflix dataset where, type = Legend & Count of type = values. TV shows types of Movies & TV shows.



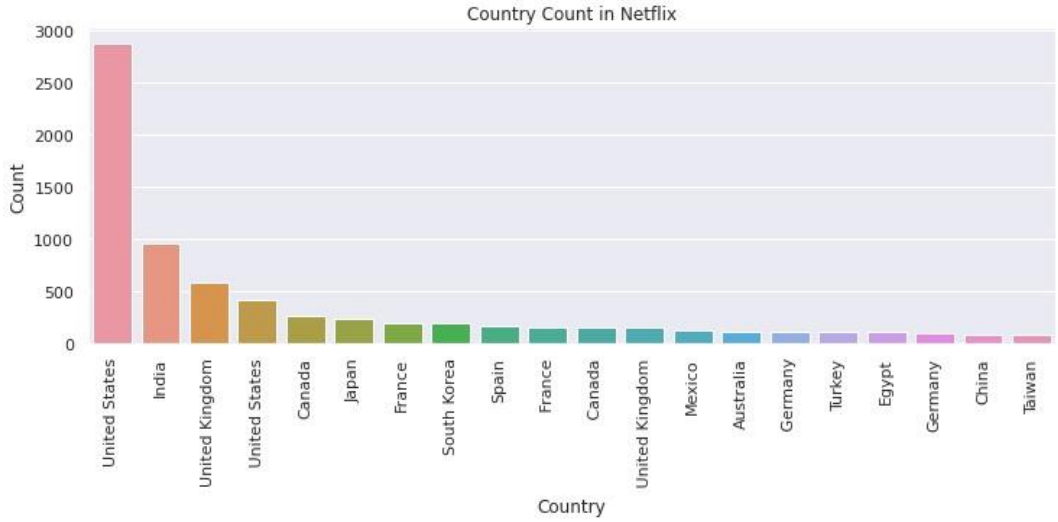
In this Chart we work on word count of movies Titles. We search which word is the most use in Netflix. So, we find 'Love' is the word which is most use in the Netflix.



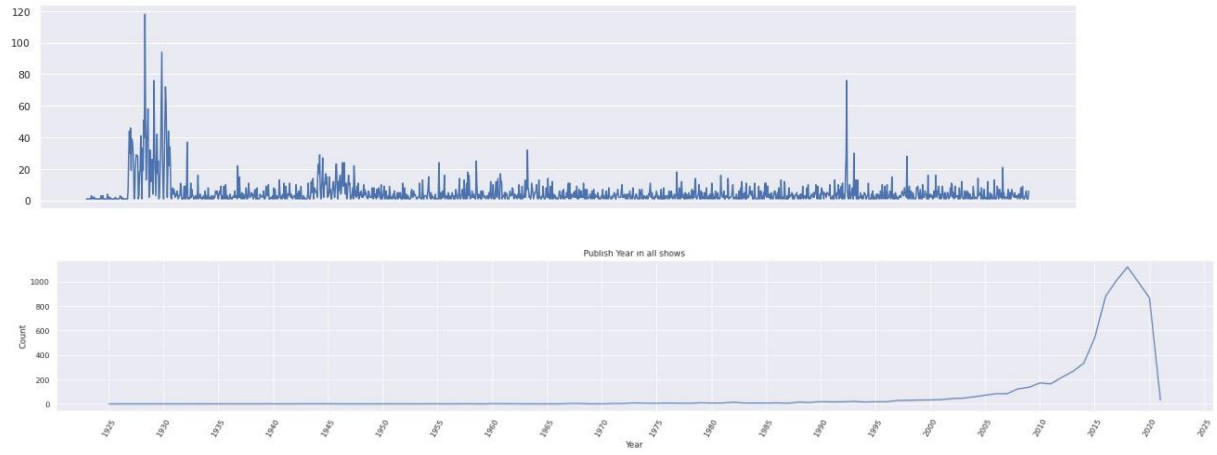
In the chart we find, what is the most release movies in Netflix by one director. We find Raul Campos. Jan Suter's movie most release on Netflix.



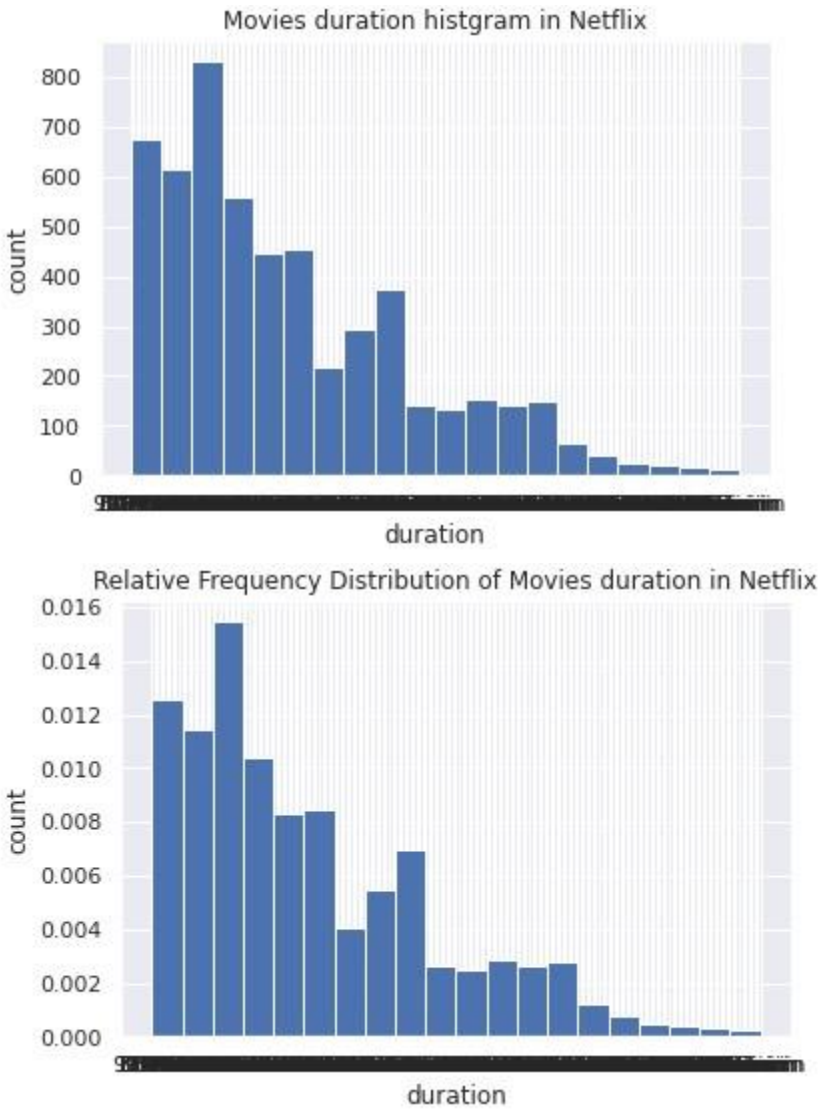
In this chart we work on Count of show_id by rating where Legend = rating & values = Count of show_id. Here highest rating show which is TV-MA & lowest rating show is less than which is NC-17.



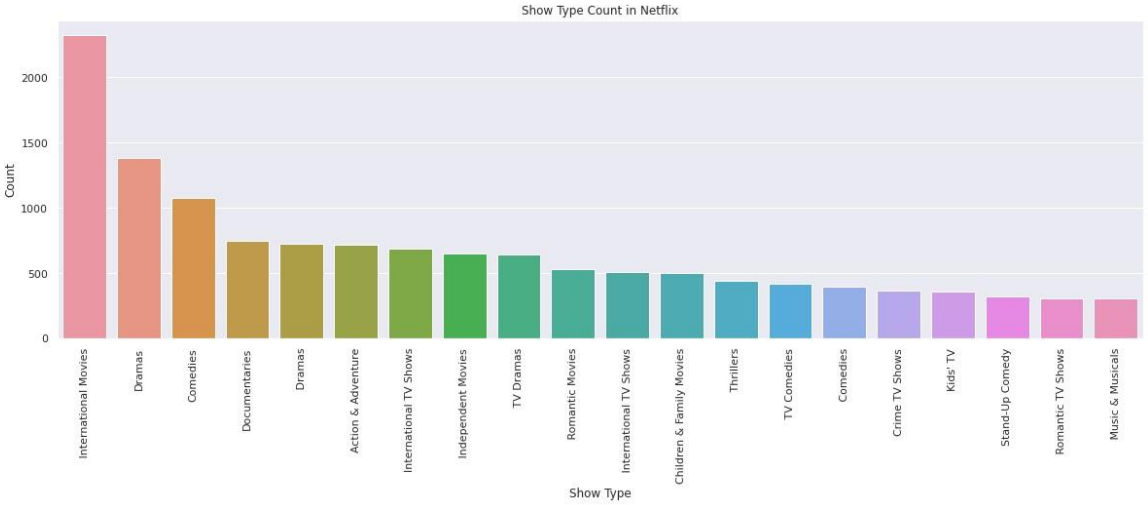
In this chart we work on country Count where Group =Country & Values = Count. In this graph we see United States is release biggest part of the Movies & TV shows on Netflix, India comes second.



In this Line chart we work on when Netflix get most publicity or which year Netflix publish most of the movies. We find from 2010 to 2020 Netflix get most of the publicity & 2020 Netflix release the most of the movies.



In this Graph we work on movies duration histogram on Netflix. It is greater than 800 min long tv series found on Netflix.



In this chart we work on which type of movies get most release on Netflix. We found International Movies got first rank on Netflix.

Conclusion:

The aim of this project was to analyze the dataset of tv shows and movies available on Netflix from 1925 - 2020 using python programming language. 10 visuals or charts are created by analyzing the entire dataset. Looking at the visuals, it is understood that from 2010 to 2020 Netflix got the most of publicity & In 2020 Netflix release most of the movies.