**Project MVP on Netflix**

First steps are done, we’ve imported the libraries & utilities we need which are numpy, pandas, matplotlib, seaborn and plotly. Next we load the dataset, do some cleaning, and check if there are any missing values.

dataset contains 12 columns/features for exploratory analysis, attributes info:

* show\_id = Movie/TV show unique identifier
* type = movie or TV show
* title = Name of the movie/TV show
* director = Movie/TV show creator
* cast = actor/actress
* country = where its produced
* date\_added = date of Movie/TV show added on Netflix
* release\_year = Movie/TV show release year
* rating = targeted viewers i.e., PG-13, TV-MA, R, etc.
* duration = Movie/TV show length
* listed\_in = Movie/TV show genre
* description = brief summary of a Movie/TV show

Now it’s time to answer some questions with visualization:

* Are there more TV shows on Netflix than Movies?

Graphical user interface

Description automatically generated

Now know that more than 2/3rd of the content on Netflix are movies and the remaining 30% are TV Shows.

* what’s the best month for a producer to release some content on Netflix?

Chart

Description automatically generated

As we can see in 2021, January, March and May have comparatively less content than other months, hence, it might be a great choice to release new content during these months.

* which year produced the most movies on Netflix?

Chart, bar chart, funnel chart

Description automatically generated

2017 is the year when most of the movies were released, 2018 came a closed second.