Dynamic Netflix Dashboard Report

Course Title: **DATA VISUALIZATION** | Course Code: **INT233**

Submitted by:

Sushant Kumar

Registration No: 12218023

Roll Number 53

Programme Name: B.Tech. CSE

Under the Guidance of

Ma'am Maneet Kaur

School of Computer Science & Engineering

Lovely Professional University, Phagwara

DECLARATION

I, **Sushant**, declare that this program, titled **Skill Based Assignment**, is solely prepared by me as a requirement for **INT:233** at Lovely Professional University. The program is based on my original work and represents my own efforts, ideas, and analysis.

I further declare that the guidance provided by my instructor, **Maneet Kaur** , was instrumental in shaping my understanding and analysis of the subject matter. However, the final program and its contents are my own and do not reflect the views of Lovely Professional University or any other organization.

I take full responsibility for the accuracy and authenticity of the information presented in this program. I have cited all sources and references used in this program in accordance with the guidelines provided by my instructor and the academic standards of Lovely Professional University.

I understand that any act of plagiarism or academic misconduct may result in severe consequences, including the revocation of my academic degree, and I affirm that I have not engaged in any such activities.

Date: 15 /11/2024

Name of Student: Sushant Kumar
Registration no: 12218023

Indexing

Pages Name
1. Introduction
2. Scope of the Analysis
3. Existing System – Drawbacks | Limitations
4. Source of Dataset
5. ELT Extract, Transform, and Load
6. Analysis On Dataset | Visuals | functions | Formulas
7. Dashboard Overview
8. List of Analysis with Results
9. Future Scope
10. References
11. Bibliography

Introduction

In the dynamic world of online streaming, Netflix has emerged as a dominant player, continually evolving to meet the diverse preferences of its global audience. This report unveils a meticulously designed Tableau dashboard that leverages the expansive Netflix dataset to offer a detailed exploration of content distribution, genre prevalence, and viewer engagement across different regions. The primary aim of this dashboard is to provide stakeholders with actionable insights that support strategic decision-making in content management and market expansion.

The advent of data-driven strategies in entertainment has revolutionized how content is curated and delivered. Recognizing the significance of such insights, this dashboard captures a comprehensive snapshot of Netflix's offerings, detailing the intricacies of content types, viewer ratings, and the geographical distribution of movies and TV shows.

Through a series of interactive visualizations, the dashboard highlights key aspects of the Netflix catalog, including:

- **Geographical Content Distribution**: A world map visualization that delineates the availability of titles across different countries, providing a clear view of Netflix's global outreach and regional content strategies.
- **Content Ratings Analysis**: A vertical bar chart that categorizes the number of titles by their maturity rating, offering insights into the target demographics and content appropriateness.
- **Content Format Breakdown**: A pie chart that contrasts the proportion of movies versus TV shows, shedding light on the content format preferences among Netflix's audience.
- **Genre Popularity**: A horizontal bar chart listing the top 10 genres, which illustrates viewer preferences and highlights potential areas for content development.
- **Historical Content Trends**: An area chart that traces the volume of content added over the years, reflecting Netflix's growth trajectory and responsiveness to market dynamics.

The integration of these visualizations into a single dashboard provides a holistic view of Netflix's content strategy. It not only enhances the understanding of current offerings but also identifies trends that could dictate future acquisitions and productions. The dashboard's interactive filters for type, rating, release year, and more, allow for a granular analysis that can be customized to specific investigative needs, making it an invaluable resource for content strategists aiming to optimize the Netflix catalog.

Scope of Analysis

The scope of the analysis for this Tableau dashboard is to provide a comprehensive exploration of the Netflix content library, focusing on assessing the distribution and diversity of titles across different metrics such as geographical reach, content type, genre popularity, and temporal trends. The objective is to enable stakeholders at Netflix and associated industries to gain actionable insights into the platform's strategic content decisions, market presence, and audience engagement. By visualizing data on a granular scale, the dashboard seeks to identify patterns and trends that could inform decisions on content curation, marketing strategies, and customer retention efforts.

This analysis will utilize interactive visualizations to delve into several key areas:

- **Geographic Distribution of Content**: Understanding which countries have the most extensive libraries can help Netflix tailor its content to suit regional preferences and regulatory environments.
- **Content Ratings and Viewer Demographics**: By analyzing the distribution of content ratings, Netflix can better align its offerings with the viewer's age and preferences, ensuring compliance with regional censorship laws.
- **Genre Analysis and Trend Identification**: Identifying which genres are most popular and tracking changes in genre popularity over time can guide Netflix in developing new content that resonates with current viewer trends.
- **Temporal Distribution of Content Additions**: Observing how the volume of content has expanded over the years will provide insights into Netflix's growth strategies and response to competitive pressures in the streaming industry.

Each aspect of the analysis is designed to provide a layer of understanding that supports strategic business decisions. By leveraging the interactive capabilities of Tableau, the dashboard not only offers a snapshot of current data but also allows users to explore historical data trends and forecast future patterns. This dynamic approach to data analysis is pivotal in helping Netflix maintain its leadership in the highly competitive streaming service market, ensuring that its content strategy is both proactive and responsive to shifting consumer dynamics.

Existing System

The existing systems for analyzing streaming content libraries often involve static reporting and basic descriptive analytics, which do not provide the depth or interactivity required for strategic decision-making in a rapidly evolving media landscape.. This approach often results in several critical limitations:

1. **Lack of Interactivity**: Traditional systems generally offer static reports that do not allow users to interact with the data, explore different scenarios, or customize views according to specific needs or questions. This can hinder the ability to perform ad-hoc analysis and make real-time decisions based on the latest data.
2. **Scalability Issues**: As data volumes grow, the older systems struggle to handle large datasets efficiently. This can lead to significant delays in loading and processing data, which in turn can impede the timeliness and relevance of the insights generated. For a company like Netflix, which continuously expands its content library and subscriber base, scalability of the analytics platform is crucial.
3. **Insufficient Visualization Capabilities**: Many existing tools lack advanced visualization capabilities, making it difficult to represent complex datasets in an intuitive and insightful manner. Effective visualization is key to understanding nuanced patterns such as geographical distribution of content, genre preferences across different demographics, and temporal trends in content consumption.
4. **Limited Data Integration**: Integrating data from multiple sources—such as viewer ratings, content metadata, and subscriber information—can be challenging in traditional systems. This limitation restricts the depth of analysis possible, as it prevents a holistic view of the data that could unveil more comprehensive insights into user behavior and content performance.
5. **Delayed Decision-Making**: With static reports, any new query or requirement for analysis has to go through a time-consuming process of data extraction, transformation, and loading, followed by report generation. This delay can be a significant disadvantage in making timely decisions, especially in a fast-paced industry where content trends and viewer preferences can shift rapidly.

The tableau dashboard developed for this analysis seeks to overcome these drawbacks by offering a dynamic, interactive, and scalable solution. It enables real-time data exploration with advanced visualization tools that can handle large volumes of data seamlessly, integrating diverse data sources for a holistic analysis. This system not only enhances the efficiency of data analysis processes but also empowers decision-makers with the ability to derive timely, actionable insights tailored to specific strategic questions.

Source of Dataset

The dataset was sourced from Kaggle, a widely recognized platform for data science competitions and collaborative projects, renowned for its diverse and extensive datasets.

1. **Dataset Overview**:
2. **Number of Rows**: The dataset contains approximately 7,800 entries.
3. **Number of Columns**: There are 12 columns in the dataset.
4. **Types of Data Contained**:
5. **Title**: The name of the movie or TV series.
6. **Director**: Names of the director(s) of the film or series.
7. **Cast**: A list of the main actors featured in the film or series.
8. **Country**: The country or countries where the movie or series was produced.
9. **Date Added**: The specific date when the title was added to the Netflix platform.
10. **Release Year**: The year when the title was originally released.
11. **Rating**: The official content rating assigned to the title, which helps viewers determine age appropriateness.
12. **Duration**: The length of the movie in minutes or the number of seasons for TV series.
13. **Listed In**: The genres or categories under which the title is listed on Netflix.
14. **Description**: A brief summary of the plot or significant themes of the title.
15. **Utility of Dataset**:

This dataset provides a comprehensive snapshot of the range of content available on Netflix, equipping analysts with the necessary data to explore trends in viewer preferences, content distribution across different regions, and changes in Netflix's catalog over time. It is a crucial resource for generating dynamic and insightful visualizations within the Tableau dashboard. The attributes included in the dataset facilitate in-depth analyses of content types, popularity trends, and strategic insights, which are vital for Netflix to adapt to market dynamics and enhance viewer engagement.

ELT

The ELT (Extract, Transform, and Load) process for the Netflix dashboard started with me extracting a comprehensive dataset from Kaggle, which offered a deep dive into Netflix titles. After securing the data, I moved into the transformation phase, where I undertook several steps to refine the dataset for thorough analysis:

- I filtered out incomplete records, specifically those missing essential information like 'Country' or 'Release Year'.
- I standardized the date formats across the 'Date Added' column to maintain consistency throughout the dataset.
- I tackled null values strategically, removing or imputing them based on their relevance to the overall analysis to ensure data integrity.
- I parsed the 'Duration' field to clearly distinguish between movies and TV shows, a critical factor for subsequent content-type analysis.
- I also introduced new columns such as 'Month Added' from the 'Date Added' data to facilitate more detailed time-series analysis.

After transforming the data, I loaded it into Tableau. Using Tableau’s robust functionalities, I manipulated the data to develop dynamic and interactive visualizations. This process was instrumental in customizing the dashboard to align with the specific requirements of our stakeholders and to provide clear insights into the trends and distribution patterns in Netflix's extensive content library. This meticulous ELT process was essential in crafting a responsive and insightful dashboard, significantly enhancing the visualization and comprehension of the data.

Analysis on Datasets

Analysis Techniques for the Netflix Dashboard

1. Filters

To enhance the interactivity and analytical depth of the dashboard, several filters were implemented:

- **Content Type Filter**: This filter allows users to select and view data specifically for Movies, TV Shows, or both, enriching the user experience by tailoring the data displayed according to the viewer's interest.
- **Date Filters**: Filters such as 'Release Year' and 'Date Added' provide insights into content trends over time, allowing stakeholders to analyze patterns based on when content was released or added to the platform.
- **Genre Filter**: This enables the isolation of data for specific genres, supporting focused analyses on genre performance and viewer preferences.

2. Calculated Fields

Calculated fields were created to derive new data points essential for targeted visualizations:

- **Duration Parsing**: This categorization was done using an IF-ELSE formula to classify content into "Short", "Medium", or "Long" based on the 'Duration' field:

plaintext

Copy code

```
IF [Duration] <= 90 THEN 'Short'
ELSEIF [Duration] <= 150 THEN 'Medium'
ELSE 'Long'
END
```

- **Content Age Category**: Content was categorized into 'New' or 'Old' to facilitate trend analysis over time, using the following calculation:

plaintext

Copy code

```
IF YEAR([Date Added]) >= YEAR(TODAY())-1 THEN 'New'
ELSE 'Old'
END
```

3. Data Aggregation and Summarization

To summarize the data effectively for visual metrics and comparisons, several aggregation functions were used:

- **Count of Titles by Country**: Implemented using COUNTD to determine the unique number of titles available per country, highlighting geographical distribution.

- **Average Ratings Calculation**: The AVG function was employed to compute the average viewer ratings across different content types or genres, aiding in quality assessment.

4. Time-Series Analysis

Time-series analysis was conducted to observe content growth and patterns:

- **Yearly Content Addition**: Titles added each year were aggregated using the YEAR function applied to 'Date Added', and results were plotted over time to trace growth trends.

5. Geographical Analysis

To analyze content availability by country:

- **Mapping Function**: Tableau's built-in geographic roles were utilized to map each 'Country' entry to its geographic location, enhancing the geographical data representation on the dashboard.

6. Visual Calculations for Dashboards

Advanced visual calculations were used to enhance user engagement:

- **Percentage Distribution of Content Types**: This was calculated to visually compare the proportion of Movies vs. TV Shows.
- **Top Genres Analysis**: The RANK() function was used in conjunction with filters to display only the top 10 genres, making the genre analysis more focused and actionable.

7. Dynamic Titles and Tooltips

Dynamic titles and tooltips were implemented to provide contextual information dynamically:

- **Dynamic Titles**: Titles such as "Top Genres in [Country]" were generated to reflect the filtered data dynamically.
- **Tooltips**: Enhanced tooltips provide detailed information like title, director, and cast upon hover, offering a deeper insight into the content without cluttering the visual space.

Datasets Overview

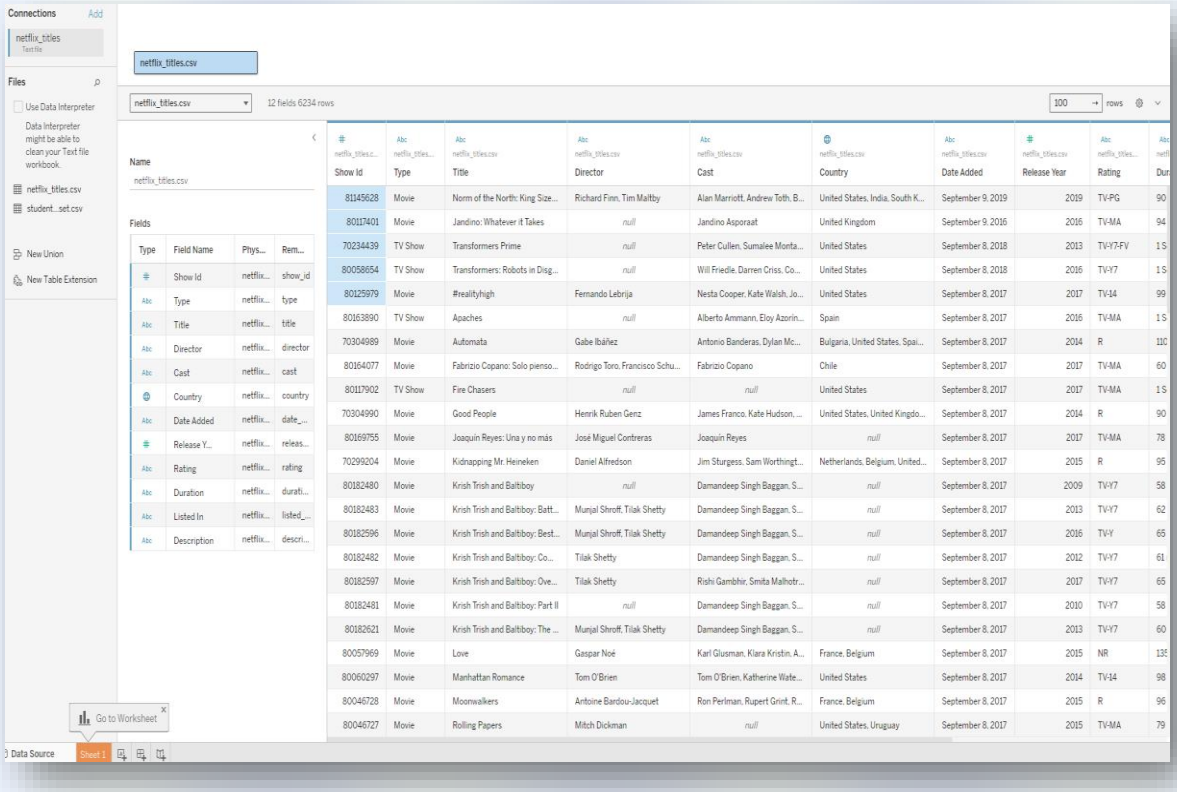
The dataset utilized for this Tableau dashboard project is derived from a comprehensive collection of Netflix titles publicly available and collated into a structured format. The dataset includes an extensive array of data points covering the vast array of content available on Netflix, such as movies and TV shows from various countries and spanning multiple genres. Each record in the dataset provides detailed attributes including the title name, release year, rating, duration, listed genres, and more. This information is crucial for dissecting the trends and patterns in Netflix's content strategy.

Key attributes within the dataset include:

- **Title**: The name of the movie or TV show.
- **Director**: The director(s) of the film or series episode.
- **Cast**: List of main actors involved.
- **Country**: The country or countries where the movie or TV show was produced.
- **Date Added**: The date the title was added to the Netflix library.
- **Release Year**: The official release year of the title.
- **Rating**: The age certification or general audience rating of the title.
- **Duration**: The length of the movie or number of seasons for TV shows.
- **Listed In**: The genres or categories under which the title is listed.
- **Description**: A brief synopsis of the title.

The dataset was sourced from Kaggle, ensuring it contains up-to-date and relevant information reflective of Netflix's current catalog. This rich dataset serves as the foundation for all analyses conducted through the dashboard, allowing for a detailed examination of the diversity and dynamics of Netflix's offerings. With this dataset, the dashboard explores various dimensions such as geographic content availability, popularity trends across different genres, content rating distributions, and temporal changes in Netflix’s content strategy.

Snapshot


Connections
Add
netflix_titles
Text file
Files
Use Data Interpreter
Data Interpreter might be able to clean your Text file workbook.
netflix_titles.csv
student...set.csv
New Union
New Table Extension
netflix_titles.csv
netflix_titles.csv
12 fields 6234 rows
100
rows
Name
netflix_titles.csv
Fields
Type
Field Name
Phys...
Rem...
Show Id
Type
Title
Director
Cast
Country
Date Added
Release Year
Rating
Go to Worksheet
Data Source
Sheet 1

Dashboard Overview

Main Dashboard

MapsCharts

Background Theme

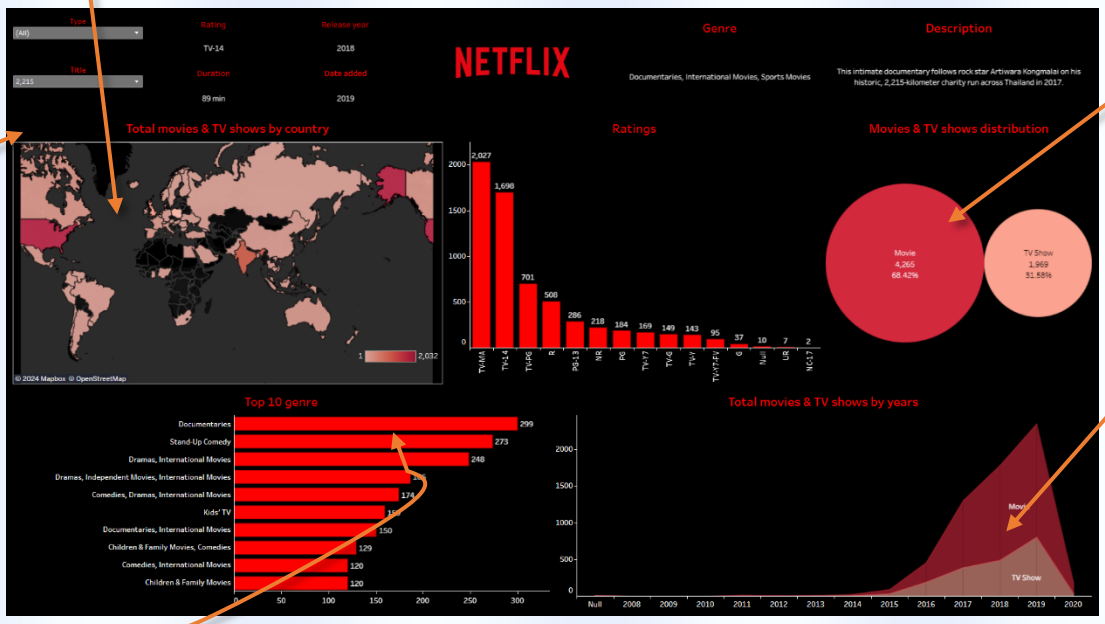
Bubble Chart

Area Graph

Descriptions

Horizontals Bars Showing Genres

Ratings

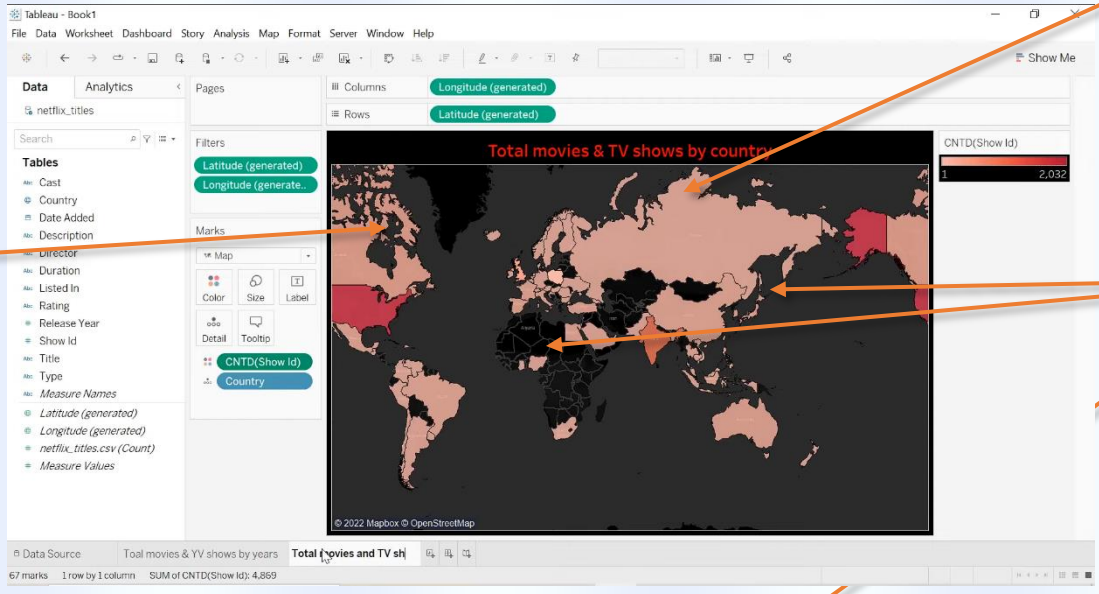


Implementation Of MapBox -

Plotting the Measures and Dimensions

Countries

Implementing Details

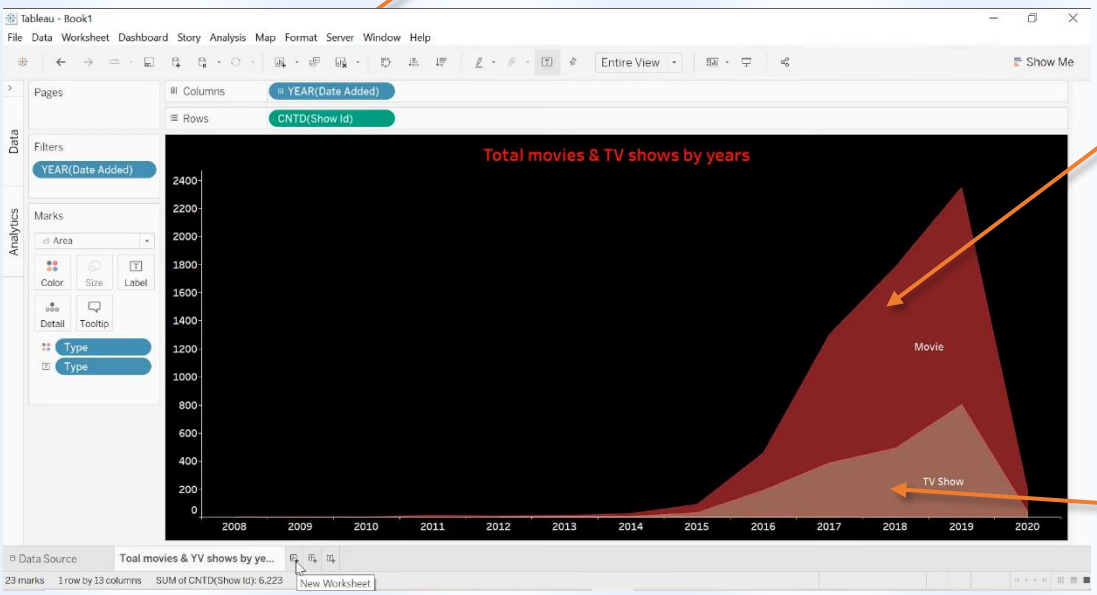


Labels

Applying Colour to differentiate

Movies

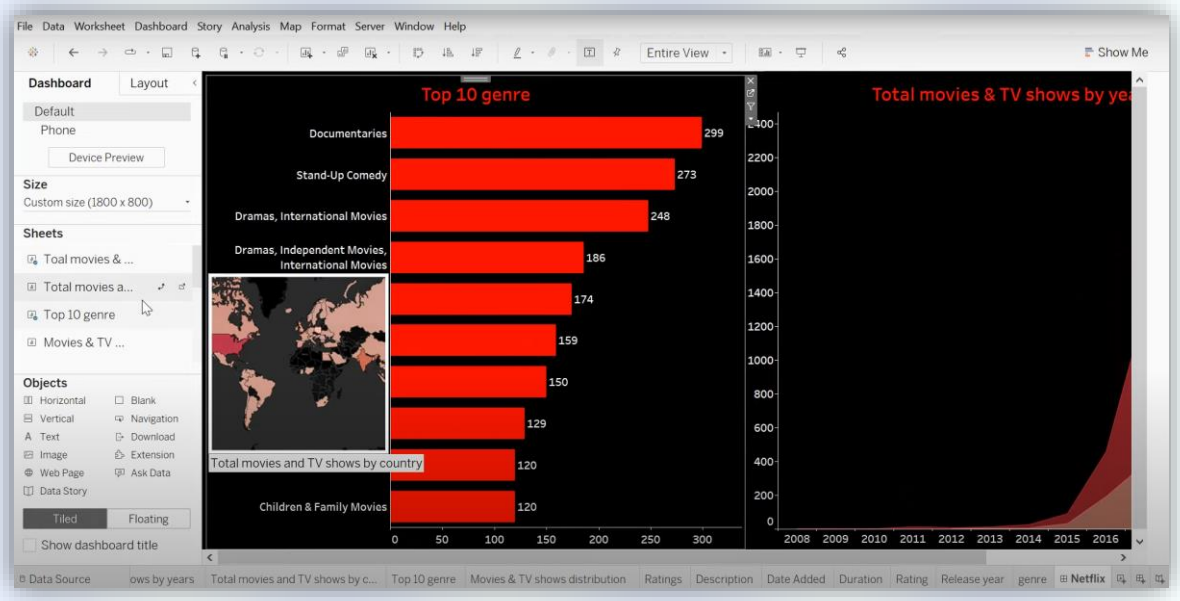
Area Chart



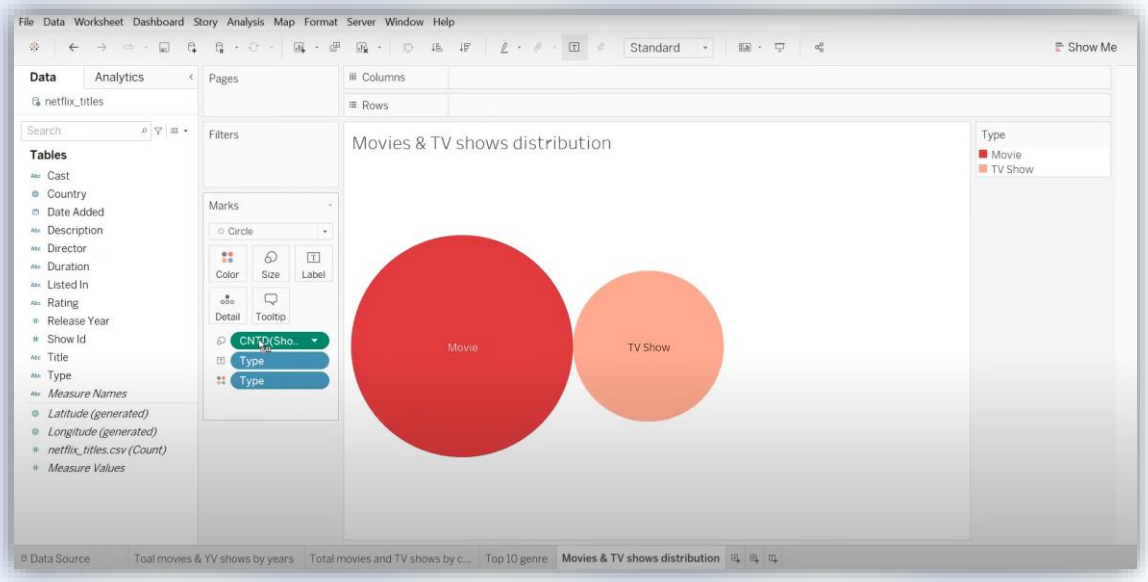
TV Shows

Dashboard Overviews

Screen 1

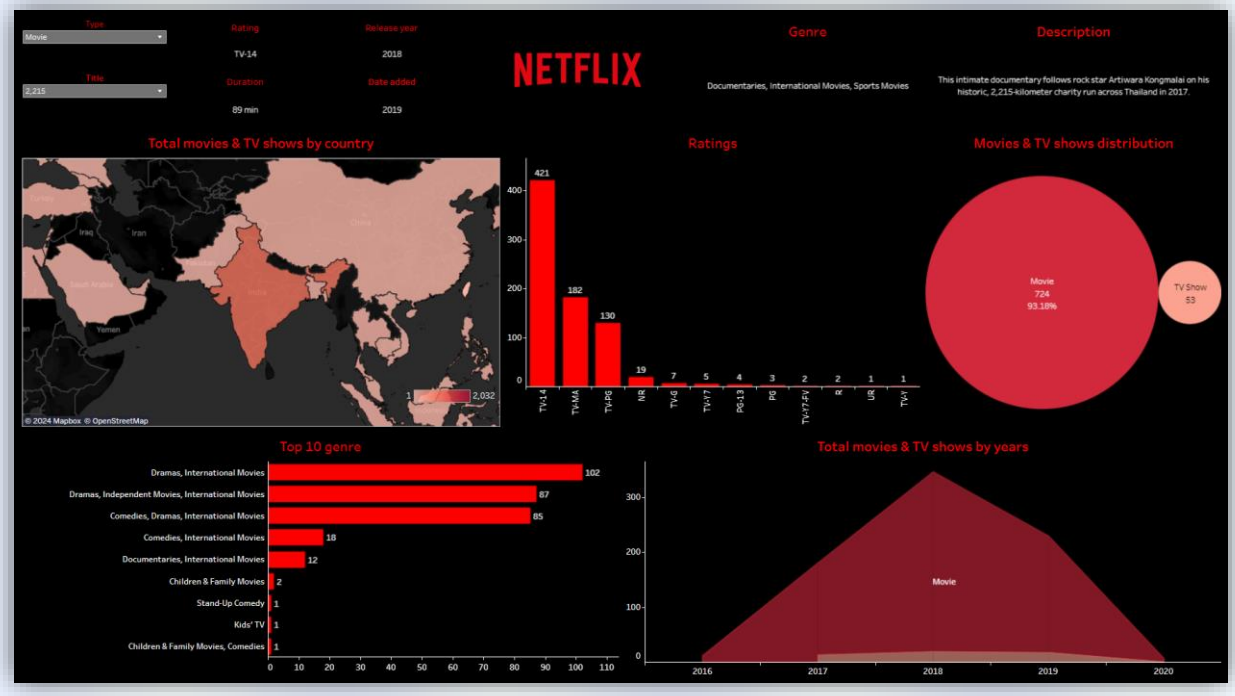


Screen 2

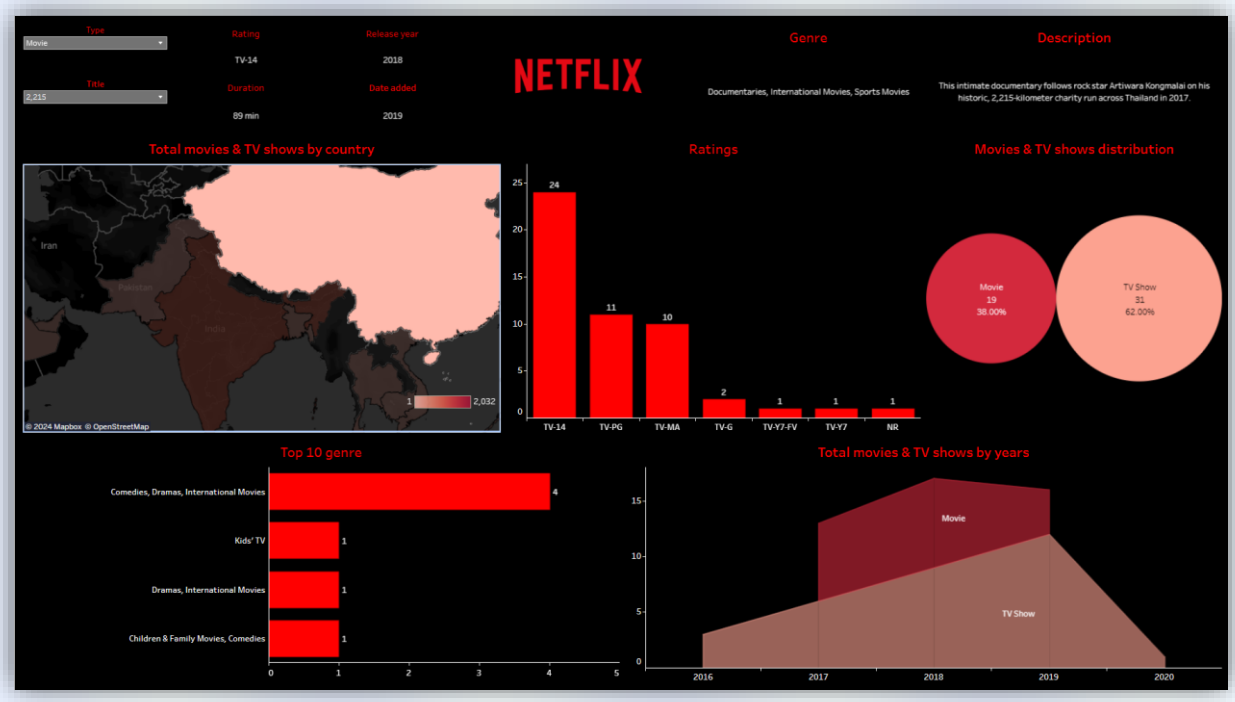


Dashboard Overviews

Filtration



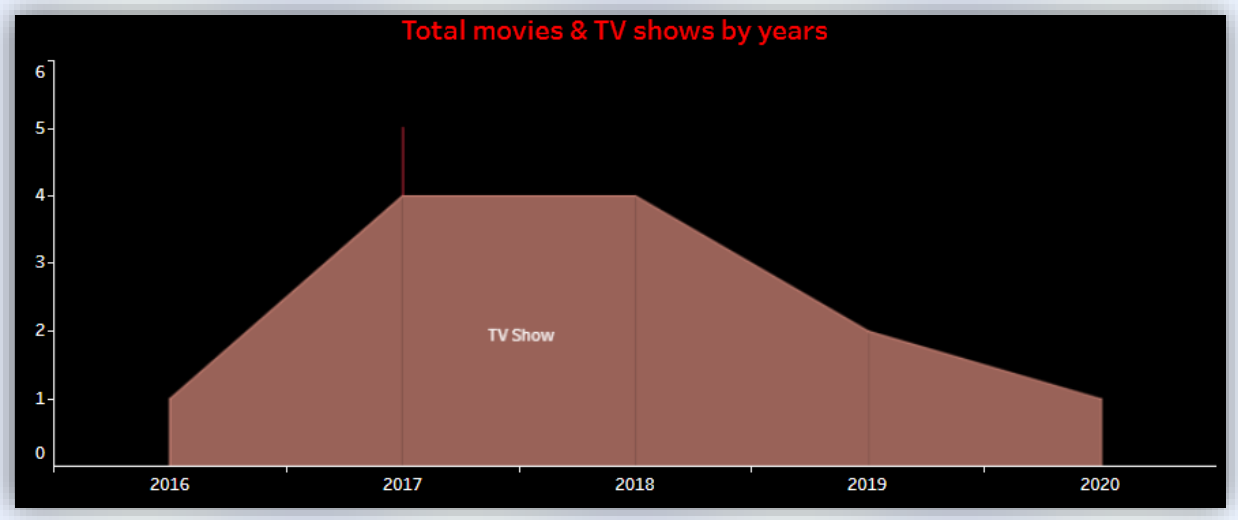
Motions



List of Analysis with Results

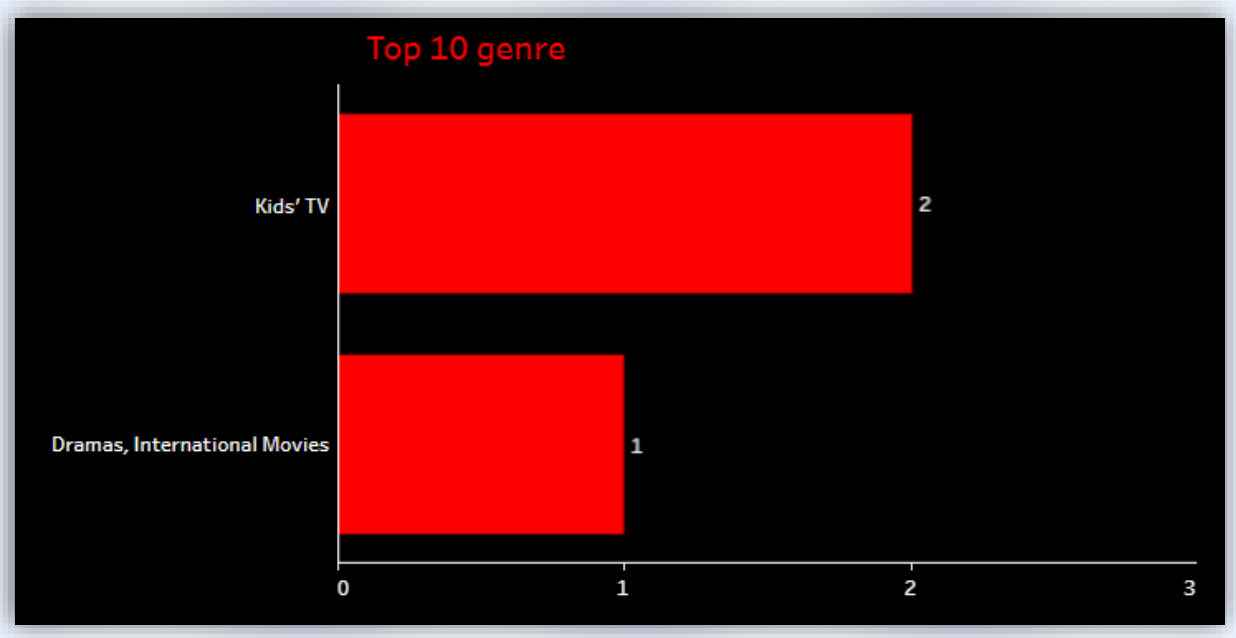
Insights

Total movies & TV Shows by years

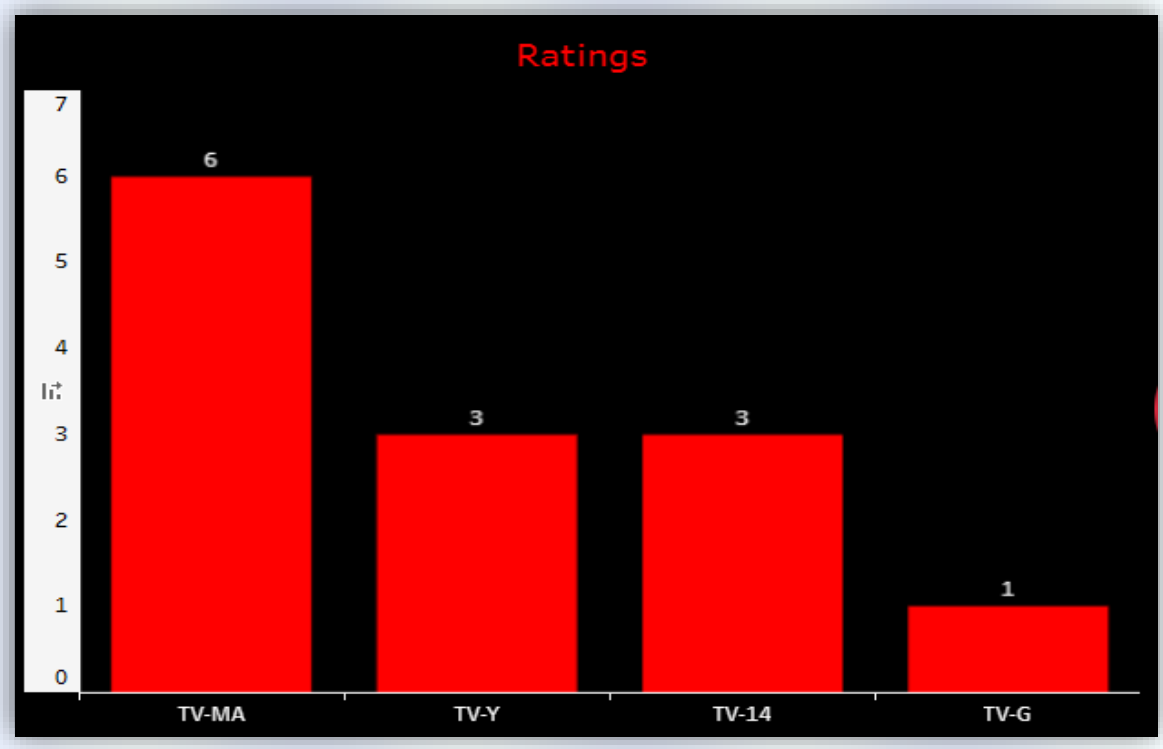


Dashboard Overviews

Top 10 Genre



Ratings



List of Analysis with Results

List of Analysis with Results

1. Total Movies & TV Shows by Country

- **Analysis**: The geographical distribution map visualizes the total number of movies and TV shows available in different countries.
- **Results**: The map highlights that the United States, India, and the United Kingdom have the highest volume of titles. This suggests a strong market presence in these regions, potentially aligning with higher subscriber bases.

2. Ratings Distribution

- **Analysis**: The ratings bar chart provides an overview of the number of titles categorized by their maturity ratings (e.g., TV-MA, TV-14).
- **Results**: The majority of content is rated TV-MA, indicating a focus on mature audiences. This may reflect a strategic alignment towards adult viewers who possibly have greater control over subscription decisions.

3. Movies & TV Shows Distribution

- **Analysis**: A pie chart compares the proportion of movies to TV shows.
- **Results**: Movies constitute 68.42% of the content, significantly outnumbering TV shows. This could indicate Netflix's strategy to prioritize films, which may attract more one-time viewership or cater to preferences for shorter viewing commitments.

4. Top 10 Genres

- **Analysis**: The top genres are analyzed using a horizontal bar chart to identify which are most prevalent.
- **Results**: Documentaries and stand-up comedies top the list, suggesting a niche but strong interest area among Netflix viewers, possibly due to the unique, informative, and entertaining content these genres offer.

5. Total Movies & TV Shows by Years

- **Analysis**: An area chart tracks the number of movies and TV shows added over the years.
- **Results**: There was a significant increase in content addition up to 2019, with a sharp rise in movies compared to TV shows. This trend may indicate a response to increasing competition in the streaming market, necessitating a robust content library.

Insights and Strategic Implications

- **Geographical Focus**: The heavy concentration of titles in specific markets like the US, India, and the UK suggests that Netflix could be optimizing its content strategy based on regional viewer preferences and market potential.
- **Content Rating Strategy**: The prevalence of TV-MA content aligns with an adult target audience, supporting content that caters to mature themes and complex narratives.
- **Content Type Strategy**: The dominance of movies over TV shows highlights Netflix's potential strategy to capture diverse viewer interests, offering films that cover various genres and story arcs in shorter viewing spans.
- **Genre Popularity**: The popularity of documentaries and comedies could guide future content development, suggesting a demand for high-quality, thought-provoking, and entertaining content.
- **Trend in Content Acquisition**: The increasing content acquisition up to 2019 reflects an aggressive growth strategy, possibly in response to emerging competitors and a growing global subscriber base seeking fresh and diverse content.

Future Scope

The analysis of Netflix's extensive content library using the current dataset has provided valuable insights, but there are several avenues for further exploration and enhancement. Future iterations of this analysis could focus on the following areas to deepen understanding and support more strategic decision-making for content management and viewer engagement:

1. Integration of Viewer Interaction Data

Future analyses could incorporate viewer ratings, watch times, and completion rates to gauge content popularity and viewer engagement more accurately. This data could help Netflix tailor its recommendations and explore the potential for renewing series or expanding movie franchises based on actual viewer behavior.

2. Predictive Analytics for Content Performance

Utilizing machine learning models to predict the success of various content types could significantly benefit Netflix’s strategic planning. Predictive analytics could help determine what genres or combinations of actor-director might perform well, allowing Netflix to proactively commission content that has a higher probability of success.

3. Sentiment Analysis of Viewer Reviews

Expanding the dataset to include viewer reviews and performing sentiment analysis could provide deeper insights into the subjective experiences of viewers. This approach would allow Netflix to capture viewer sentiment, which could be invaluable for content curation, promotional strategies, and improving customer satisfaction.

4. Demographic Analysis

Further analysis could segment viewer preferences by demographics such as age, location, and viewing device type. This segmented analysis would enable Netflix to craft more targeted marketing campaigns and content offerings, enhancing personalization and increasing viewer retention rates.

5. Cross-Platform Content Strategy Evaluation

Considering the competitive landscape, analyzing how Netflix’s content strategy stacks up against competitors could provide strategic insights. Comparative analyses involving similar datasets from other platforms like Amazon Prime, Hulu, or Disney+ could highlight strengths and weaknesses in Netflix’s current strategy and prompt more informed strategic decisions.

6. Economic and Regulatory Impact Analysis

Exploring the economic and regulatory impacts on content strategy, especially in diverse international markets, could further refine content localization strategies. This includes understanding how different content ratings affect viewership in various regions and adjusting content offerings to align with local regulations and cultural preferences.

7. Real-Time Data Utilization

Implementing a real-time data analysis framework could enable Netflix to make faster decisions. As the streaming wars intensify, the ability to quickly adapt to changing viewer preferences and market dynamics could become a critical competitive advantage.

8. Environmental and Social Governance (ESG) Factors

Finally, considering the growing importance of ESG factors, future analyses could incorporate assessments of how well Netflix's content and operations align with broader social, environmental, and corporate governance goals. This could include studies on the diversity of casts and crews, the environmental impact of productions, and the company's contributions to social issues.

References

1. **Kaggle**: The Netflix dataset used in this analysis was sourced from Kaggle, a platform known for its comprehensive datasets and community-driven data science resources.
2. **YouTube**: Tutorials and insights from YouTube were instrumental in understanding advanced Tableau features, including calculated fields, dynamic filters, and interactive visualizations.
3. **Tableau Documentation**: Official Tableau documentation provided guidance on using Tableau functions, mapping tools, and best practices for creating interactive dashboards.
4. **Data Science Blogs**: Various data science blogs and articles offered valuable perspectives on data cleaning, transformation processes, and visualization techniques, especially within the context of streaming data analysis.
5. **Research Papers on Streaming Analytics**: Academic research on streaming service analytics provided insights into industry trends, viewer engagement strategies, and the role of data in shaping content strategies for platforms like Netflix.

Bibliography

1. **Kaggle.** *Netflix Movies and TV Shows Dataset*. Retrieved from Kaggle.com, providing extensive data on Netflix content including titles, genres, ratings, release years, and more. This dataset served as the foundation for the dashboard's analysis.
2. **Tableau Software**. *Tableau Desktop Documentation*. Retrieved from Tableau.com. Official documentation on Tableau's functionalities, assisting in creating interactive visualizations and implementing calculated fields, filters, and mapping tools.
3. **YouTube**. *Tableau Tutorials and Dashboard Creation*. A collection of video tutorials that provided practical knowledge on using Tableau for data visualization, advanced filtering, and dashboard design principles.
4. **McKinsey & Company**. *The Age of Streaming: Trends and Insights in Media Consumption*. An industry report on streaming service trends, offering a broader understanding of viewer engagement and market demands, which influenced the strategic insights derived from the dashboard.
5. **Harvard Business Review**. *Data-Driven Decision Making in Media and Entertainment*. Discusses the impact of data analytics on content strategy in the entertainment industry, which provided context for analyzing Netflix's content curation through data.
6. **Data Science Central.** *Best Practices for Cleaning and Transforming Streaming Data*. Article on data transformation techniques and best practices in preparing data for visualization, guiding the ELT process for the Netflix dataset.
7. **Pew Research Center**. *Video Streaming Service Preferences by Demographics*. A research report providing insights into demographic-based content preferences, highlighting the importance of segmentation in streaming analytics.

Social Media Insights

Student Profile

https://www.linkedin.com/in/07sushant/

Social Media Insights

Post Insights

https://www.linkedin.com/in/07sushant/

Thank You