DATA VISUALISATION PROJECT REPORT

(Project Semester: January-May 2023)

NETFLIX TITLES

Submitted by

Nirbhay Singh

12019256

Program and Section: B.Tech(CSE), KM018

Course Code: INT233

Under the Guidance of

Nidhi Arora : 28373

Discipline of CSE/IT

Lovely School of Computer Science & Engineering

Lovely Professional University, Phagwara

CERTIFICATE

This is to certify that Nirbhay Singh bearing Registration no. 12019256 has completed INT233 project titled, **"NETFLIX TITLES "** under my guidance and supervision. To the best of my knowledge, the present work is the result of his original development, effort, and study.

Nidhi Arora
School of Computer Science & Engineering
Lovely Professional University Phagwara,
Punjab.

Date: 12-4-2023

DECLARATION

I, Nirbhay Singh, student of B.Tech CSE under CSE/IT Discipline at Lovely Professional University, Punjab, hereby declare that all the information furnished in this project report is based on my own intensive work and is genuine.

Date: 12-04-2023

Registration No.: 12019256 Nirbhay Singh

ACKNOWLEDGEMENT

I would like to express my special thanks of gratitude to my teacher Mrs. Nidhi Arora who gave me the golden opportunity to do this wonderful project of analysis of the data of a netflix namely "*ANALYSIS OF NETFLIX TITLES*" which also helped me in doing a lot of research and I came to know about so many new things. I am thankful to them. Secondly, I would also like to thank my parents and friends who helped me a lot in finalizing this project within the limited time frame.

INDEX

INTRODUCTION

Data Analysis is a process of inspecting, cleansing, transforming, and modeling data with the goal of discovering useful information, informing conclusions, and supporting decisionmaking. Data analysis has multiple facets and approaches, encompassing diverse techniques under a variety of names, while being used in different business, science, and social science domains.

Netflix is one of the world's leading video streaming services, offering a wide variety of movies, TV shows, and documentaries to millions of viewers worldwide. As the platform continues to grow, so does the number of titles available to stream. With so much content to choose from, it can be difficult for users to find the perfect title to watch.
For this reason, the aim of this Tableau dashboard project is to analyse Netflix's vast library of titles and provide insights into their distribution across various categories, such as genre, release year, and country of origin. By visualizing this data, we can help users make more informed decisions about what to watch next and gain a deeper understanding of Netflix's content strategy.

The Netflix titles database keeps track of the following data fields:

- Show id – Unique ID given to every Movies and TV shows.
- Type – It specifies the type its Movie or TV show.
- title – title of movie or TV show.
- Date added – give the Release date on Netflix ott.
- Release year- That show movie or tv show Release year.
- Number of Vehicles- Total number of vehicles involved in the accident.
- Rating – Its shows what Rating get for particular movie or tv show
- Duration –Its gives the timing or season of movies or tv shows.
- Listed In – This shows distribution on the behave of casting of movie or tv show .
- Description – this gives the hint for which type of movie is this in summary form.

SCOPE OF ANALYSIS

The scope of analysis for the Netflix Titles topic in the Tableau dashboard project could include:

- Title Count by Country: Analyzing the number of titles available on Netflix by country of origin to gain insights into the streaming platform's global content distribution strategy.

- Genre Distribution: Examining the distribution of genres in the Netflix library to understand which types of content are most popular with viewers and which are underrepresented.

- Release Year: Analyzing the distribution of titles by their release year to identify trends in content production and understand how the platform has evolved over time.

- Content Rating: Exploring the distribution of titles by their content rating to gain insights into the audience demographic that Netflix is targeting.

- Runtime Distribution: Examining the distribution of titles by their runtime to understand the popularity of different types of content length.

- Popular Titles: Identifying the most popular titles on Netflix by analyzing user ratings, number of views, and other metrics.

- Director and Actor Analysis: Analyzing the distribution of titles by their directors and actors to understand the influence of these individuals on the platform's content.

By visualizing and analyzing these various aspects of the Netflix titles, we can create a comprehensive dashboard that provides valuable insights into the platform's content strategy and user preferences.

EXISTING SYSTEM

There are various existing systems for data analysis and visualization, including software like

Excel : helps to fetch the data from the user and store it in order.

Tableau Prep: helps to clean the data and to show only the relevant amount of data according to user preference.

Tableau Desktop: for virtualization and making interactive dashboards this software provides users with a simple interface using which they can visualize their data just by dragging and dropping.

These systems are designed to enable users to extract insights from data and communicate them effectively through visual representations. However, there are some drawbacks associated with these systems. For example,

- some systems may have limited functionality and may not support advanced data analysis techniques.
- Others may require significant technical expertise to use effectively, making them less accessible to users without specialized skills.
- some systems may have limited scalability, making it difficult to analyze large datasets or to support multiple users simultaneously.
- data security and privacy can be a concern with these systems, particularly when dealing with sensitive or confidential information.

As such, it is important to carefully evaluate existing systems for data analysis and visualization to ensure they meet the needs of the organization and address any potential drawbacks.

ANALYSIS OF DATASET

1. Total movies & tv shows by year

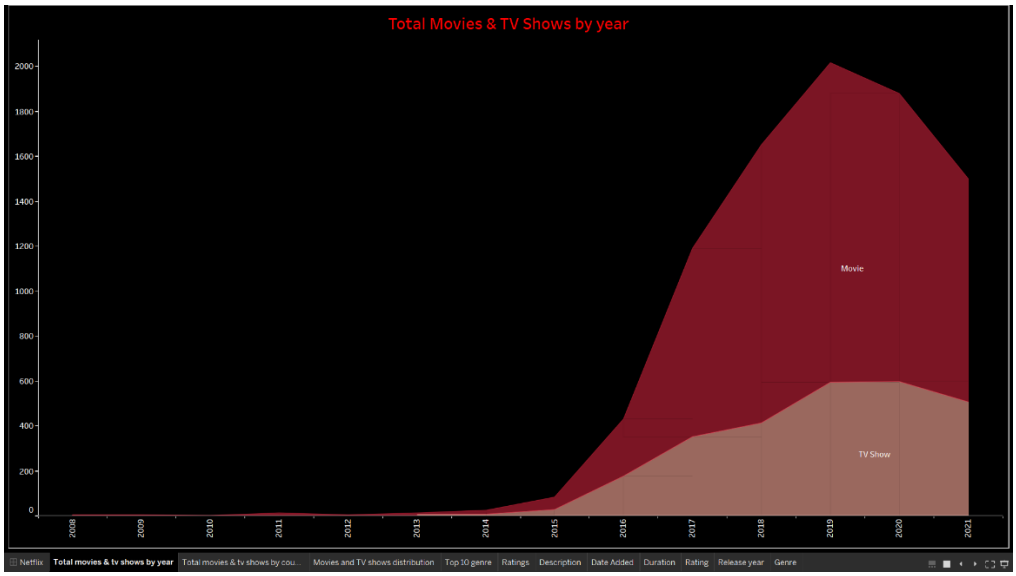
Description:

By knowing about the Movie and TV Shows Releasing over the month in a particular year we can know the month by month and finally year by year Release of Movie or TV shows which one is greater in amount.

After finding the Movies and TV shows we visualize the result with the help Area chart.

Visualization:

The results are then visualized in the form of a Area chart.



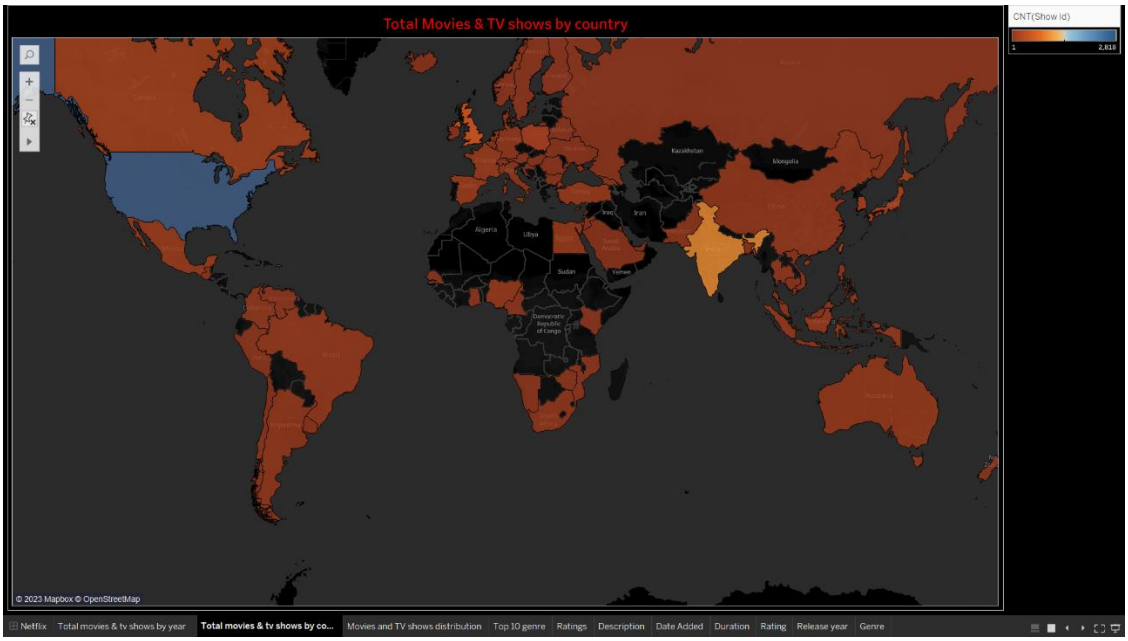
2. Total movies & tv shows by Country:

Description:

This analysis gives the total number of Count of show ID that were Available in the particular country. In this analysis we have also shown Colours wise distribution in a particular year.

Visualization:

We will use a map chart to show the trend.



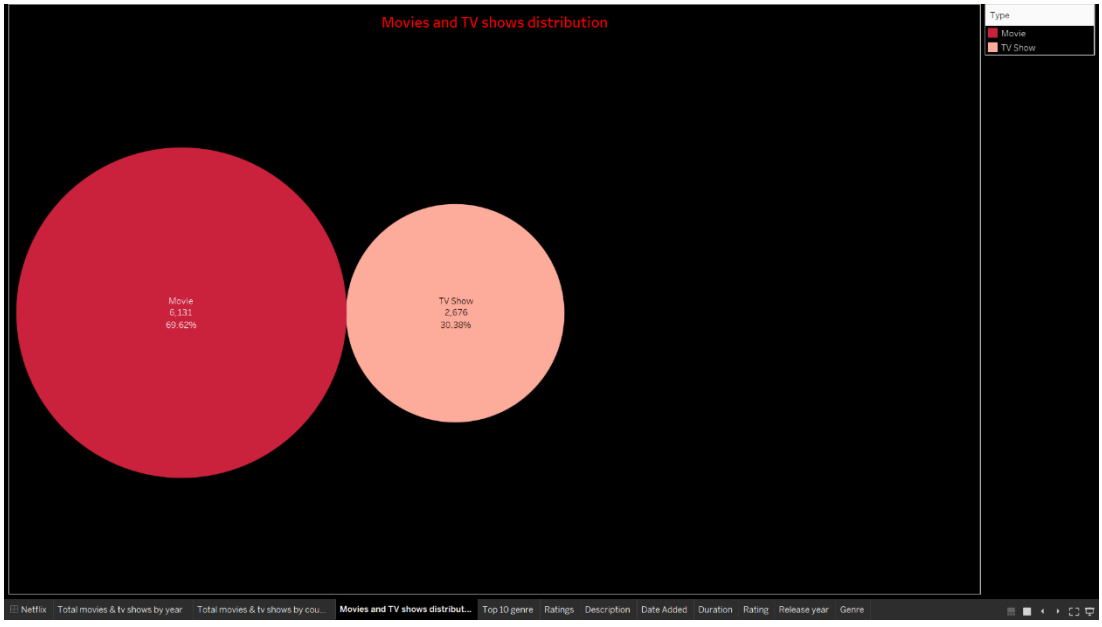
3. Total movies & tv shows Distribution :

Description:

This analysis gives the total count of shows ID that are in the category of Movies and TV shows that occurred in a particular year.

Visualization:

The results are visualized with the help of a packed buuble chart. Showing the trend over year



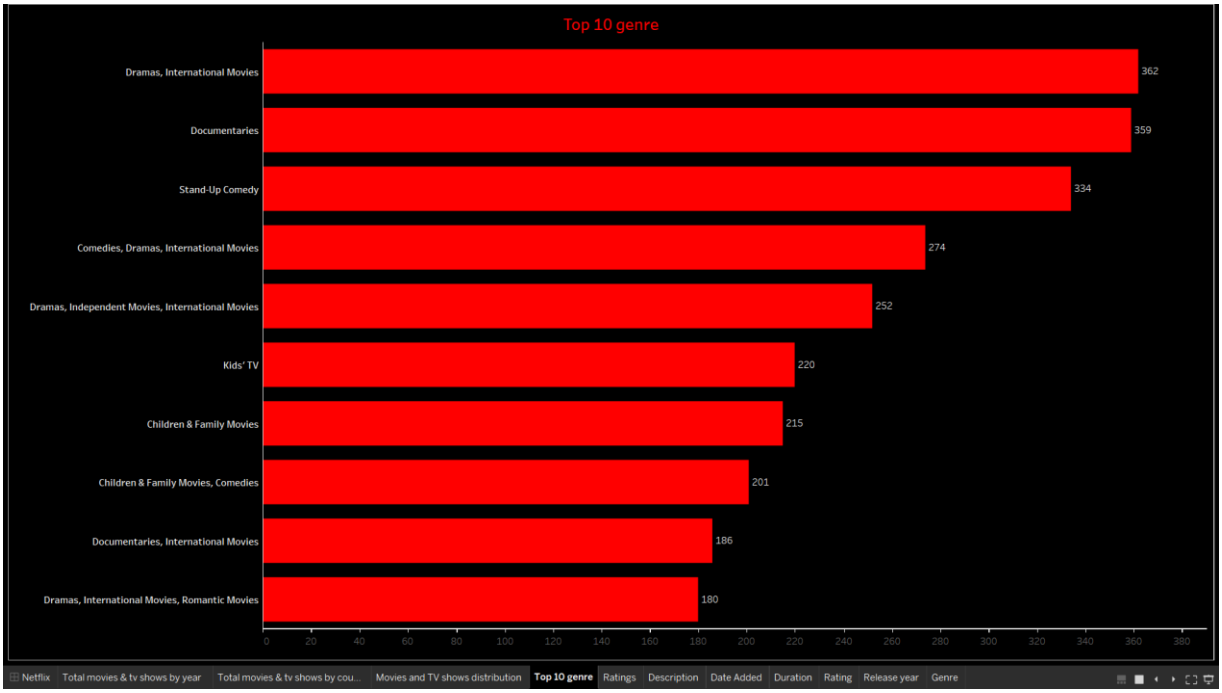
4. **Top 10 Genre:**

Description:

This analysis gives the Top 10 Genre depend on rating. In this we have use filter over Listed in for the top 10 values.

Visualization:

We visualize the above results with the shapes of Horizontal bar.



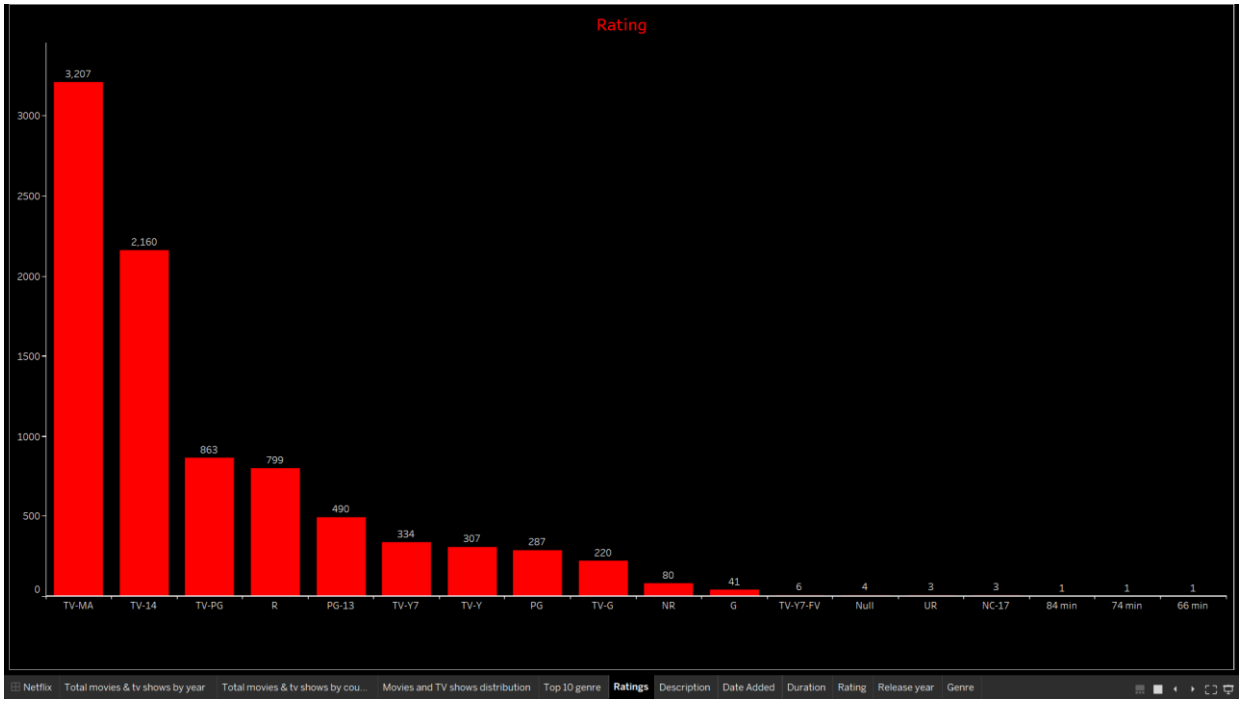
5. **Ratings:**

Description:

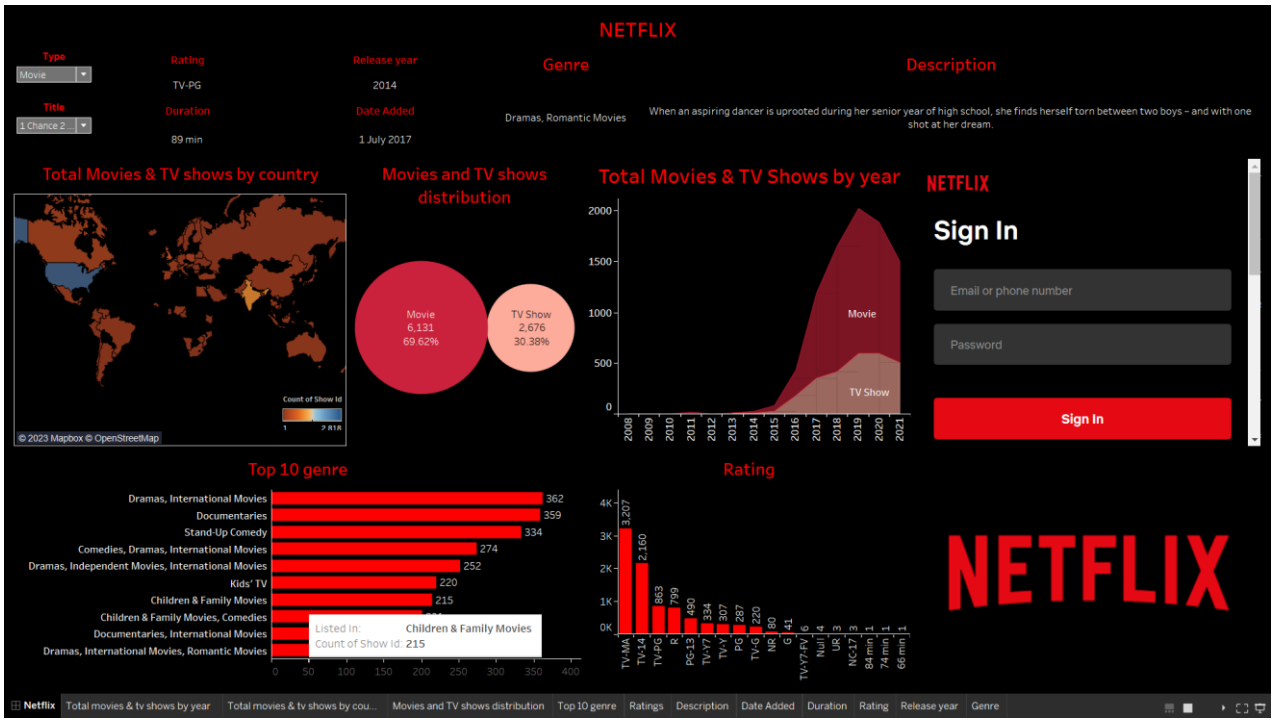
This analysis gives the Ratings of the count of show ID that happen in different Type of Listed in Name.

Visualization:

The results are visualized in the form of Vertical bar chart.



FINAL DASHBOARD



FUTURE SCOPE

To scope a future Netflix titles Tableau dashboard project, here are some potential areas to consider:

1. Content type: With the increase in original programming, it would be interesting to analyze the proportion of Netflix's content that is produced in-house versus licensed from other companies. This could involve creating a pie chart or stacked bar chart to show the breakdown of content by type.
2. Recommendation engine: Netflix's recommendation engine is one of its biggest selling points. It would be interesting to analyze how accurate the recommendations are and how much they influence what viewers watch. This could involve creating a heatmap or network graph that shows how often recommended titles are actually viewed.
3. Content acquisition and distribution: With the increasing competition among streaming services, it would be interesting to analyze Netflix's content acquisition and distribution strategies. This could involve creating a line chart or area chart that shows how much Netflix is spending on content acquisition, or a map that shows which regions have the most Netflix subscribers.
4. User behavior: It would be interesting to analyze user behavior on Netflix, such as how often they watch, which titles they watch most often, and how long they typically spend watching. This could involve creating a histogram or box-and-whisker plot that shows how long users typically spend watching a single title, or a bubble chart that shows which titles are most frequently watched together.
5. Content trends: It would be interesting to analyze trends in the types of content that Netflix produces or licenses, such as the popularity of certain genres or formats. This could involve creating a line chart or bar chart that shows how the popularity of different genres has changed over time.
6. Impact on the film and TV industry: With the rise of streaming services like Netflix, it would be interesting to analyze how these services are affecting the traditional film and TV industry. This could involve creating a comparison chart that shows how many viewers traditional networks are losing to streaming services, or a map that shows where Netflix is producing its original content.

REFERENCES AND BIBLIOGRAPHY

- YouTube (https://www.youtube.com/@datatutorials1)
- Coursera (https://www.coursera.org/learn/dataviz-dashboards)
- Kaggle(https://www.kaggle.com/datasets/shivamb/netflix-shows?resource=download)