

Type of Netflix Content through the Years and Variations between Countries

“Netflix Movies and TV Shows” Dataset

“Netflix Movies and TV Shows” dataset contains information about TV Shows and Movies available on Netflix as of 2021, such as title of the content, director, actors involved and country where the content was produced. Netflix major business is a subscription-based streaming service offering online streaming from a library of films and television series, including those produced in-house. The dataset was extracted from Kaggle which is an online community of data scientists and machine learning professionals.

Using this dataframe, it's possible to proceed with some interesting questions which are:

- Understanding what content is available in different countries;
- Movie and TV Show distributed on Netflix;
- Top 20 genres on Netflix.

The available visualizations were created to understand the impact on the added content (for example, movies and TV shows) released into Netflix between 2008 and 2021. Our interest in this dataset provided insights about this platform and in what way and conditions its popularity was increasing.

It is in the follow-up that a set of visualizations are build:

1. **World Netflix Released Content** – Number of added content (cumulative) into Netflix between 2008 and 2021. The filter used is ‘Filter by Year’ (Kaggle, s.d.);
2. **Movie or TV Show** – The graph can vary according to the year selected in the ‘Filter by Year’ and ‘Filter by Country’ (Netflix Data Visualization, s.d.);
3. **Keywords in the description of Movies and TV Shows** – Indicates the most repeated words in the content description (Netflix Data Visualization, s.d.);
4. **Best Month For New Content** – Indicates the number of released content in each month according to the filters ‘Filter by Year’, ‘Filter by Country’, ‘Filter by Genre’ and ‘Filter by Type’;
5. **Countries Movies & TV Show Distribution** – Indicates the percentage of movies and tv shows according to the selected country (‘Filter by Country’) (Netflix Data Visualization, s.d.);
6. **Average Gap Release And Netflix Added Movies** – Show the average gap between the release year and date added to Netflix according to the selected year;
7. **Genres on Netflix** – Show the top countries that show the higher quantity of movies and tv shows, according to the selected genre, in ‘Filter by Genre’.

Technical Aspects

For our project we had to adapt the datasets that we found in Kaggle about Netflix, and we created new ones with these datasets for easily handle the data. From Kaggle we used two datasets, the 'netflix_df.csv' and the 'netflix_titles.csv', from 'netflix_df.csv' we create two new datasets 'genres.csv' and 'df_realise.csv'.

We selected all data related to the films and TV shows from 2008 onwards to create our project.

We created five filters, filter by year, by country, by genre, by type and by title of the TV Show or Movie, depending on the previous selection that was made in the type filter. For these filters we use Dropdown, Slider and RadioItems.

Regarding the title chosen by the person, information about the year, duration, genre, director and cast that the film / TV show had was shown.

Having created our graphics and filters, we created the app structure, in this we put as a title an image of the Netflix logo. Then we put the filters and the description of the film / TV show side by side in boxes. In another box we put the map and a graph that shows the percentage of how many films / tv shows which vary by year and an image with keywords in the description of films / tv shows.

Below we place three graphs side by side, each separated in boxes. And we conclude with a graph about the genres of Netflix movies / tv shows.

The code for this work is available on our GitHub which can be seen at this [link](#): .

Discussion

With the presentation of all these interactive visualizations, it is possible to better understand how the interest of users of the platform (Netflix) is distributed by country, as well as understand the type of movie/tv show genre that each population around the globe prefers. With this better understanding, it will be possible to plan more efficiently the availability of new content, meeting what each type of population wants to have available on the platform.

The company may be interested in exploring this topic further, to create, for example, special subscription packages tailored to each country, with content more directed to a specific target. Also, a deeper analysis regarding the best periods to launch new content could be of special interest, since the circular view presented only shows a set of numbers that are already defined, without really being able to understand why those are the most favorable months for that factor. Assuming a group, this fact may be related to vacation periods in which users, in general, spend more time watching series and/or television programs available on the platform.

References

(s.d.). Obtained from Netflix Data Visualization:

<https://www.kaggle.com/joshuaswords/netflix-data-visualization>

Kaggle. (s.d.). Obtained from Netflix EDA with Plotly:

<https://www.kaggle.com/dmitryuarov/netflix-eda-with-plotly>

Kaggle. (s.d.). Obtained from Netflix Shows: <https://www.kaggle.com/shivamb/netflix-shows>