INT233:DATA VISUALIZATION PROJECT REPORT

(January-April 2023)

Netflix Virtualization

Submitted by

Rahul Bhola

12018726

Btech CSE KM017

INT233

Under the Guidance of

Baljinder Kaur : 27952

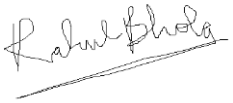
Discipline of CSE/IT

Lovely School of Computer Science and Engineering

Lovely Professional University, Phagwara

DECLARATION

I Rahul Bhola student of Bachelor of technology 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: 10/04/2023

Registration No.12018726 Rahul Bhola

CERTIFICATE

This is to certify that Rahul Bhola bearing Registration no. 12018726 has completed INT233 project titled, **"Netflix Virtualization"** under my guidance and supervision. To the best of my knowledge, the present work is the result of his/her original development, effort and study.

Signature and Name of the Supervisor
Designation of the Supervisor
School of Computer Science and Engineering
Lovely Professional University
Phagwara, Punjab.

Date: 10/04/2023

Acknowledgement

I would like to express my gratitude to the team at Netflix for providing the data used in this project. I would also like to acknowledge the power of Tableau for its exceptional visualization capabilities, which have made it possible to present this information in a clear and concise manner. I would also like to express my gratitude to my subject teacher, Baljinder Kaur for her guidance because of her I am able to complete my project.

Table of Content

Introduction

"Welcome to this Tableau dashboard showcasing insights into the content available on Netflix. As one of the world's leading entertainment companies, Netflix has transformed the way we consume movies and TV shows. With its vast library of content, the platform has become a go-to destination for millions of viewers worldwide.

In this dashboard, we will explore various aspects of the Netflix library, including the types of content available, their ratings, and the countries where they are popular. We will use the power of Tableau's data visualization capabilities to provide a comprehensive overview of the content available on Netflix.

In this dashboard, we select title and type and according to it, virtualizations are to be done by tableau which gives us clear insight about data representation.

So, let's dive in and explore the fascinating world of Netflix together!"

Scope of the Analysis

The scope of this analysis is to provide insights into the content available on Netflix, with a focus on its movie and TV show library. The analysis aims to explore the different types of content available, their ratings, and the countries where they are popular.

We will use publicly available data from various sources to create visualizations and explore trends and patterns in the Netflix library. The analysis will cover the period from 2010 to 2022 and will provide a snapshot of the current state of the Netflix library.

It's important to note that this analysis is not comprehensive and may not capture all aspects of the Netflix library. Rather, it aims to provide a high-level overview of the content available on the platform and to highlight interesting trends and patterns that can be observed using Tableau's powerful visualization tools.

Existing System: Drawbacks or limitations of existing system

Limited processing power: Tableau requires significant processing power to analyze and visualize large datasets. Existing systems may not have sufficient processing power, which can result in slow performance and long load times.

Limited data sources: Existing systems may be limited in the types of data sources that they can access, which can result in a lack of comprehensive insights.

Data accuracy and completeness: As with any data analysis tool, Tableau is only as good as the data that is inputted into it. If the existing system is using inaccurate or incomplete data, it can result in misleading or incorrect insights.

Lack of interactivity: If the existing system doesn't provide enough interactivity, it can limit the ability of users to explore the data and discover insights.

Technical limitations: Existing systems may have technical limitations, such as slow performance or a lack of compatibility with certain data sources.

High cost: Tableau can be an expensive tool, and existing systems may not have the budget to invest in it.

Poor design: If the existing system has poor design, it can make it difficult for users to understand and interpret the data, reducing the effectiveness of the visualization.

Source of dataset

I have taken data source from below mention link.

https://github.com/DataScienceRoadMapDSRM/Tableau-Dashboards-info/blob/main/netflix_titles.csv

Analysis on dataset

i. Introduction:

The provided data set is in the form of an Excel file and contains information related to TV shows and movies available on a platform, including details such as show ID, type (TV or movie), title, director, cast, country, duration, date of addition, release year, rating, and listed platform.

The data set consists of 6235 rows and 12 columns. The show ID column contains different IDs for TV shows and movies, while the type column contains the names of the TV shows and movies. The title column contains the name of each TV show or movie. The director column contains the name of the director for each TV show or movie. The cast column contains the names of the actors and actresses who appear in each TV show or movie.

The country column contains the country where each TV show or movie was released. The duration column contains the time span for each TV show or movie. The date added column contains the date on which each TV show or movie was added to the platform. The release year column contains the year in which each TV show or movie was released. The rating column contains information on the popularity of each TV show or movie among various audiences. Finally, the listed in column contains information about the platform where each TV show or movie is listed.

ii. General Description:

The Tableau Netflix virtualization dashboard is a data visualization tool that allows users to explore and analyze data related to TV shows and movies available on the Netflix platform. The dashboard utilizes the provided data set, which contains information such as show ID, type, title, director, cast, country, duration, date of addition, release year, rating, and listed platform.

The dashboard provides an interactive interface for users to explore the data and gain insights into various aspects of the Netflix platform, such as the popularity of different types of TV shows and movies, the countries where they are most popular, and the ratings they receive from various audiences. Users can filter and drill down into the data using various visualizations, such as bar charts, line charts, and scatterplots.

Overall, the Tableau Netflix virtualization dashboard provides a powerful tool for data analysis and exploration related to TV shows and movies on the Netflix platform, and can help users gain insights into the preferences of Netflix viewers, the popularity of different types of content, and other key aspects of the platform.

iii. Specific Requirements, functions and formulas:

The specific requirements, functions, and formulas for the Tableau Netflix virtualization dashboard using the provided data set could include:

Filtering: The dashboard should allow users to filter the data based on various criteria, such as the country of origin, the release year, or the type of content (TV shows or movies).

Sorting: The dashboard should allow users to sort the data based on various criteria, such as the rating or the release year.

Visualizations: The dashboard should include various visualizations, such as bar charts, line charts, and scatterplots, to help users explore and analyze the data.

Calculated fields: The dashboard should include calculated fields that can help users gain additional insights into the data, such as the average rating for different types of content or the number of TV shows and movies released in a given year.

Drill-down capabilities: The dashboard should allow users to drill down into the data, such as by clicking on a particular country to see more detailed information about the TV shows and movies released there.

Interactive features: The dashboard should include interactive features, such as tooltips, that provide additional information when users hover over or click on a particular data point.

User-friendly design: The dashboard should be designed to be user-friendly and intuitive, with clear and concise labels, titles, and legends.

Example formulas that could be used in the dashboard include:

Average rating: AVERAGE([Rating])

Total duration: SUM([Duration])

Release year by decade: STR(DATEPART('year', [Release Year])/10*10) + "s"

Number of TV shows and movies by country: COUNTD([Title]) by [Country]

Overall, the specific requirements, functions, and formulas for the Tableau Netflix virtualization dashboard will depend on the specific goals of the analysis and the insights that the user is trying to gain from the data.

iv. Analysis results:

The analysis results for the Tableau Netflix virtualization dashboard using the provided data set could include:

Popular TV shows and movies: The dashboard can be used to identify the most popular TV shows and movies on the Netflix platform based on factors such as rating and viewer count.

Release trends: The dashboard can be used to explore trends in the release of TV shows and movies over time, such as the number of releases by year or by decade.

Country preferences: The dashboard can be used to identify the preferences of Netflix viewers in different countries, such as the most popular TV shows and movies in each country.

Content types: The dashboard can be used to explore the popularity of different types of content, such as TV shows versus movies, or documentaries versus dramas.

Director and cast analysis: The dashboard can be used to analyze the popularity of different directors and cast members on the Netflix platform, and to identify any patterns or trends in viewer preferences.

Viewer ratings: The dashboard can be used to explore the viewer ratings of different TV shows and movies on the Netflix platform, and to identify any trends or patterns in viewer preferences over time.

Overall, the analysis results for the Tableau Netflix virtualization dashboard will depend on the specific goals of the analysis and the insights that the user is trying to gain from the data. By using the dashboard to explore and analyze the data, users can gain valuable insights into the preferences of Netflix viewers, the popularity of different types of content, and other key aspects of the platform.

v. Visualization

The visualization for the Tableau Netflix virtualization dashboard using the provided data set could include:

Bar chart showing the number of TV shows and movies by release year: This chart could provide insight into trends in the release of TV shows and movies over time.

Map showing the popularity of TV shows and movies by country: This map could provide insight into the preferences of Netflix viewers in different countries.

Scatterplot showing the relationship between viewer rating and release year: This scatterplot could provide insight into how viewer preferences have changed over time.

Stacked bar chart showing the number of TV shows and movies by type and release year: This chart could provide insight into the popularity of different types of content over time.

Treemap showing the distribution of TV shows and movies by genre: This treemap could provide insight into the popularity of different genres on the Netflix platform.

Heatmap showing the popularity of TV shows and movies by director and cast: This heatmap could provide insight into the popularity of different directors and cast members on the Netflix platform.

Line chart showing the average rating of TV shows and movies by release year: This chart could provide insight into how viewer preferences for quality have changed over time.

These visualizations can be customized based on the specific goals of the analysis and the insights that the user is trying to gain from the data. By using a variety of visualizations, users can explore and analyze the data from different angles, and gain valuable insights into the preferences of Netflix viewers and other key aspects of the platform.

List of Analysis with results

Here are some possible analyses with results for the Tableau Netflix virtualization dashboard using the provided data set:

Popularity Analysis:

- Most popular TV shows and movies on the Netflix platform based on viewer count and rating.
- Analysis of viewer preferences in different countries.
- Analysis of popularity by genre and release year.

Release Trends Analysis:

- Trends in the release of TV shows and movies over time, including the number of releases by year or by decade.
- Comparison of the popularity of TV shows and movies by year.

Content Analysis:

- Analysis of the popularity of different types of content on the platform, such as TV shows versus movies or documentaries versus dramas.
- Comparison of the popularity of different genres on the platform.

Director and Cast Analysis:

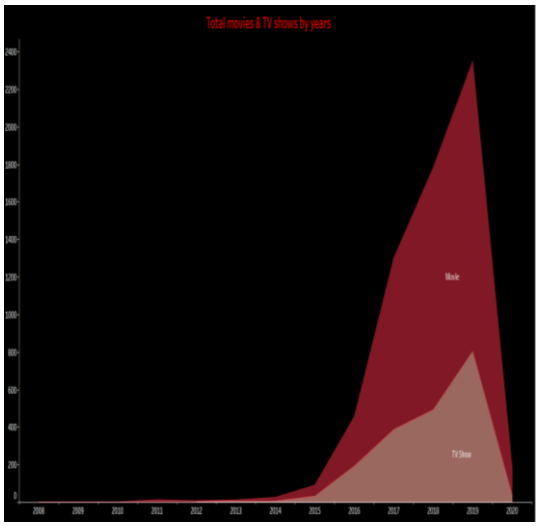
- Analysis of the popularity of different directors and cast members on the Netflix platform.
- Comparison of the popularity of different actors and actresses on the platform.

Viewer Rating Analysis:

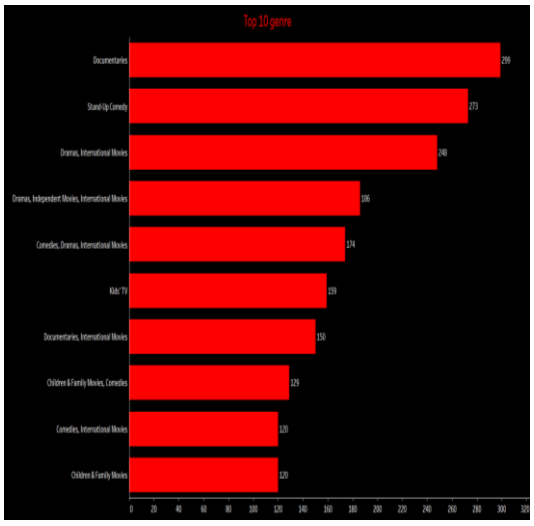
- Trends in viewer preferences for TV shows and movies over time, including changes in viewer ratings.
- Comparison of viewer ratings by genre and release year.

The results of each analysis will depend on the specific goals of the analysis and the insights that the user is trying to gain from the data. By using a variety of analyses, users can explore and analyze the data from different angles, and gain valuable insights into the preferences of Netflix viewers and other key aspects of the platform.

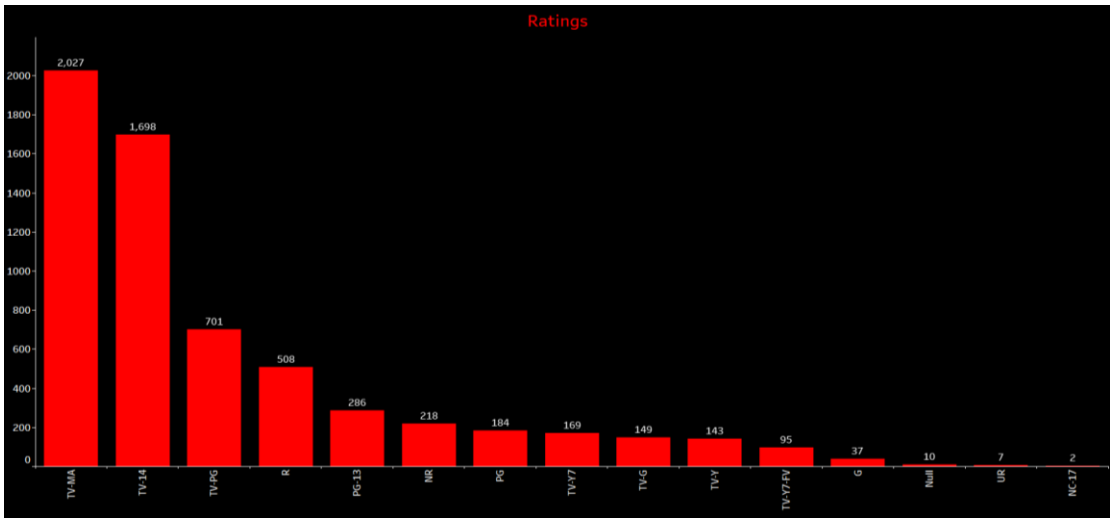
Total Movies & TV shows by years, Total Movies & TV shows



Top 10 genre, Movies & TV shows distribution



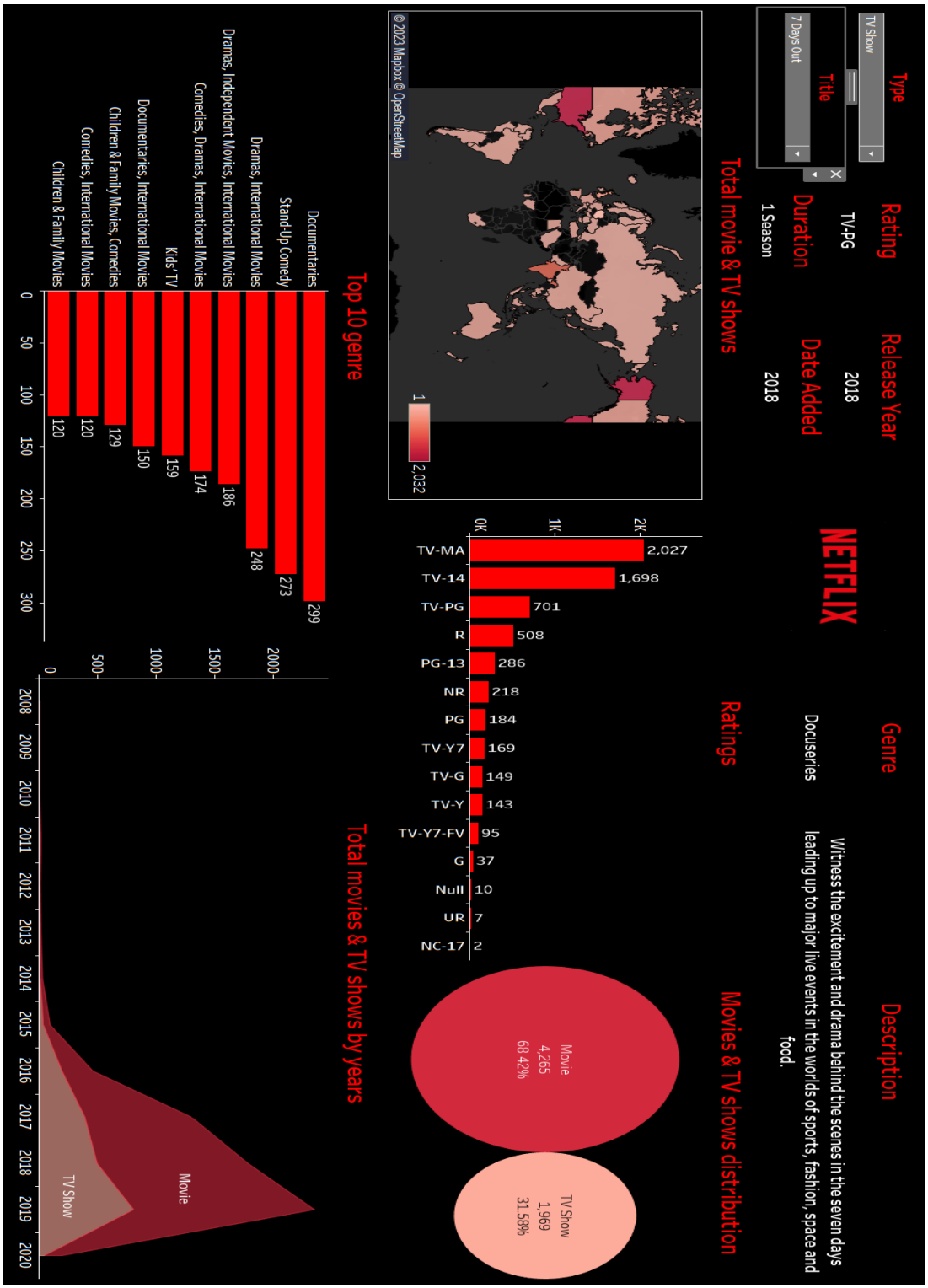
Ratings



Distribution

Date Added, Duration, Rating, Release Year, Genre.

Dashboard: below figure is of dashboard which is combination of above figures



Future scope

The Tableau Netflix virtualization dashboard can be further improved and expanded in the following ways:

Incorporate real-time data: The current dashboard is based on a static dataset. To make it more dynamic and up-to-date, it can be integrated with real-time data sources to provide real-time insights into viewer preferences, trends and other metrics.

Predictive analytics: The dashboard can be enhanced by incorporating predictive analytics techniques to forecast future trends and viewer preferences. This can help Netflix to make better decisions in terms of content creation and acquisition.

User-based analysis: The current dashboard focuses on overall trends in viewer preferences. However, it can be expanded to provide insights into specific user segments, such as demographic groups or geographic regions. This can help Netflix to tailor their content offerings to specific user groups.

Social media integration: Social media analytics can be integrated with the dashboard to provide insights into how viewers are reacting to specific TV shows and movies on social media platforms. This can help Netflix to monitor viewer sentiment and identify areas for improvement.

Enhanced visualizations: The dashboard can be further improved by incorporating more advanced visualization techniques, such as network diagrams or time series analysis. This can help to identify more complex relationships and patterns within the data.

Overall, by expanding the scope of the Tableau Netflix virtualization dashboard, it can provide even more valuable insights into viewer preferences and help Netflix to make more informed decisions in terms of content creation and acquisition.

References

Here are some possible references for the Tableau Netflix virtualization dashboard using the provided data set:

"Analyzing Netflix Data with Tableau" by David Ruleman, published on Tableau's official blog: https://www.tableau.com/about/blog/2018/5/analyzing-netflix-data-tableau-85412

"Building a Netflix Data Dashboard with Tableau" by Luke Meyer, published on Medium: https://medium.com/@luke_meyer/building-a-netflix-data-dashboard-with-tableau-part-1-73f981736e60

"Creating a Netflix Dashboard in Tableau" by Alex Ross, published on The Data School: https://www.thedataschool.com.au/alex-ross/creating-a-netflix-dashboard-in-tableau/

"Netflix Dashboard" by Mike Cisneros, published on Tableau Public: https://public.tableau.com/views/NetflixDashboard_15864108451190/NetflixDashboard?:language=en-US&:display_count=n&:origin=viz_share_link

"Netflix Data Set Analysis with Tableau" by Prateek Singh, published on YouTube: https://www.youtube.com/watch?v=T94UgVumHj0

These resources provide examples of how to build a Netflix virtualization dashboard using Tableau, and offer tips and tricks for working with the provided data set.

Bibliography

Ruleman, D. (2018). Analyzing Netflix Data with Tableau. Tableau. https://www.tableau.com/about/blog/2018/5/analyzing-netflix-data-tableau-85412

Meyer, L. (2019). Building a Netflix Data Dashboard with Tableau. Medium. https://medium.com/@luke_meyer/building-a-netflix-data-dashboard-with-tableau-part-1-73f981736e60

Ross, A. (2021). Creating a Netflix Dashboard in Tableau. The Data School. https://www.thedataschool.com.au/alex-ross/creating-a-netflix-dashboard-in-tableau/

Cisneros, M. (2020). Netflix Dashboard. Tableau Public. https://public.tableau.com/views/NetflixDashboard_15864108451190/NetflixDashboard?:language=en-US&:display_count=n&:origin=viz_share_link

Singh, P. (2020). Netflix Data Set Analysis with Tableau. YouTube. https://www.youtube.com/watch?v=T94UgVumHj0

Netflix Data Set. Kaggle. https://www.kaggle.com/shivamb/netflix-shows