Data Visualization using Tableau

By: Anukriti Singh (singh72)

About Tableau:

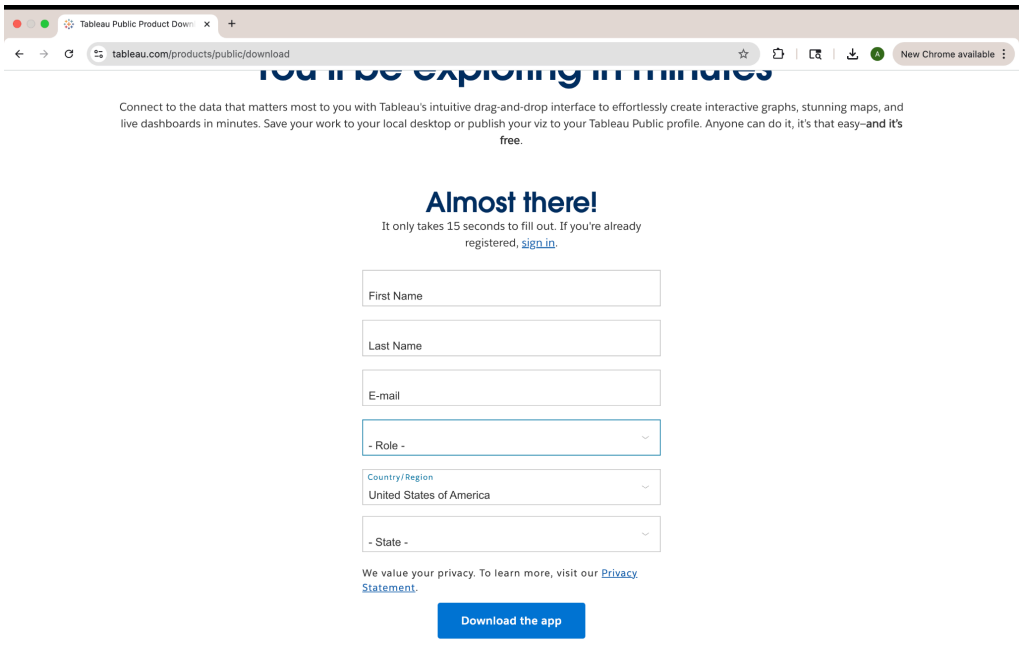
Not everyone is a Python or R-savvy data scientist, but everyone instinctively knows that a picture can say a thousand words. Tableau bridges that gap: it transforms raw spreadsheets or database tables into rich, interactive visuals with a few mouse drags, no code required.

As we shall see in action, Tableau's lightning-fast Hyper in-memory engine enables the user to load millions of rows in a matter of seconds and instantaneously transform them into bars, lines, heat maps, treemaps, and even geographic maps by just putting fields onto a canvas.

Thus, it's safe to say that Tableau strikes a balance between the ease of use of Excel and the capability of Python/R, enabling analysts, marketers, journalists, and anybody else with a question and a dataset to access professional-grade data storytelling.

Tableau installation guide:

1. Go to [Tableau Public]
2. Fill the form



3. Run the Installer, follow the instructions

4. Launch Tableau, then sign in with a free *Tableau Public* account or your licensed *Tableau Desktop* credentials.

Project Overview:

This project stitches together five publicly available Netflix datasets. By cleaning and joining these datasets, we can make an interactive Tableau dashboard that tracks viewership trends (Top 10 genre and reach in different countries), IMDb sentiment, and platform reach(on which streaming platform a show/movie is available).

CSV (these have been renamed for clarity)	Purpose	Source
netflix_full.csv	Full catalogue: title, release year, IMDb rating + votes	Full Netflix Dataset
movies_all	Platform flags for movies (Netflix, Prime, Hulu, Disney+)	Movies on Netflix, Prime Video, Hulu and Disney+
shows_all	Platform flags for TV shows (Netflix, Prime, Hulu, Disney+)	https://www.kaggle.com/datasets/ruchi798/tv-shows-on-netflix-prime-video-hulu-and-disney
all-weeks-global.csv	Weekly global Top-10 with hours viewed	https://www.kaggle.com/datasets/dhruvildave/netflix-top-10-tv-shows-and-films/data?select=most-popular.csv
all-weeks-countries.csv	Weekly Top-10 per country	https://www.kaggle.com/datasets/dhruvildave/netflix-top-10-tv-shows-and-films/data?select=most-popular.csv
most-popular.csv	28-day launch watch time	https://www.kaggle.com/datasets/dhruvildave/netflix-top-10-tv-shows-and-films/data?select=most-popular.csv

Data Preparation:

Problems with the raw data:

- Inconsistent format across files: The Top 10 lists (global, weekly, most popular) had different naming conventions and missing values in important fields like *season_title*.

- Mixed column meanings: The category column combined type (Movie/TV) with language (English/Non-English), making filtering and analysis difficult.

- Missing join keys: There was no standard column like title_key for reliable merging across datasets.

- IMDb and vote info missing or unstructured: The IMDb ratings data was only available for Netflix content, and not always in a numeric form.

- Inconsistent or missing season titles: TV shows had seasons listed inconsistently; films had null values.

- Column names with spaces and cases: Headers had uppercase letters and spaces, which are problematic in most data tools.

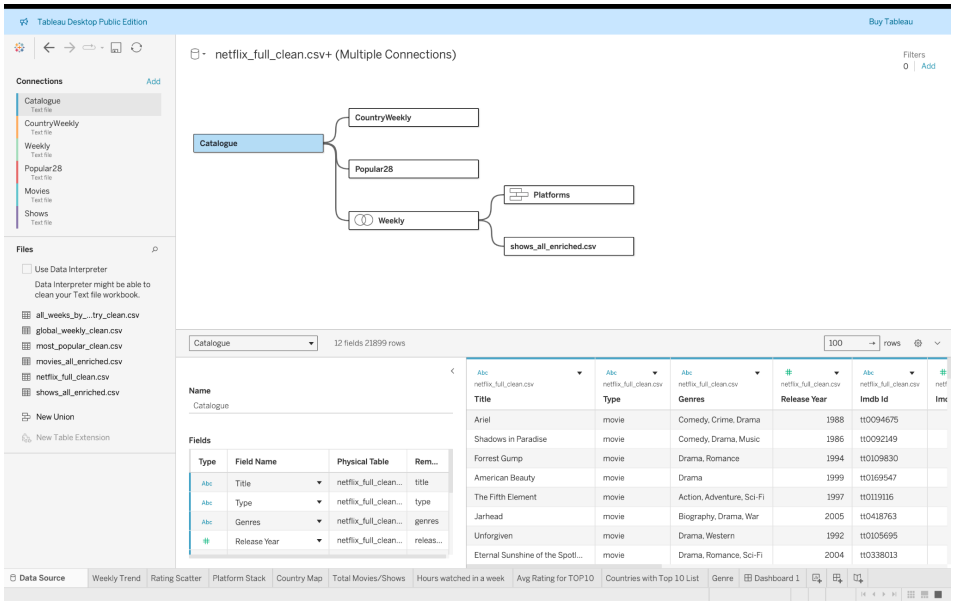
Steps followed for data cleaning and preparation:

Step	Description
Standardized Column Names	Converted all headers to lowercase snake_case.
Split Category Field	Extracted “Movie/TV” and “English/Non-English” from a single column.
Season Handling	Replaced null *season_title* with S0 for films and extracted season numbers for shows.
IMDb Data Cleaning	Parsed IMDb ratings and vote counts into usable numeric columns.
Created Join Key	Added title_key column across all tables using lowercase-trimmed show titles.

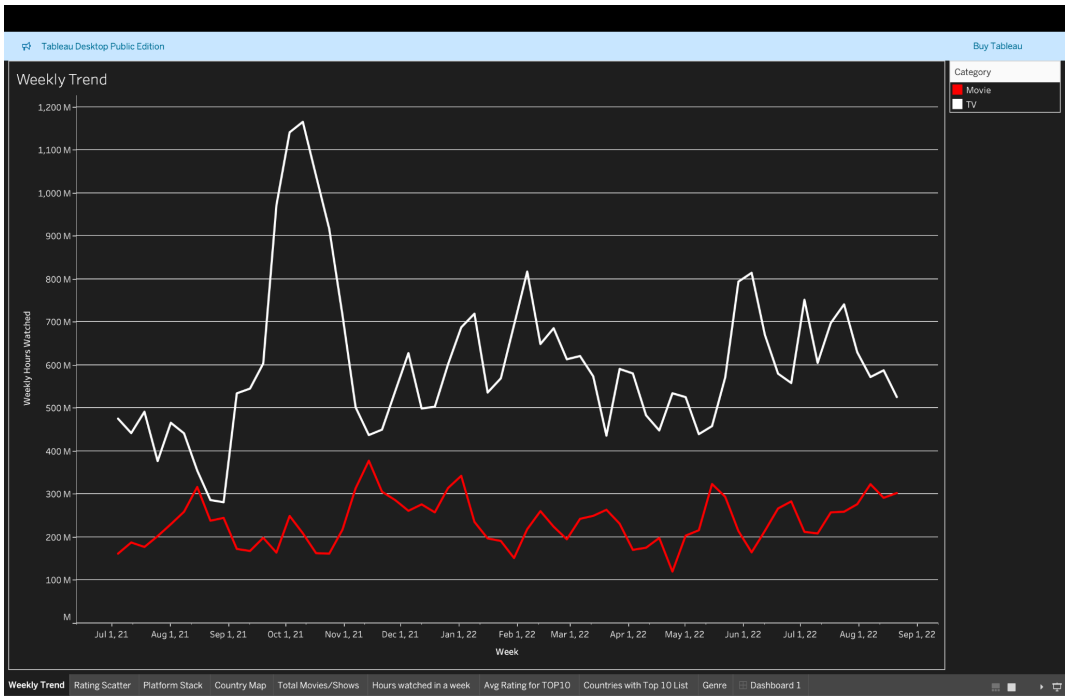
Refer to the code tutorial for a detailed walkthrough of these steps.

Loading the data in tableau:

1. Open Tableau Public → *Connect* → *Text/CSV* → pick the files.
2. Union the two platform files (movies + shows) into a single Platforms table
3. Create physical joins.
4. Click Sheet 1.



Worksheet 1: Weekly Trends (Line chart):



The “Weekly Trend” view plots global Netflix engagement week by week. Along the horizontal axis sit consecutive reporting weeks (June 2021 – September 2022), while the vertical axis shows total hours watched, in millions, for titles that entered the Top 10 during each of those weeks.

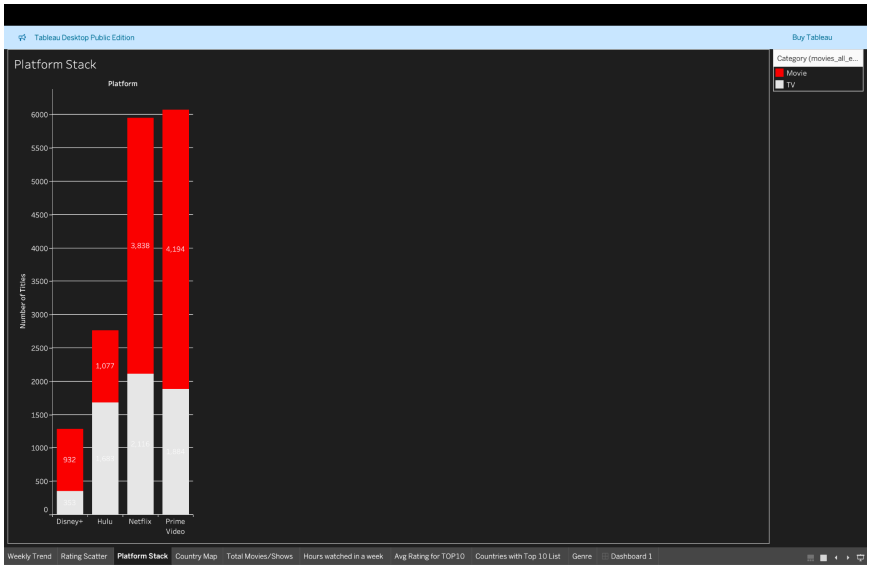
White line: TV series

Red line: movies.

Insights :

- TV series consistently outperform movies every week.
- Holiday weeks (late Dec) and early-summer weeks show the highest spikes.
- Movie spikes are sharp but brief; TV peaks plateau longer.

Worksheet 2: Platform Stack (Stacker bar):



The “Platform Stack” worksheet displays a single stacked bar for each streaming service i.e., Disney+, Hulu, Netflix, and Prime Video, showing how many titles each offers. The bar height represents total catalogue size, while its two coloured sections break that total into categories: the red portion counts movies and the light-grey portion counts TV series.

Insights:

- Netflix holds the largest overall catalogue.
- Prime Video is a close second in total size but skews even more toward movies.
- Hulu's library is TV-heavy: TV titles outnumber movies by about 60 %.
- Disney+ is the smallest overall, yet maintains a higher movie-to-TV ratio than Hulu.

Worksheet 3: Genre (Treemap):

This is a treemap where each rectangle represents a primary genre.
The rectangle size equals total weekly hours watched.

insights

- Drama occupies the single largest block, indicating it draws the highest share of watch-hours.

- Action, Crime, and Comedy together fill most of the remaining area.
- Niche genres (Horror, Documentary, Musical, etc.) appear as small blocks.

I have made some more visualizations which are included in the Tutorial video.

Final Dashboard

Tableau dashboards can hold multiple visualizations so that the user can see the story in a glance. I wanted to play around with different colors and to keep the look on-brand, and add a creative touch, so I styled the entire dashboard in Netflix's signature blacks and reds.

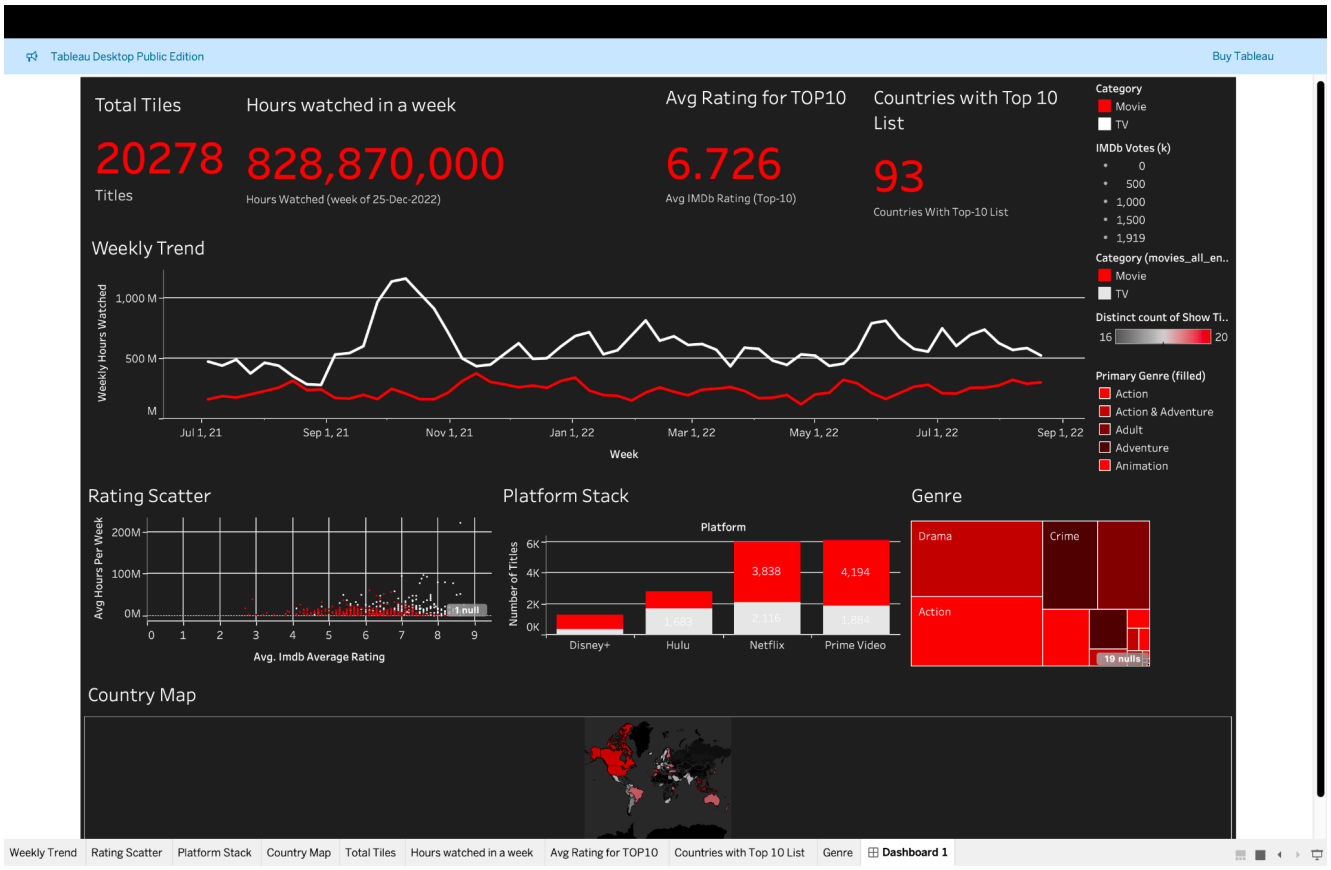


Tableau links for this project:

1. Dashboard
2. Weekly Trend
3. Rating Scatter
4. Platform Stack
5. Country map
6. Genre

References:

- https://www.geeksforgeeks.org/tableau/tableau-tutorial/
- https://help.tableau.com/current/pro/desktop/en-us/calculations_calculatedfields_create.htm
- https://help.tableau.com/current/pro/desktop/en-us/dashboards.htm
- https://help.tableau.com/current/pro/desktop/en-us/joining_tables.htm
- https://help.tableau.com/current/pro/desktop/en-us/buildexamples_line.htm