Dashboard for netflix on tableau

**Contents**

[**List of tables & Figures:** 3](#_Toc158330077)

[**Executive Summary** 4](#_Toc158330078)

[**1.** **Introduction** 5](#_Toc158330079)

[**1.1 Influence of Big Data Analytics on Netflix's Content Management** 5](#_Toc158330080)

[**1.2 Importance of Big Data in Netflix's Content Strategy** 6](#_Toc158330081)

[**1.3 Big Data Analytics Best Practices in the Entertainment Industry** 7](#_Toc158330082)

[**2.** **Data Brief** 9](#_Toc158330083)

[**2.1 Netflix Dataset Overview:** 9](#_Toc158330084)

[**3.** **Data Cleansing and Preparation** 10](#_Toc158330085)

[**4.** **Key Metrics and Relationships** 12](#_Toc158330086)

[**5.** **Strategic Insight** 19](#_Toc158330087)

[**6.** **Conclusion and Recommendations** 21](#_Toc158330088)

[**References:** 23](#_Toc158330089)

[**Appendices:** 24](#_Toc158330090)

**List of tables & Figures:**

[Figure I Big data & Netflix Source: (Dwivedi, 2022) 5](#_Toc158330091)

[Figure II Predictive analytics in Netflix 6](#_Toc158330092)

[Figure III Data Opened on Excel for Manipulation and cleaning source: excel worksheet data from: (Netflix Movies and TV Shows, 2021) 10](#_Toc158330093)

[Figure IV Metric 1 13](#_Toc158330094)

[Figure V Metric 2 14](#_Toc158330095)

[Figure VI Metric 3 15](#_Toc158330096)

[Figure VII Metric 4 16](#_Toc158330097)

[Figure VIII Metric 5 16](#_Toc158330098)

[Figure IX Metric 6 16](#_Toc158330099)

[Figure X Metric 7 17](#_Toc158330100)

[Figure XI Metric 8 17](#_Toc158330101)

[Figure XII Metric 9 18](#_Toc158330102)

[Figure XIII Metric 10 18](#_Toc158330103)

[Figure XIV Metric 11 19](#_Toc158330104)

[Figure XV Dashboard for Our Netflix Dataset on Tableau software 25](#_Toc158330105)

[Table I Big data importance 6](#_Toc157906461)

# **Executive Summary**

This report presents a comprehensive analysis of Netflix's content landscape, leveraging data insights derived from the provided dataset and visualized using Tableau software. The analysis focuses on key metrics such as total shows by year, shows by country, top genres, content distribution, ratings, and more, to uncover trends and inform strategic decisions.

**Key findings from the analysis include:**

* A substantial increase in both movies and TV shows in 2019, highlighting Netflix's robust content expansion strategy.
* The dominance of the USA and India as primary contributors to the content library, emphasizing the importance of global content localization.
* The popularity of documentaries and the need for genre diversity to cater to varied viewer interests.
* A balanced mix of movies and TV shows, with movies constituting the majority of the content library.
* The prevalence of TV-MA-rated content on the platform, signaling viewer preferences for mature audience-oriented content.

To capitalize on these insights and further enhance Netflix's content strategy, the following recommendations are proposed:

* Continue data-driven content acquisition and production to align with viewer preferences and industry trends.
* Invest in local content for key markets and foster collaborations with creators worldwide to diversify the content portfolio.
* Utilize interactive features and personalized recommendations to enhance user engagement and satisfaction.
* Implement real-time analytics to adapt quickly to changing viewer preferences and industry dynamics.

# **Introduction**

In today's ever-changing entertainment landscape, over-the-top (OTT) streaming services like Netflix are dominating the space thanks in large part to a series of strategic implementations of big data analytics. While the instant gratification that consumers feel when they stream a movie at the romantic evening's peak is immeasurable, we understand the impetus for big data analytics implementation on behalf of a streaming service like Netflix (Dwivedi, 2022). This section will explore big data analytics, specifically as it relates to content management, and how Netflix implements this tech to not only inform which titles it offers subscribers, but also to streamline various processes of the platform's operations (Gorgoglione et al., 2019).



Figure Big data & Netflix Source: (Dwivedi, 2022)

## **1.1 Influence of Big Data Analytics on Netflix's Content Management**

Netflix leverages big data analytics to anticipate the preferences of viewers and personalize its extensive collection to provide a tailored experience for each individual (Dwivedi, 2022). This data-centric approach enables Netflix to create highly popular shows and movies, ensuring that viewers have a delightful binge-watching experience and that Netflix maintains its leadership position in the streaming industry (Maddodi & K., 2019). By utilizing big data analytics, Netflix can anticipate viewers' interest’s hours before they become aware of them, featuring a diverse range of content including international selections, foreign language films, and cult classics (Dwivedi, 2022). This article delves into how big data analytics drives Netflix's content management strategy, positioning them ahead of their rivals (Maddodi & K., 2019).

The scope of this exploratory study is to ascertain how big data analytics serves at the core of Netflix content management that aids platform owners stay ahead in terms and trends, provide individual preference based considerations and continue enjoy competitive edge (Dwivedi, 2022). After that, subsequent parts will delve further into specific aspects of these influences but for now we should have this as a backdrop (Gorgoglione et al., 2019).

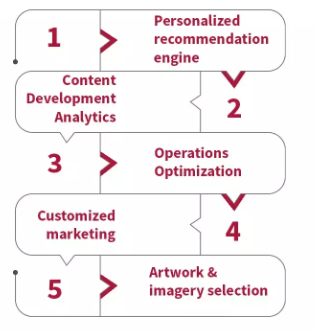


Figure II Predictive analytics in Netflix

Source: (Predictive Analytics: Grow Your Business like Netflix! | Engati, n.d.)

## **1.2 Importance of Big Data in Netflix's Content Strategy**

Table I Big data importance (Source: MS word)

|  |
| --- |
|  |

|  |  |
| --- | --- |
| **Aspect** | **Description** |
| **Content Selection** | Big data analytics plays a pivotal role in identifying content gaps, understanding genre popularity, and strategically investing in original productions (Mayorga Escalada, 2019). |
| **Recommendation Algorithms** | Advanced algorithms analyze user behavior to curate highly personalized recommendations based on viewing history, genre preferences, and other factors (Mayorga Escalada, 2019). |
| **User Experience Enhancement** | Insights into user interactions aid in optimizing the platform interface, improving streaming quality, and tailoring promotional efforts for an enhanced user experience (Mayorga Escalada, 2019). |

## **1.3 Big Data Analytics Best Practices in the Entertainment Industry**

According to (Mayorga Escalada, 2019), in today's rapidly-evolving and fiercely competitive entertainment landscape, it is imperative for companies to adeptly utilize big data analysis tools in their strategic decision-making process in order to maintain a competitive edge (Maddodi & K., 2019). Netflix, a leading streaming platform, has established industry-best practices by employing such data to inform its content strategies and enhance user experience (Maddodi & K., 2019).

**User Segmentation and Personalization:**

* + Segmenting users based on demographics, viewing habits, and preferences allows streaming platforms to deliver personalized content recommendations tailored to individual tastes (Gorgoglione et al., 2019).
  + By analyzing user data in real-time, platforms can adapt their content offerings and marketing strategies to target specific audience segments effectively (Gorgoglione et al., 2019).

**Predictive Analytics for Content Acquisition:**

* + Predictive analytics models analyze historical viewership data, market trends, and audience engagement metrics to forecast the potential success of new content acquisitions or original productions (Gorgoglione et al., 2019).

**Content Performance Measurement and Optimization:**

* + Utilizing metrics such as viewer engagement, retention rates, and content consumption patterns, streaming platforms assess the performance of individual titles and content categories.
  + Continuous monitoring and analysis of content performance enable platforms to optimize their content libraries, remove underperforming titles, and prioritize high-demand genres or formats.

**Dynamic Pricing Strategiess**

* + Big data analytics enables streaming platforms to implement dynamic pricing models based on user behavior, demand fluctuations, and market conditions (Gorgoglione et al., 2019).
  + By adjusting subscription fees, promotional offers, and pricing tiers in real-time, platforms can maximize revenue while maintaining user satisfaction and retention (Maddodi & K., 2019).

**Content Discovery and Recommendation Engines:**

* + Advanced recommendation algorithms leverage machine learning and collaborative filtering techniques to deliver personalized content recommendations to users (Maddodi & K., 2019).
  + By analyzing user interactions, viewing histories, and feedback, recommendation engines continuously refine their algorithms to enhance content discovery and user engagement (Gorgoglione et al., 2019)..

**Data Privacy and Security Measures:**

* + Streaming platforms prioritize data privacy and implement robust security measures to safeguard user information and comply with regulatory requirements (Maddodi & K., 2019).
  + Transparent data policies, encryption protocols, and regular security audits ensure the integrity and confidentiality of user data, fostering trust and loyalty among subscribers.

# **Data Brief**

**Source of data: kaggle**

****

**Source: (*Kaggle SVG Logo*, n.d.)**

## **2.1 Netflix Dataset Overview:**

The Netflix dataset contains information about the various shows it offers viewers, both in the form of movies and TV series. The dataset brings you valuable information about the content library including show ID, type (movie or tv shows), title, director, cast, country where it was produced, date added, release year, rating and duration (*Netflix Movies and TV Shows*, 2021). The record for every field gives a special take on its particular show so we can really look at Netflix as a whole by looking at all these pieces information together (*Netflix Movies and TV Shows*, 2021).

*Note: our data is from period 2008-2020. (Kaggle, 2023) and contains 8808 rows of data.*

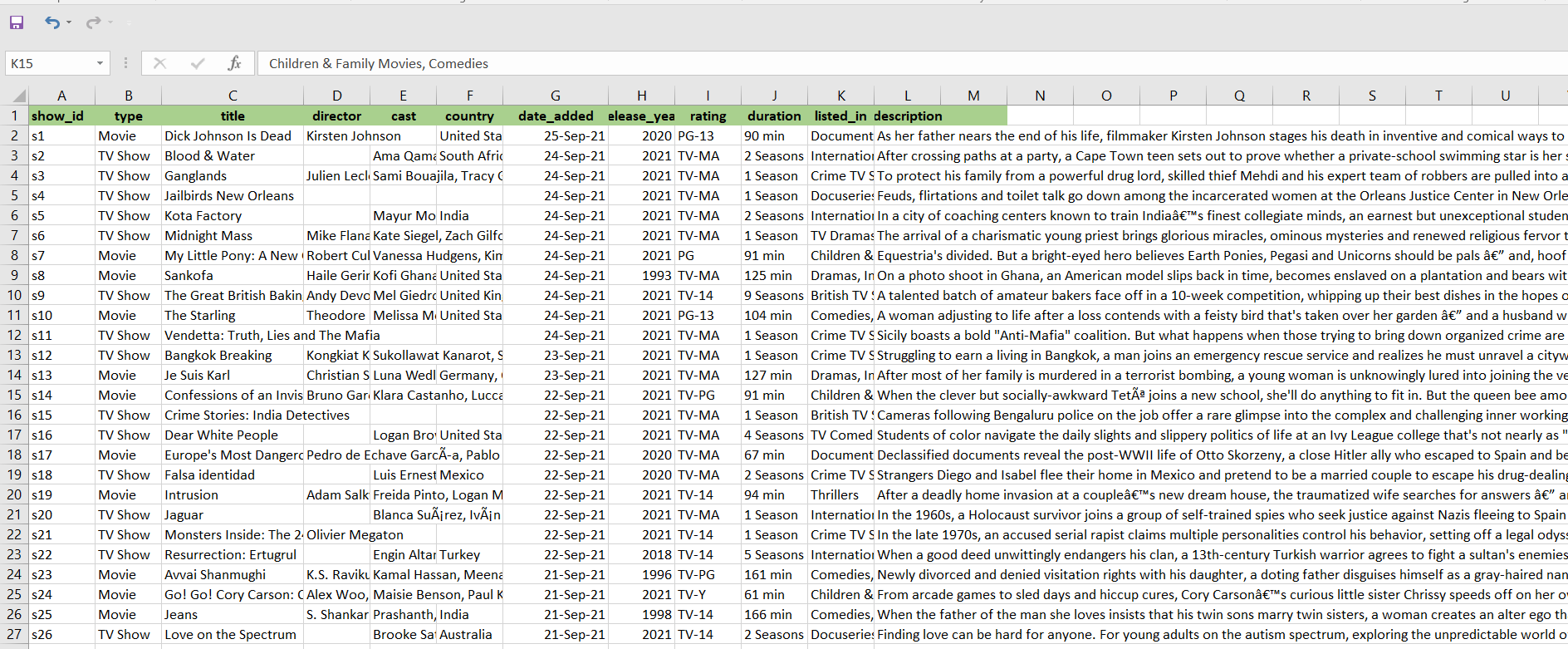
**

Figure Data Opened on Excel for Manipulation and cleaning source: excel worksheet data from: (Netflix Movies and TV Shows, 2021)

**Key Fields in the Netflix Dataset:**

1. **show\_id:** Unique identifier for each show in the dataset.
2. **Type:** Indicates whether the entry is a movie or a TV show.
3. **Title:** The title of the show.
4. **Director:** The director(s) of the show.
5. **Cast:** The cast members starring in the show.
6. **Country:** The country or countries associated with the show.
7. **Date added:** The date when the show was added to the Netflix platform.
8. **Release year:** The year when the show was released.
9. **Rating:** The content rating assigned to the show (e.g., PG, TV-MA).
10. **Duration:** The duration of the show (e.g., "1 Season", "90 min").
11. **Listed in:** The category or categories the show is listed under (e.g., "Comedies", "Documentaries").
12. **Description:** A brief description of the show's plot or premise.

This dataset serves as a valuable resource for exploring the content landscape of Netflix, enabling analyses related to content distribution, trends over time, genre preferences, and more.

# **Data Cleansing and Preparation**

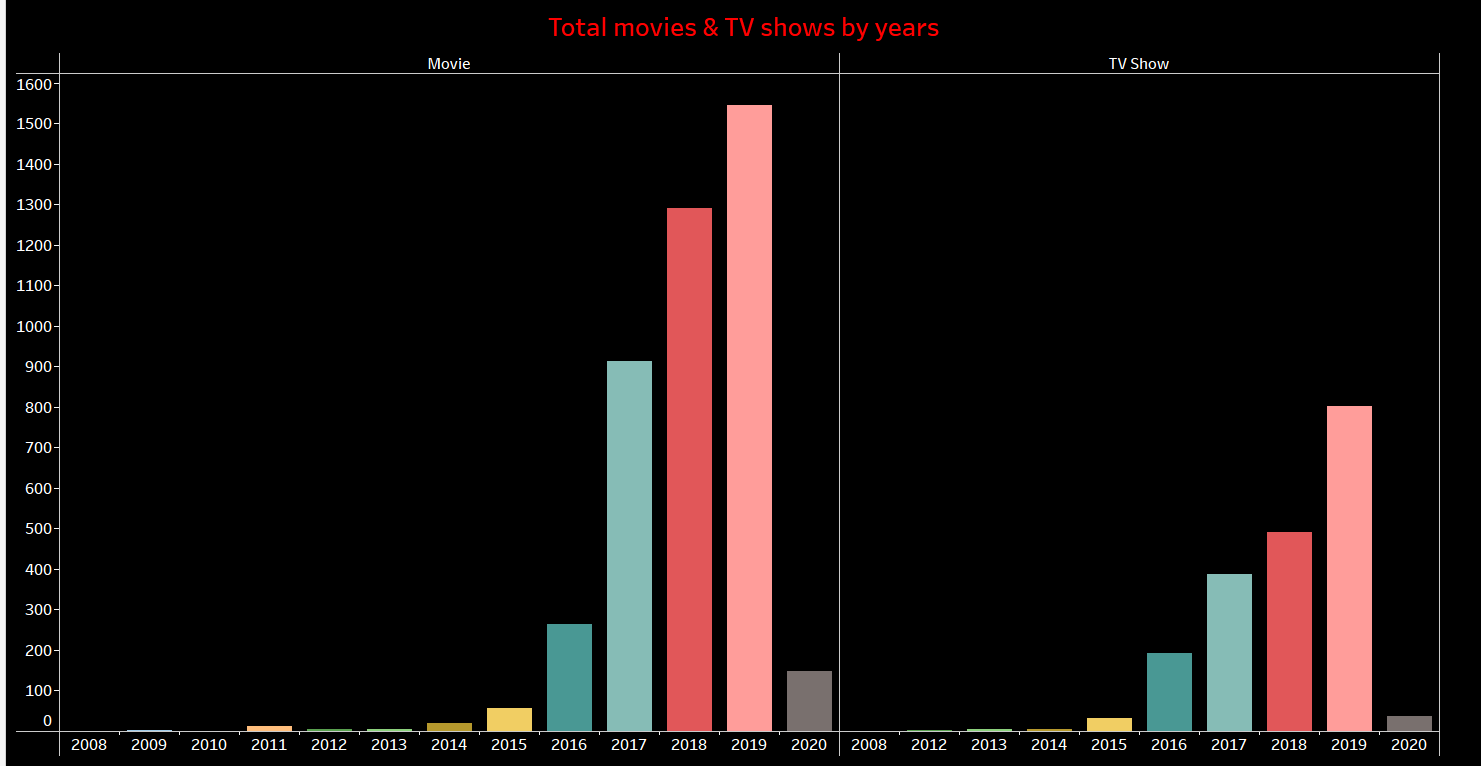
Cleaning and transforming to the Netflix dataset involved several steps, primarily executed in Microsoft Excel before importing the refined dataset into Tableau. This process aimed to ensure data accuracy and consistency, laying the groundwork for reliable insights within the Tableau environment for ease.

1. **Handling Missing Values:**
   * In Excel, missing values in crucial fields like 'director' and 'cast' were identified and addressed. Some entries were manually filled based on available information, ensuring a more comprehensive dataset. The manual values which were filled were confirmed from Google’s reputable sites.
2. **Standardizing Data Formats:**
   * Date formats in the 'date\_added' field were standardized in Excel to a consistent format (e.g., DD/MM/YYYY).
3. **Addressing Duplicates:**
   * Duplicate entries were identified and removed using Excel's data deduplication command/feature.
4. **Dealing with Inconsistencies:**
   * Categorical variables such as 'type' and 'rating' were standardized for consistency i.e PG-13,TV-MA etc. Excel's find and replace functionalities were employed to streamline naming conventions.
5. **Parsing and Extracting Information:**
   * In Excel, relevant information was parsed and extracted from fields like 'cast' and 'listed\_in' to create structured categorical variables, improving categorization and analytical capabilities in Tableau.
6. **Converting Data Types:**
   * Data types were verified and converted where necessary. For instance, the 'release\_year' field was ensured to be in a numeric format to facilitate numerical operations in Tableau.
7. **Handling Outliers:**
   * Outliers in numerical fields like 'duration' field one were identified and reviewed in Excel. While some extreme values were adjusted based on domain knowledge, the process was documented for transparency.
8. **Text Cleaning in Descriptions:**
   * Excel was used to clean and preprocess text data in the 'description' field, removing special characters and irrelevant symbols to enhance the accuracy of it.
9. **Creating Derived Features:**
   * Additional features were created in Excel, such as extracting the month from the 'date\_added' field. This facilitates monthly trend analysis within Tableau.
10. **Encoding Categorical Variables:**
    * Excel was used to encode categorical variables like 'type' and 'rating' into numerical representations, employing techniques such as label encoding for compatibility with Tableau.
11. **Ensuring Data Quality:**
    * Rigorous checks were implemented in Excel to ensure data quality at each stage. The refined dataset was then imported into Tableau, providing a reliable foundation for subsequent analyses and visualizations.

# **Key Metrics and Relationships**

In analyzing the Netflix dataset within Tableau, several key metrics and relationships were explored to gain insights into the content landscape and user preferences. The chosen metrics and relationships are essential for understanding trends, patterns, and correlations within the dataset. Below is a discussion of these metrics and relationships as reflected in the Tableau work-sheets and dashboard (Gorgoglione et al., 2019).

1. **Total Movies & TV Shows by Years:**
   * This metric provides insights into the growth and evolution of Netflix's content library over time (*Netflix Movies and TV Shows*, 2021).
   * By visualizing the total number of movies and TV shows added each year, trends in content production and acquisition can be identified (*Netflix Movies and TV Shows*, 2021).



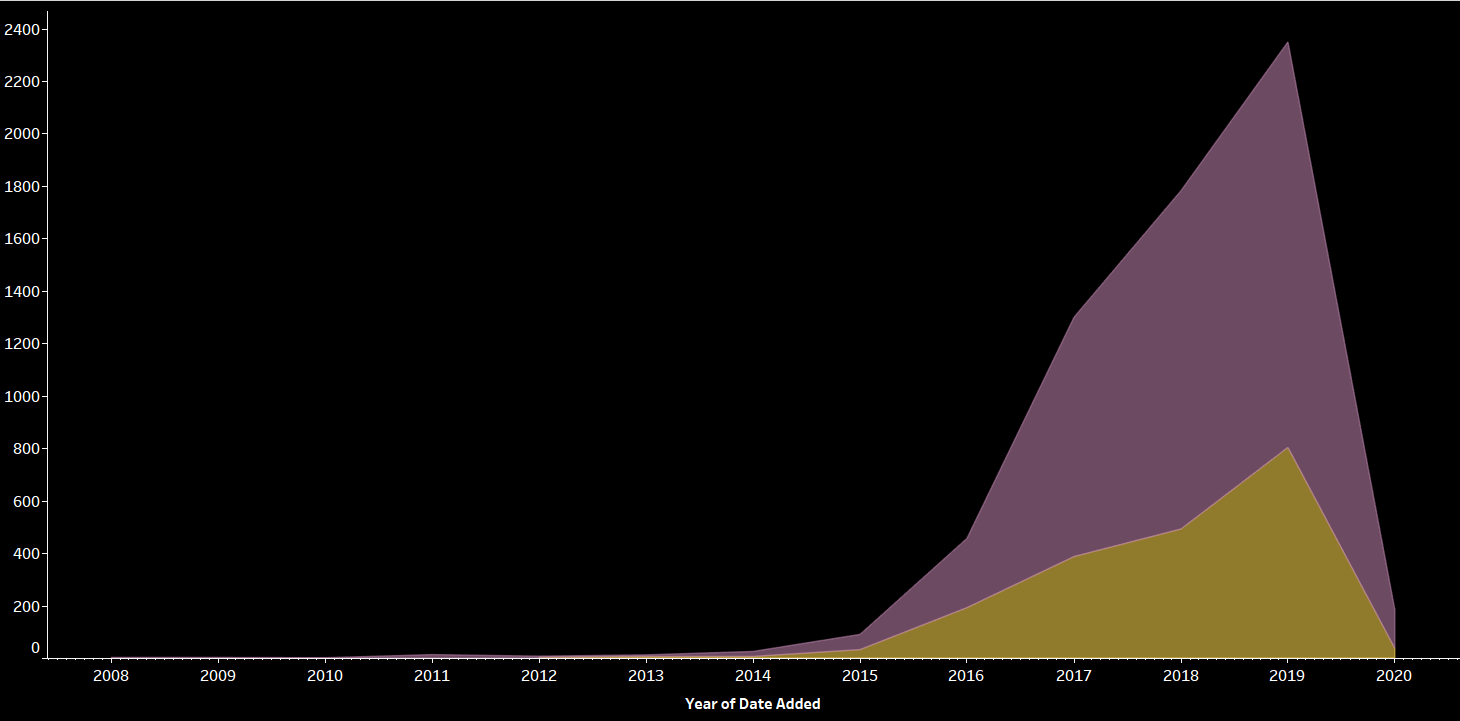
.

Figure Metric 1

1. **Total Movies & TV Shows by Country:**
   * Examining the distribution of shows by country highlights Netflix's global reach and localization efforts (*Netflix Movies and TV Shows*, 2021).
   * This metric reveals which countries contribute the most content to the platform and may indicate regional content preferences (*Netflix Movies and TV Shows*, 2021).

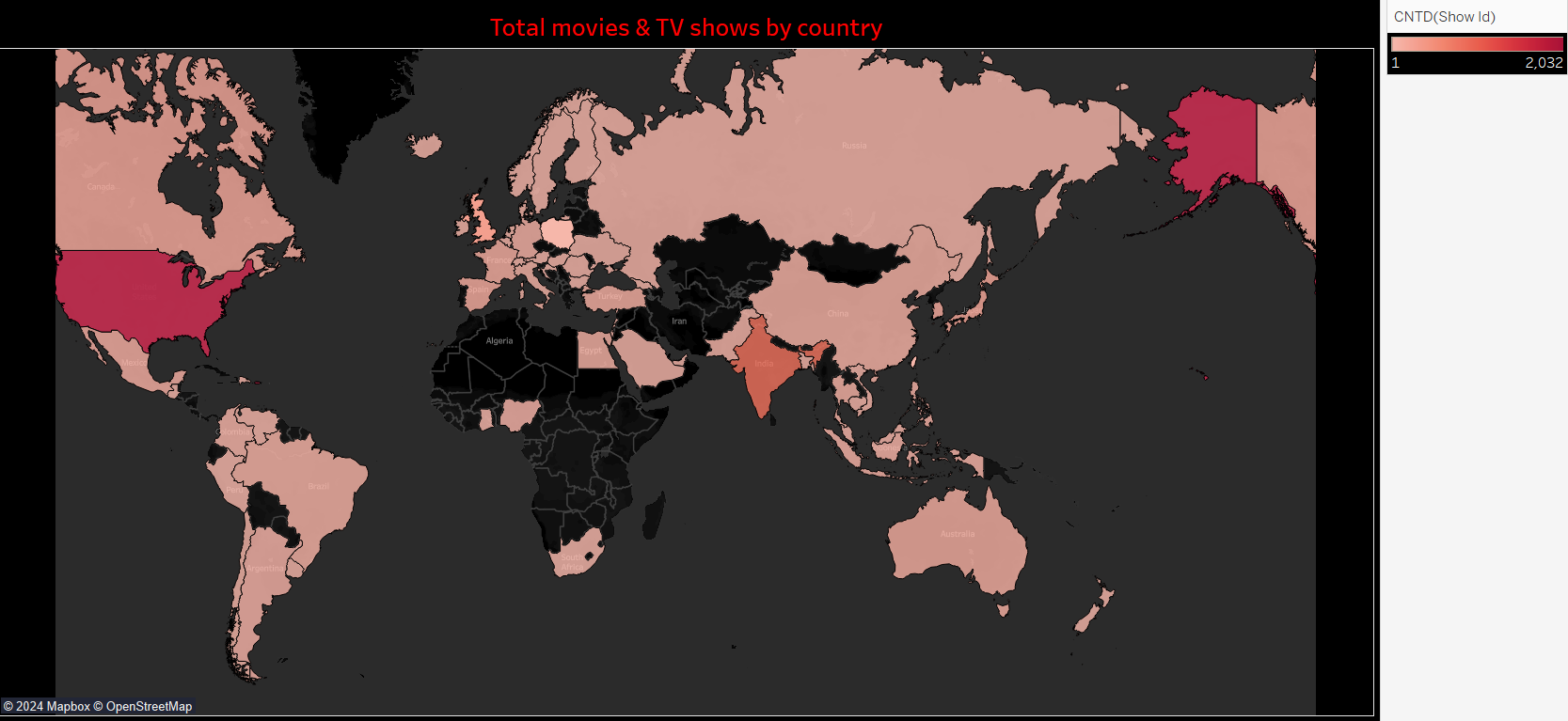


Figure Metric 2

1. **Top 10 Genre:**
   * Identifying the top genres among Netflix's content library offers valuable insights into audience preferences and content trends (*Netflix Movies and TV Shows*, 2021).
   * This metric helps prioritize content acquisition and production efforts in popular genres (*Netflix Movies and TV Shows*, 2021).

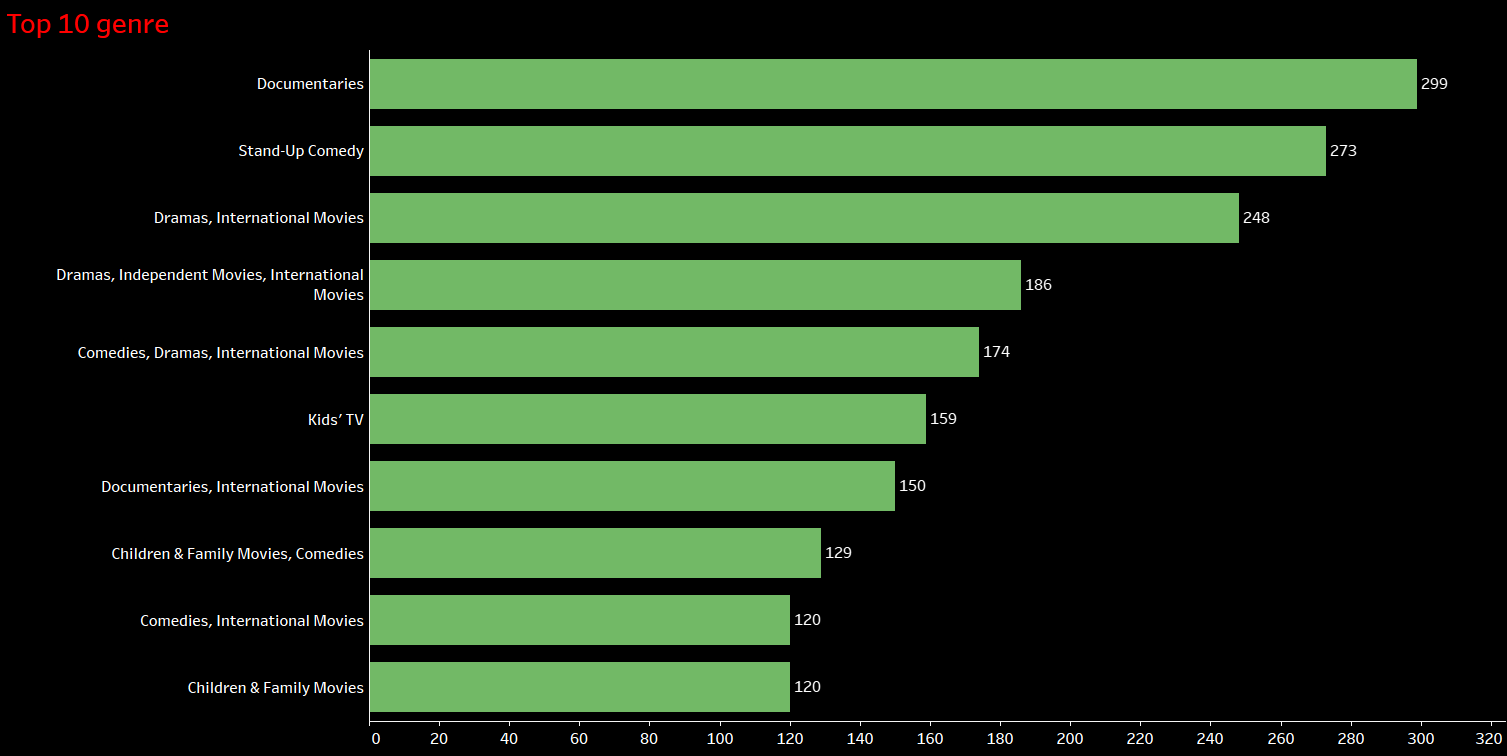


Figure Metric 3

1. **Movies and TV Shows Distribution:**
   * Analyzing the distribution of movies and TV shows within the dataset allows for a comparison of content types and their respective proportions. Understanding this distribution informs content strategy and resource allocation (*Netflix Movies and TV Shows*, 2021).

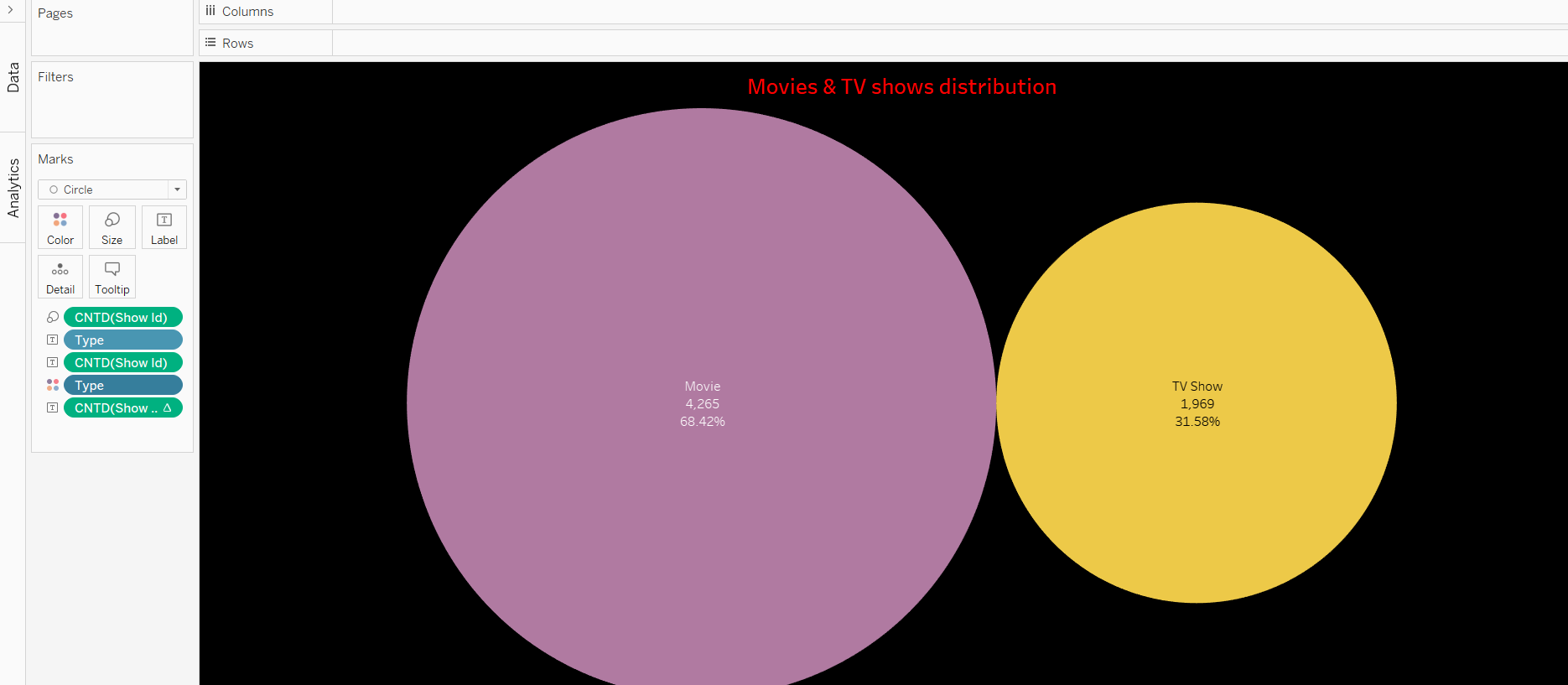


Figure Metric 4

1. **Ratings:**
   * Ratings provide an indication of audience reception and satisfaction with Netflix's content offerings (*Netflix Movies and TV Shows*, 2021). Analyzing the distribution of ratings across movies and TV shows helps identify viewer preferences and content quality trends (*Netflix Movies and TV Shows*, 2021).

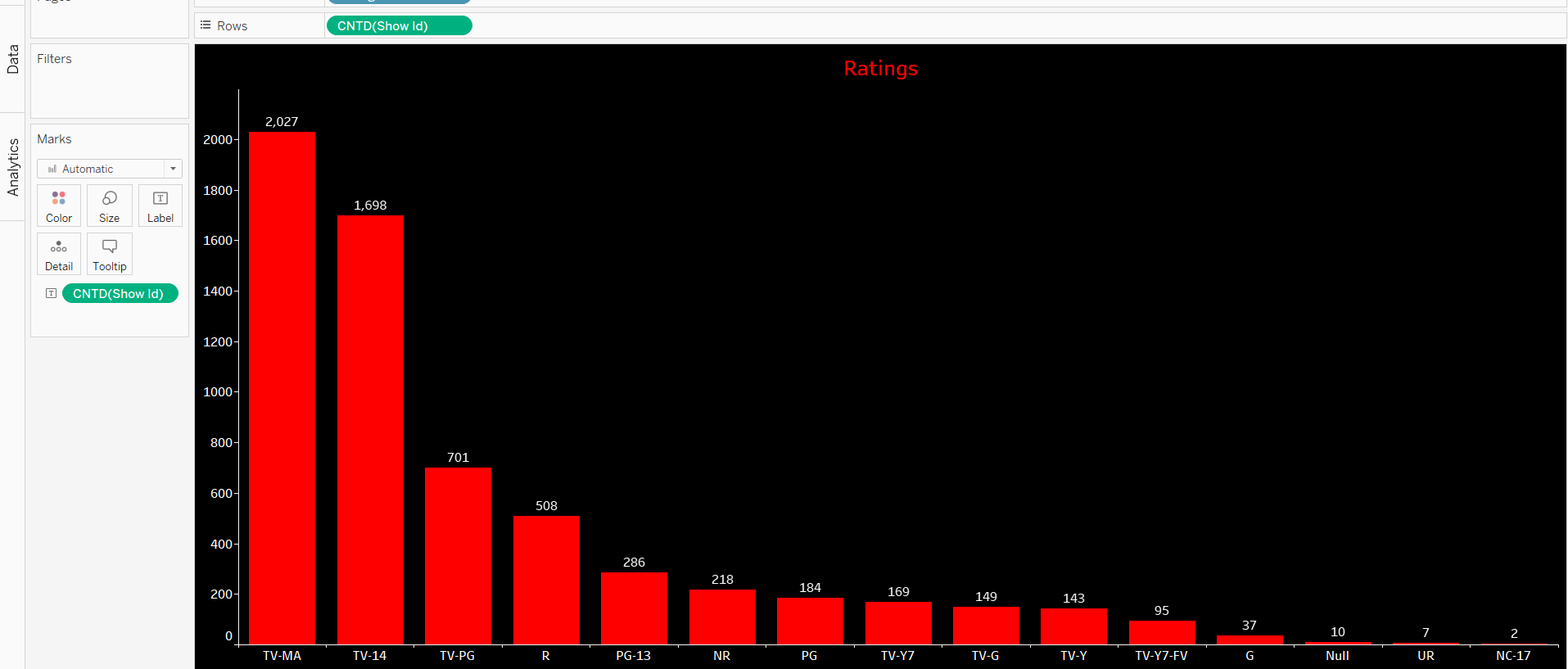


Figure Metric 5

1. **Description According to Type and Title Dropdown:**
   * This metric enables users to explore show descriptions based on type (movie or TV show) and title. It enhances user engagement by providing detailed information about specific titles within the dataset (*Netflix Movies and TV Shows*, 2021).



Figure Metric 6

1. **Date Added According to Type and Title Dropdown:**
   * By examining the dates when shows were added to Netflix, trends in content acquisition and release patterns can be observed (*Netflix Movies and TV Shows*, 2021). This metric helps identify peak periods of content addition and may inform content scheduling strategies (*Netflix Movies and TV Shows*, 2021).



Figure Metric 7

1. **Duration According to Type and Title Dropdown:**
   * Understanding the duration of movies and TV shows provides insights into content format preferences and viewer engagement (*Netflix Movies and TV Shows*, 2021). This metric aids in identifying trends in content length and consumption patterns (*Netflix Movies and TV Shows*, 2021).

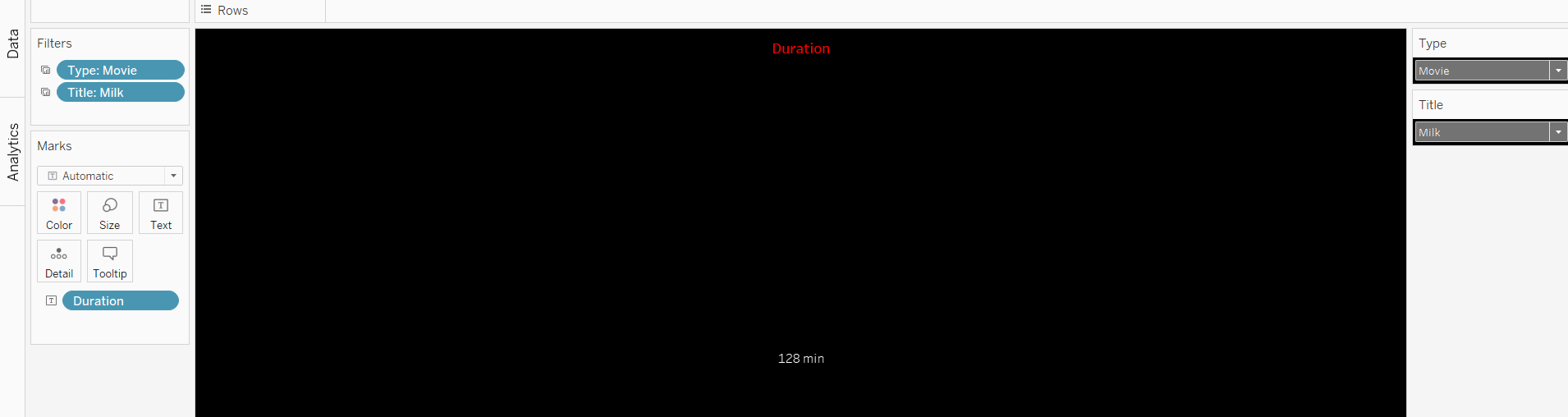


Figure Metric 8

1. **Rating According to Type and Title Dropdown:**
   * Exploring ratings based on type and title allows for a detailed assessment of audience reception and content quality (*Netflix Movies and TV Shows*, 2021). This metric facilitates comparisons between individual titles and informs content recommendation algorithms (*Netflix Movies and TV Shows*, 2021).



Figure Metric 9

1. **Release Year According to Type and Title Dropdown:**
   * Analyzing the release years of movies and TV shows helps identify content trends and historical patterns within the dataset (*Netflix Movies and TV Shows*, 2021). This metric aids in understanding the evolution of content preferences over time (*Netflix Movies and TV Shows*, 2021).

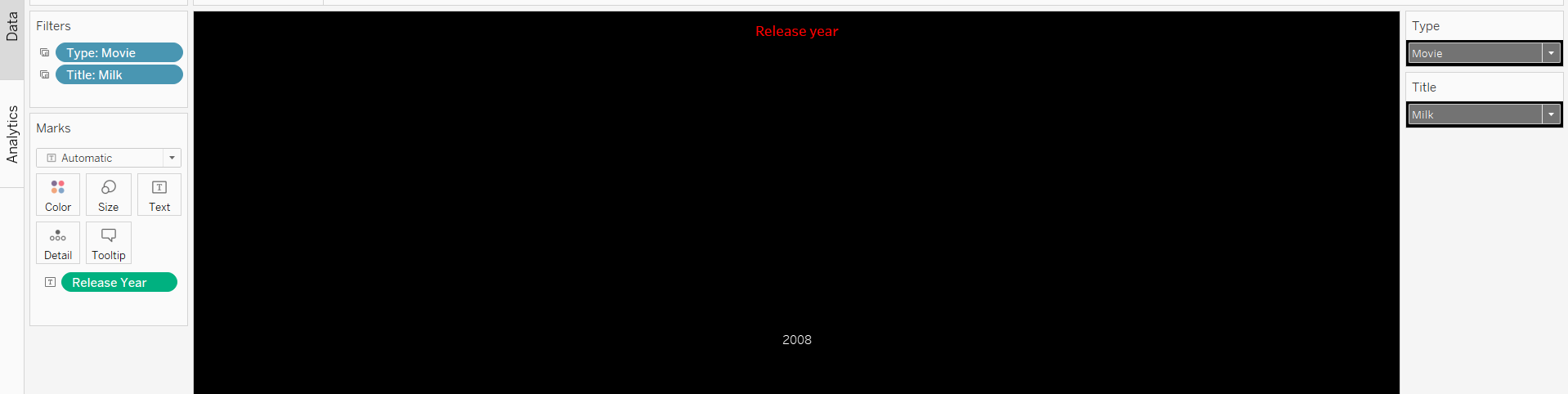


Figure Metric 10

1. **Genre According to Type and Title Dropdown:**
   * Examining genres based on type and title enables users to explore content categorization and genre preferences (*Netflix Movies and TV Shows*, 2021). This metric enhances content discovery and recommendation capabilities within the platform (*Netflix Movies and TV Shows*, 2021).



Figure Metric 11

# **Strategic Insight**

**Data Insights/Strategic Insights**

After a comprehensive analysis of the Netflix dataset using various Tableau worksheets, key insights have been derived, shedding light on content distribution, audience preferences, and strategic considerations.

1. **Total Movies & TV Shows by Years:**
   * The analysis of movies and TV shows by year, represented in side-by-side bars and line graphs, revealed a significant trend in 2019. In that year, Netflix added 1546 movies and 803 TV shows to its library, showcasing a robust content expansion strategy (*Netflix Movies and TV Shows*, 2021).
2. **Total Movies & TV Shows by Country:**
   * Symbol maps showcased the geographical distribution of shows, and India and the United States emerged as frontrunners. The USA led with 2,032 show IDs, closely followed by India with 777. This underscores the global diversity of Netflix's content, with a strong presence in these key markets (*Netflix Movies and TV Shows*, 2021).
3. **Top 10 Genre:**
   * Horizontal bars highlighted the top genres, with documentaries claiming the lead with 299 entries. This insight emphasizes the popularity and demand for factual and informative content among Netflix viewers.
4. **Movies and TV Shows Distribution:**
   * The scatter plot revealed that movies constitute a significant portion of the content library, with 4,265 entries, accounting for 68.42%. TV shows, while fewer in number (1,969), still contribute substantially at 31.58%. This insight informs content strategy, indicating a preference for a diverse mix of both movie and TV show offerings (*Netflix Movies and TV Shows*, 2021).
5. **Ratings:**
   * Horizontal bars showcased that content with a 'TV-MA' rating dominates the platform, with 2,027 entries. This insight indicates a significant portion of mature audience-oriented content in Netflix's library.
6. **Description Worksheet:**
   * The drop-down menu for descriptions, according to type and title, provides users with detailed information about specific shows. This enhances user engagement by offering comprehensive insights into the plot or premise of each title.
7. **Date Worksheet:**
   * The drop-down menu for dates added, according to type and title, enables users to explore the timeline of content additions. This insight aids in understanding the periodicity of content acquisition and release patterns.
8. **Duration Worksheet:**
   * The drop-down menu for duration, according to type and title, allows users to delve into the length of individual movies and TV shows. This information is valuable for understanding viewer preferences for content of varying lengths.
9. **Rating Worksheet:**
   * The drop-down menu for ratings, according to type and title, facilitates exploration of content based on assigned ratings. This insight is crucial for viewers seeking content aligned with specific maturity or content rating preferences.
10. **Release Year Worksheet:**
    * The drop-down menu for release years, according to type and title, enables users to explore content trends over time. This insight aids in understanding the evolution of content preferences and industry dynamics.
11. **Genre Worksheet:**
    * The drop-down menu for genres, according to type and title, allows users to explore content based on specific genre preferences. This insight enhances content discovery and recommendation capabilities, catering to diverse viewer tastes.

# **Conclusion and Recommendations**

Through the analysis of the Netflix dataset, valuable insights have been garnered to develop a strategic plan for content. Based on the findings, the following recommendations are suggested:

1. **Content Expansion**: Utilize data-driven insights to better understand viewers and their preferences.
2. **Global Localization:** Invest in producing more local content that is specifically tailored to diverse audiences in countries such as America and India.
3. **Genre Expansion:** Consider investing in a wider range of genres to cater to varying viewer tastes, including the popularity of nonfiction films.
4. **Balanced Programming Mix:** Maintain a balanced mix of movies and television shows to satisfy different viewer preferences.
5. **Increase Audience Participation**: Engage viewers and enhance their experience through interactive features and personalized recommendations.
6. **Continuous Observation:** Stay up-to-date with industry trends and remain ahead of the competition by utilizing real-time analytics.
7. **Co-Creation:** Collaborate with diverse program designers to increase diversity and further solidify Netflix's position as a leading global entertainment company.

**References:**

1. Maddodi, S., & K, K. P. (2019, January 1). *Netflix Bigdata Analytics- The Emergence of Data Driven Recommendation*. Social Science Research Network. <https://doi.org/10.2139/ssrn.3473148>
2. *Netflix Movies and TV Shows*. (2021, September 27). Kaggle. https://www.kaggle.com/datasets/shivamb/netflix-shows?resource=download
3. Maddodi, S., & K., K. P. (2019, October 20). Netflix Bigdata Analytics- The Emergence of Data Driven Recommendation. *International Journal of Case Studies in Business, IT, and Education*, 41–51. <https://doi.org/10.47992/ijcsbe.2581.6942.0050>
4. Dwivedi, R. (2022, January 21). *Deciphering the unstoppable Netflix and the role of Big Data*. Muvi One. <https://www.muvi.com/blogs/deciphering-the-unstoppable-netflix-and-the-role-of-big-data/>
5. Mayorga Escalada, S. (2019, July 31). Netflix, strategy and brand management around the relevance of the content. *AdComunica*, 219–244. <https://doi.org/10.6035/2174-0992.2019.18.11>
6. Gorgoglione, M., Panniello, U., & Tuzhilin, A. (2019, September 1). *Recommendation strategies in personalization applications*. Information & Management. <https://doi.org/10.1016/j.im.2019.01.005>
7. *Predictive Analytics: Grow your business like Netflix! | Engati*. (n.d.). Engati. <https://www.engati.com/blog/predictive-analytics>
8. Kaggle SVG Logo. (n.d.). VectorLogoZone. <https://www.vectorlogo.zone/logos/kaggle/index.html>

**Appendices:**

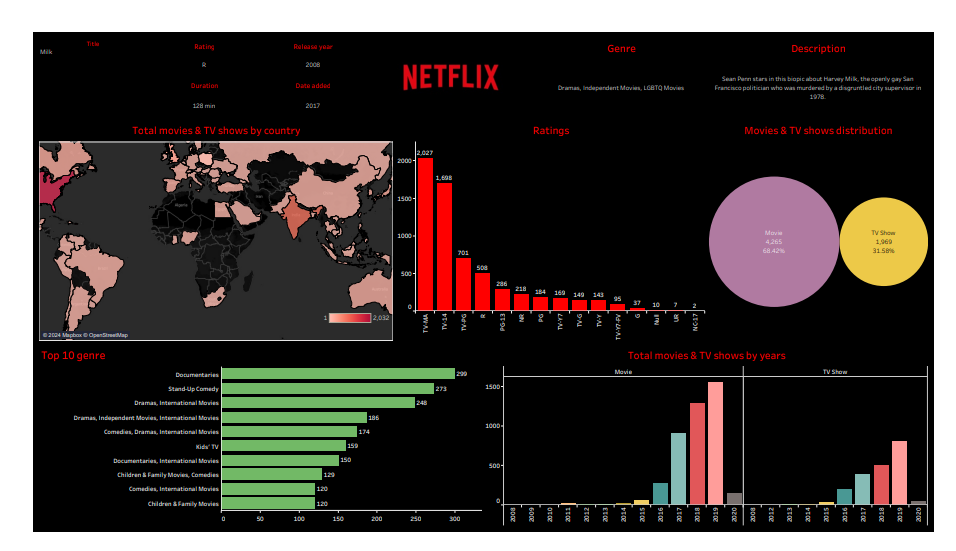


Figure Dashboard for Our Netflix Dataset on Tableau software