**Netflix Survey Data Analysis – Supporting Document 📊🎥**

**Overview**

This document outlines the data cleaning and transformation process applied to the Netflix dataset and includes details on key measures and visualizations. The Power BI dashboard provides insights into Netflix shows by type, rating, release year, and country of production.

**Data Cleaning and Transformation Process**

1. **Added Additional Column**:
   * Created the column **"Murder Mystery?"** to categorize shows based on whether they fall under the murder mystery genre.
2. **Replace Values**:
   * Standardized text data across the columns, particularly in categories and ratings, for consistency (e.g., replaced variations of rating labels).
3. **Removed Unwanted Columns**:
   * Removed columns that did not contribute to the analysis, including metadata like **‘description’** and other irrelevant information.
4. **Split Columns with Delimiter**:
   * Split the **‘duration’** column to extract useful information and create the **‘duration in minutes’** column for more precise analysis.
5. **Tackled Blank and Duplicated Rows**:
   * Removed rows with missing key values (like title or release year) and eliminated any duplicates to ensure data quality.
6. **Additional Column Transformation**:
   * Added other necessary columns for better filtering and segmentation of the data based on the requirements of the dashboard.

**Columns Used in Analysis**

* **Show Details**:
  + **Listed In**: Genre of the show.
  + **Show ID**: Unique identifier for each show.
  + **Cast**: Cast members involved.
  + **Country**: Country of origin of the show.
  + **Director**: Director of the show.
  + **Title**: Name of the show.
* **Categorization & Metrics**:
  + **Duration**: Duration of the show (split into minutes for detailed analysis).
  + **First Category**: Primary category/genre of the show.
  + **Rating**: Parental rating of the show.
  + **Release Year**: Year the show was released.
  + **Type**: Movie or TV show.

**Measures Created**

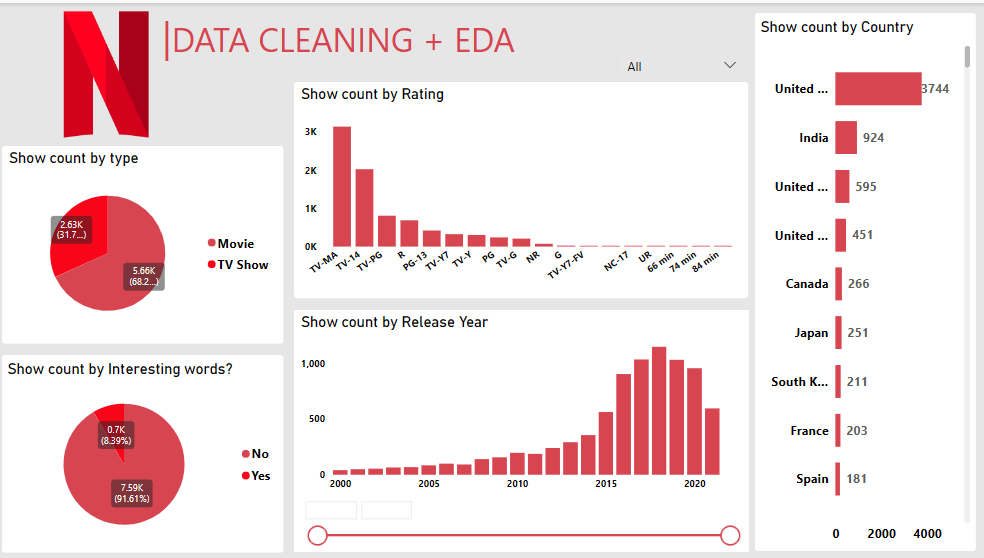
* **Total Shows**:  
  A measure to calculate the total number of shows listed, segmented by type, genre, and country.
* **Date Added**:  
  A calculated measure to track when the show was added to the Netflix platform, used for time-based visualizations.

**Visualization Overview**

The Power BI dashboard features interactive visualizations, including bar charts, pie charts, and filters, that allow users to explore the Netflix dataset effectively. Below is a sample screenshot of the dashboard:

**Screenshots**

* **Show Count by Type**: A pie chart showing the proportion of movies versus TV shows.
* **Show Count by Rating**: A bar chart visualizing the number of shows per rating category (e.g., TV-MA, PG, R).
* **Show Count by Release Year**: A bar chart showing the number of shows released each year.
* **Show Count by Country**: A vertical bar graph that breaks down shows by their country of origin.



**License**

This project is licensed under the **MIT License** – you are free to use, modify, and distribute this work with proper attribution. See the LICENSE file for more details.

**Installation**

**To use this dashboard:**

1. **Download the .pbix file.**
2. **Open it in Power BI Desktop.**
3. **Replace the data source with your own data for analysis.**
4. **Modify measures and columns as needed to suit your requirements.**

**Contributing**

**Feel free to contribute to the project by creating issues, suggesting features, or submitting pull requests. Make sure to follow the contribution guidelines outlined in the repository.**

**Conclusion**

The Netflix dataset provides a wealth of information on user preferences and show details. Through effective data cleaning and transformation steps, we’ve structured the data for meaningful analysis. The dashboard visualizes key metrics such as show types, ratings, release years, and country distribution, allowing users to explore the Netflix content library efficiently.