THE BEST STREAMING PLATFORM

NETFLIX: CONTENT ANALYISIS 2008-2021

1.- **INDEX:**

1. Description of the Data Theme:

The following is an analysis of all the streams of various movies and series on Netflix to gain a general idea of what is most liked and most viewed by the audience. The data source was downloaded from Kaggle, and adjustments were made, such as the number of streams, to manage its information effectively.

2. Description of the Hypothesis:

The goal is to demonstrate and understand how Netflix is one of the largest streaming platforms and one of the leading providers of entertainment worldwide. Some questions to analyze include the count of streams, ratings, and revenue of each movie added to the platform, which countries consume this type of content the most, which year was the most active, among many others.

3. Project Objective and Scope:

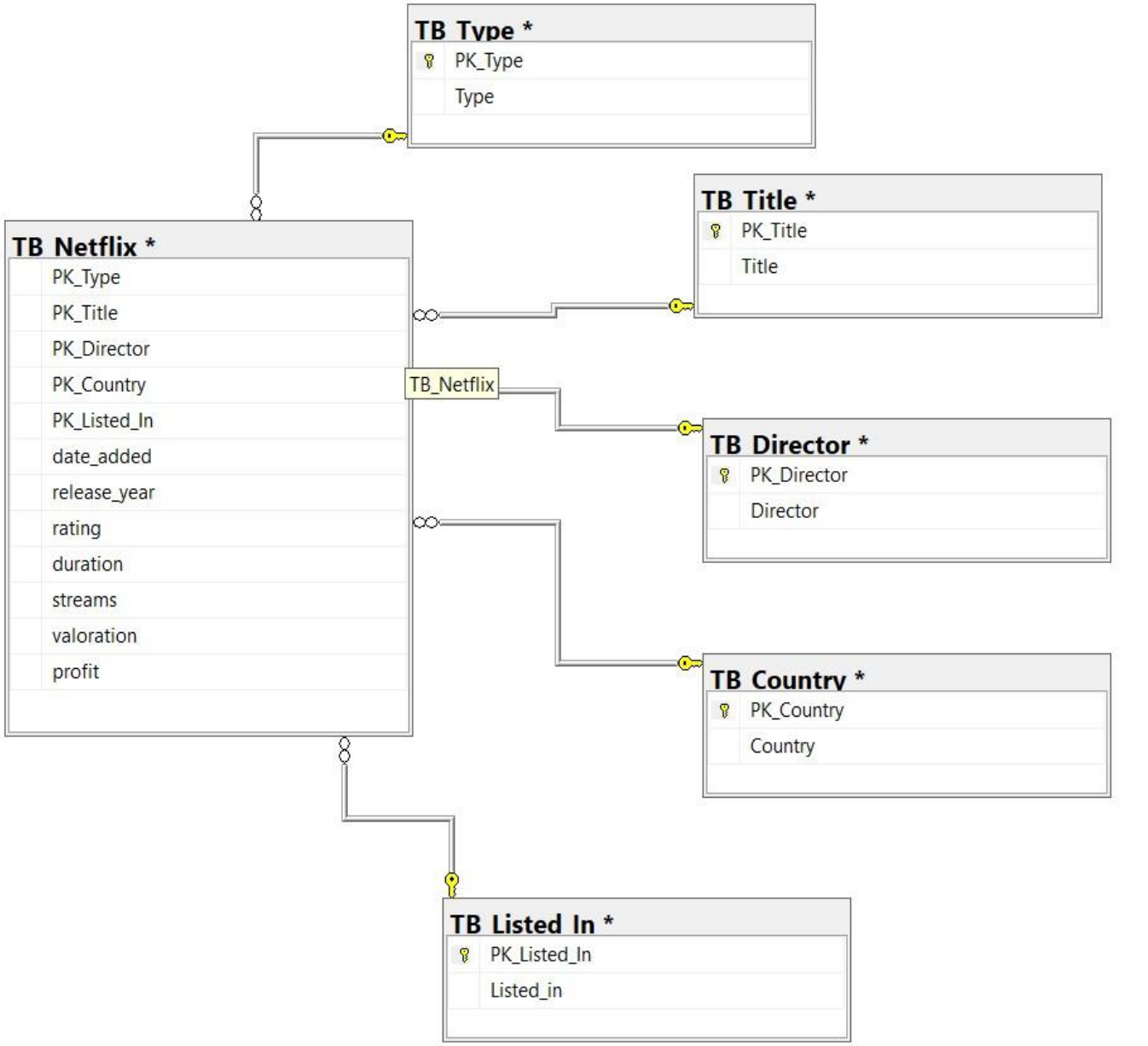
This analysis aims to understand the growth and success of the platform from past times to a significant future projection through descriptive and predictive analysis. The intended level of scope is to pave the way for investments and global marketing strategies, fostering the growth and positioning of the platform in the entertainment market.

4. End User and Level of Analysis Application:

This analysis is primarily aimed at individuals within the strategic level of the company so that, through the proper use of data, creative and administrative decisions can be made within the entertainment platform business. Additionally, it could also be used by users at the tactical and operational levels who want to see with greater granularity how the platform's data behaves.

5. Entity–Relationship Diagram:

The following was created using SQL.



6. List of Tables:

Below is a brief description of the tables available within the dataset, along with their primary and foreign keys and the tables they reference.

TB_Netflix:

This is the main table in the relationship model. All primary keys (PK) and foreign keys (FK) originate from here, as well as columns such as:

-Date_added

-Release year

-Rating

-Duration

-Streams -

Valoration -Profit.

TB_Type:
This table contains the primary keys (PK) for the types of movies/series, segmenting the data into "Movie" or "TV Show." It is related to **TB_Netflix**.

TB_Title:
This table contains the primary keys (PK) for the names of movies/series. It includes all the movies/series registered during the specified time period. It is related to **TB_Netflix**.

TB_Director:
This table contains the primary keys (PK) for the names of directors. It includes the first and last names of those responsible for directing the work. It is related to **TB_Netflix**.

TB_Country:

This table contains the primary keys (PK) for the names of countries. It includes the names of the countries where the work was produced. It is related to **TB_Netflix**.

TB_Listed_in:

This table contains the primary keys (PK) for the categories of various movies/series. It includes segmented data for the different categories. It is related to **TB_Netflix**.

7. Technological Tools Used and Transformations Performed:

Kaggle: The CSV file for this project was downloaded from Kaggle.

Excel: This was used to read and clean the dataset (removing duplicates, establishing table relationships, etc.).

SQL Server Management Studio: This was used to create the entity-relationship diagram.

Power BI Desktop/Power Query: This was used to create the dashboard, as well as perform other transformations (removing blanks, creating tables, etc.).

8.- Created Tables - Created Calculated Measures and Their Formulas

- **India Total Streams:** Calculates the total streams in India.
  India Total streams = CALCULATE(SUM(TB_Netflix[streams]), TB_Country[Country] = "India")

- **Average Streams:** Returns the overall average of total streams.
  Average streams = AVERAGE(TB_Netflix[streams])

- **Streams 2019:** Calculates the total streams for the year 2019.
  Streams 2019 = CALCULATE(SUM(TB_Netflix[streams]), TB_Calendario[Year] = 2019)

- **Total Movies/TV Shows:** Returns the total count of movies/TV shows.
  Total movies/tv shows = COUNT(TB_Netflix[PK_Title])

- **Total Profit:** Returns the total sum of profit.
  Total profit = SUM(TB_Netflix[profit])

- **Total Streams:** Returns the total sum of streams.
  Total streams = SUM(TB_Netflix[streams])

- **USA Total Streams:** Calculates the total streams in the United States.
  USA Total streams = CALCULATE(SUM(TB_Netflix[streams]), TB_Country[Country] = "United States")

- **Perfect Rating:** Counts the number of movies or series with a rating of 5.
  Rating 5 =
  CALCULATE(COUNT(TB_Netflix[valoration]),
  TB_Netflix[valoration] = 5)

- **Movie Profit:** Calculates the total profit of movies on Netflix.
  Movie profit =
  CALCULATE(SUM(TB_Netflix[profit]),
  TB_Type[Type] = "Movie")

- **TV Show Profit:** Calculates the total profit of TV shows on Netflix.
  TV show profit =
  CALCULATE(SUM(TB_Netflix[profit]),
  TB_Type[Type] = "TV SHOW")

Calendar Table

- **Year:** *Year = YEAR(TB_Calendario[date_added])*

- **Day of the Week (Number):** *Weekday Number = WEEKDAY(TB_Calendario[date_added], 2)*

- **Date ID:** *Date ID = FORMAT(TB_Calendario[date_added], "DD/MM/YYYY")*

- **Day of the Week (Full Name):** *Day of the Week = FORMAT(TB_Calendario[date_added], "DDDD")*

- **Day of the Week (Short Name):** *Short Day Name = FORMAT(TB_Calendario[date_added], "DDD")*

- **Month Name:** *Month Name = TB_Calendario[date_added].[Month]*

- **Month Name (Short):** *Short Month Name = FORMAT(TB_Calendario[date_added], "MMM")*

- **Month Number:** *Month Number = MONTH(TB_Calendario[date_added])*

- **Period (Year-Month):** *Period = FORMAT(TB_Calendario[date_added], "YYYY-MM")*

- **Quarter:** *Quarter = "Q" & FORMAT(TB_Calendario[date_added], "Q")*

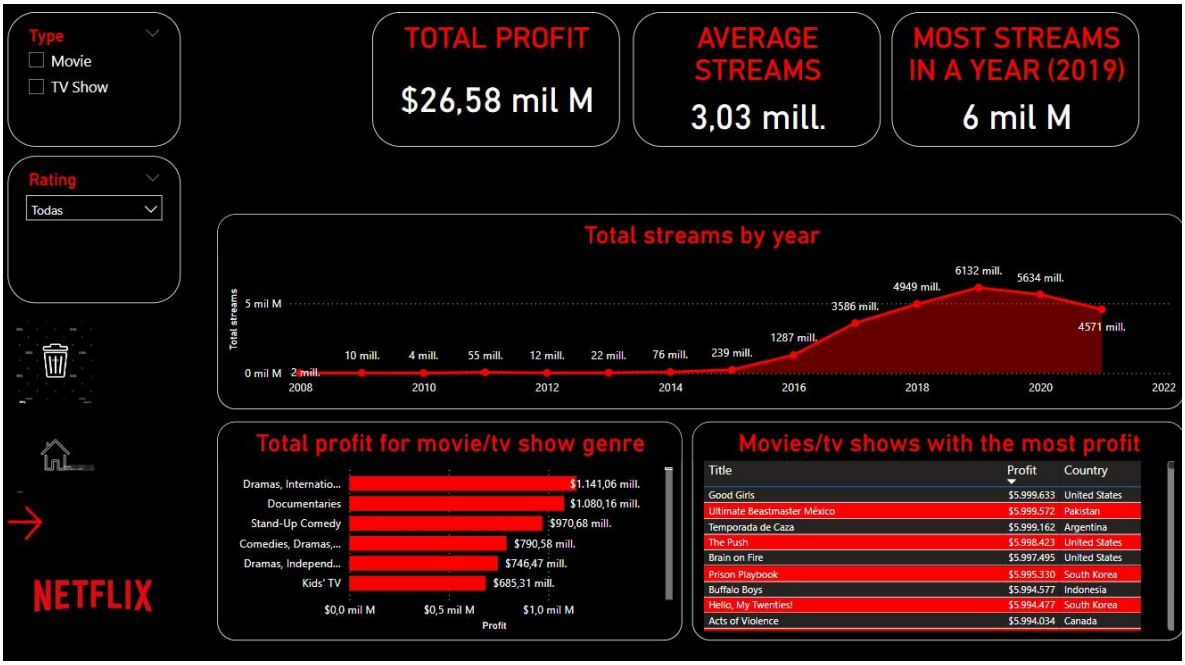
9.- Dashboard:

Cover Page:

This cover page aims to introduce the project and its vision through storytelling.

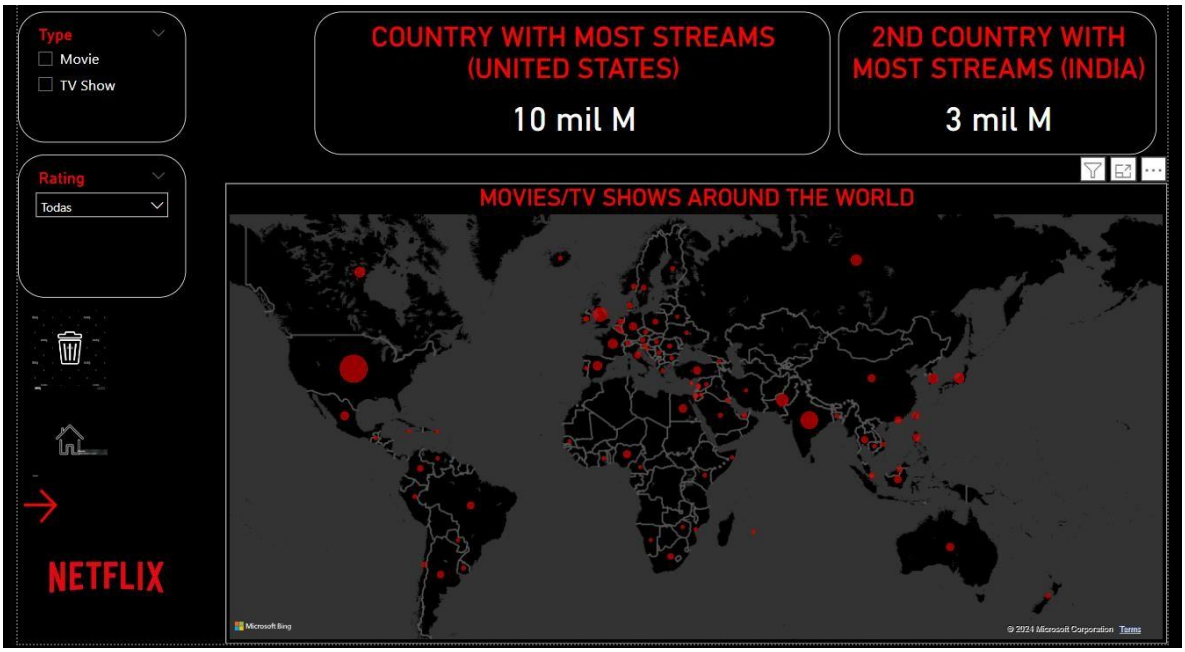
Streams by year:

This section primarily displays the number of streams per year using a line chart. It also includes the profits of various movies/series by genre, highlighting the most profitable ones. Additionally, it provides key insights such as Netflix's total profit, the overall average number of streams, and the year with the highest streams.



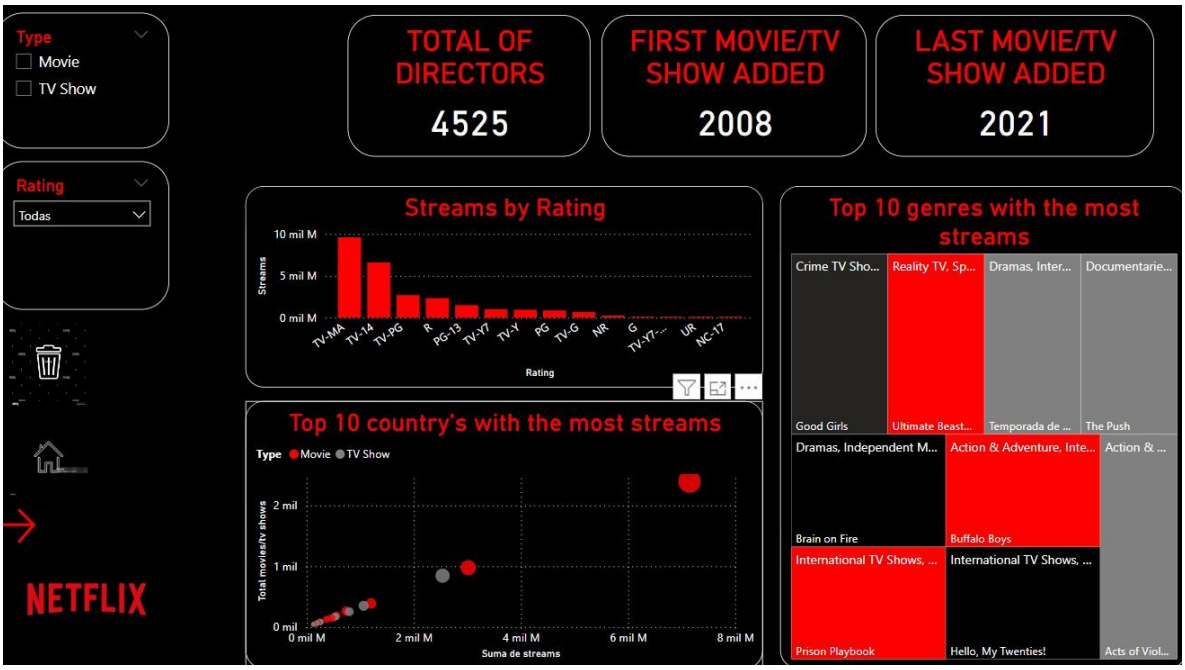
Netflix activities for country:

This section showcases Netflix's global impact by displaying the production volume of different countries worldwide. It also highlights the two countries that benefit the most from Netflix's activities.



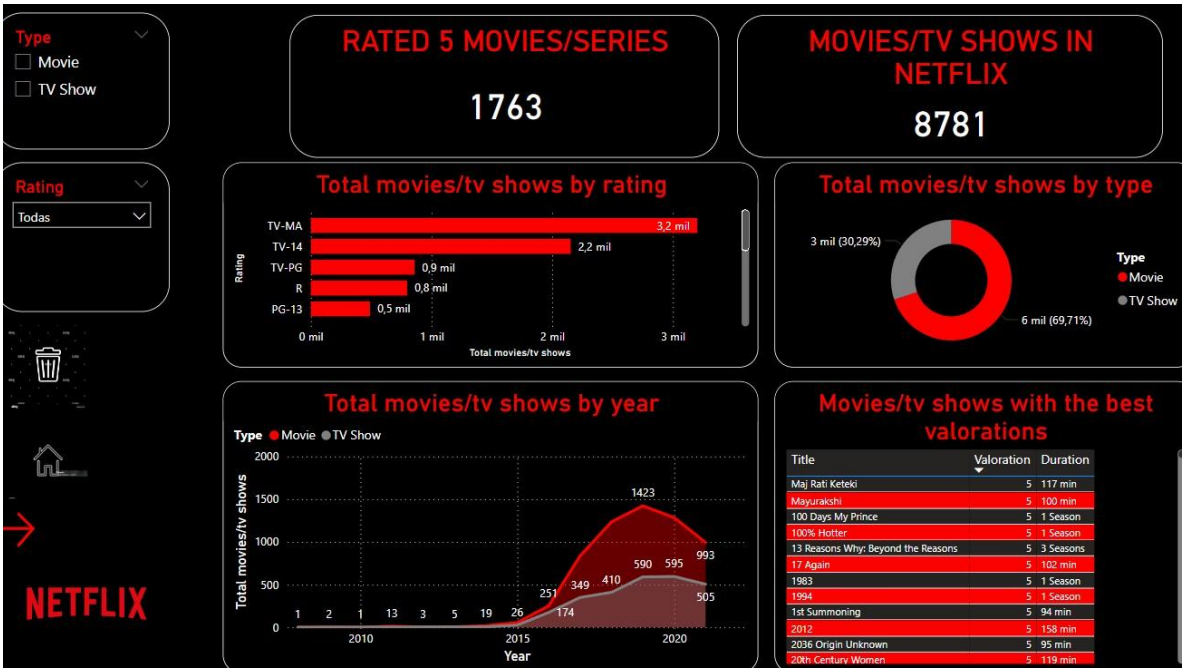
Top 10:

As the name suggests, this section presents various top 10 rankings, including genres with the most streams, countries with the highest streams, and the category with the largest number of streams.



Movies/TV shows in netflix:

This section displays the number of movies and series, categorized by type, yearly distribution, and those with the highest ratings.



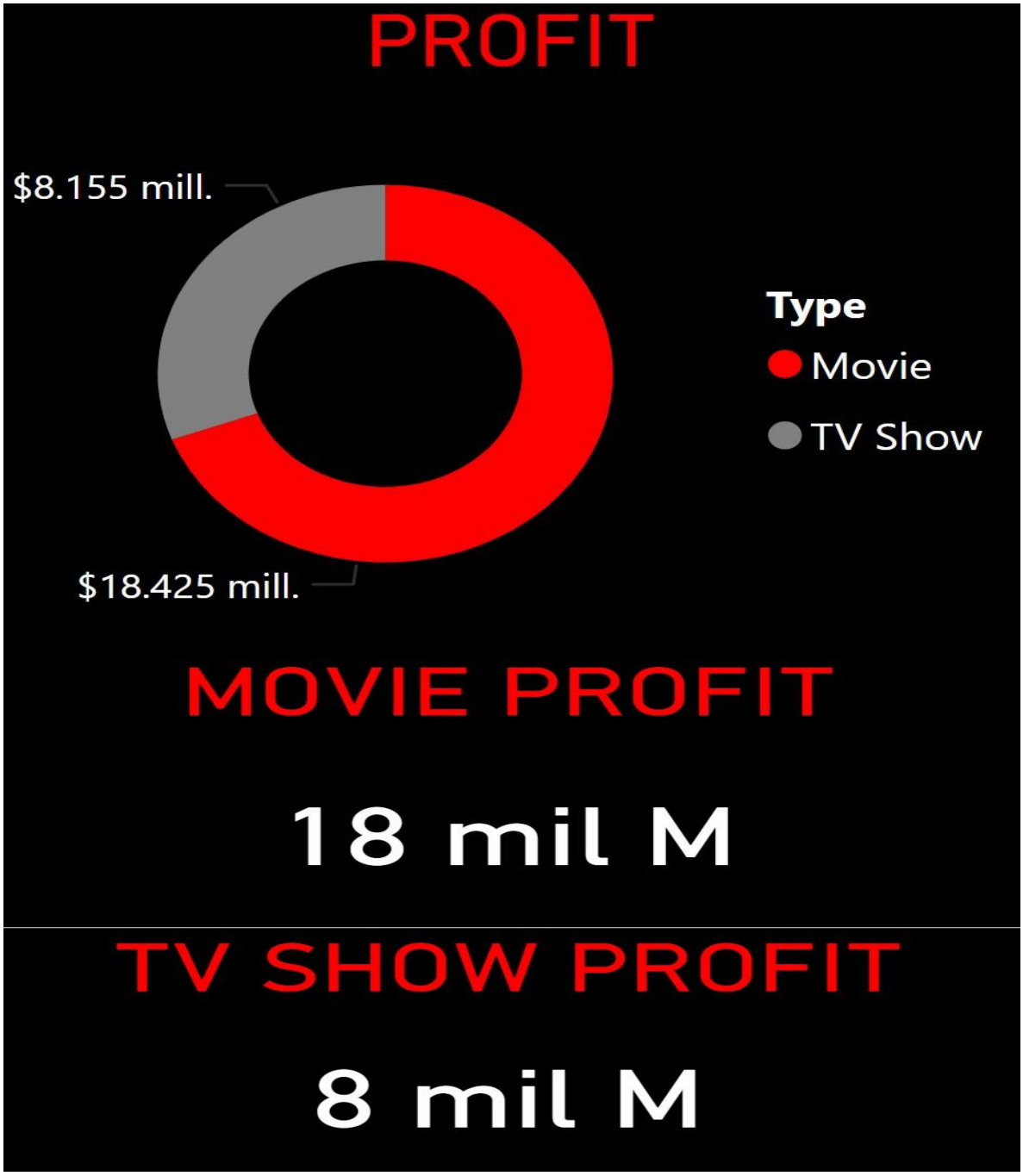
Segmentations and Tooltips:

Two data segmentations and two tooltips were created to enhance visualization:

- **Data Segmentation:** Filters the charts to display either "Movie" or "TV Show" or to filter based on category.
- **Tooltip:** Shows the profit of different types.

- **Data Segmentation:** Filters the charts to display either "Movie" or "TV Show" or to filter based on category.

- **Tooltip:** Shows the profit of different types.



10.- Conclusion:

Based on the analysis in this project, it is clear that Netflix is the largest streaming platform ever created. Its profit, global reach, and production scale extend worldwide, with a presence on every continent.