**NETFLIX DATA ANALYSIS**



**ABOUT**

Netflix is one of the largest providers of online streaming services. It collects a huge amount of data because it has a very large subscriber base. We can analyse a lot of data and models from Netflix because this platform has consistently focused on changing business needs by shifting its business model from on-demand DVD movie rental and now focusing a lot about the production of their original shows.

Some of the most important tasks that we can analyse from Netflix data are:

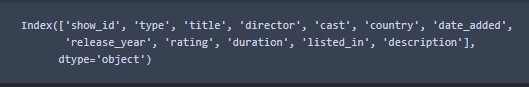
1. understand what content is available
2. understand the similarities between the content
3. understand the network between actors and directors
4. what exactly Netflix is focusing on
5. and sentiment analysis of content available on Netflix.

**DATASET**

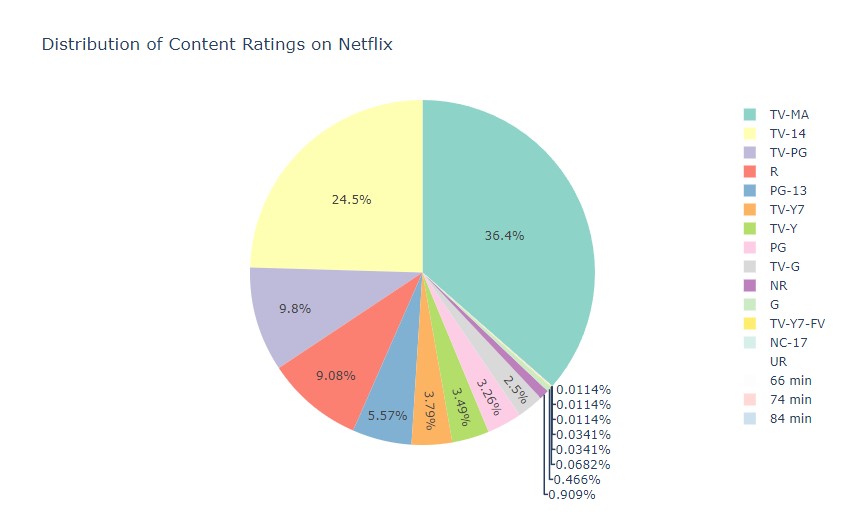
The dataset I use for the Netflix data analytics task consists of TV shows and movies streamed on Netflix as of 2019. The dataset is provided by [Flixable](https://flixable.com/) which is an engine of third-party research available on Netflix.

**IMPLEMENTATION**

1. I will start this Netflix data analysis task with Python by importing the dataset and all the Python libraries needed for this task.
2. So, the data consists of 6234 rows and 12 columns, now let us look at the column names.

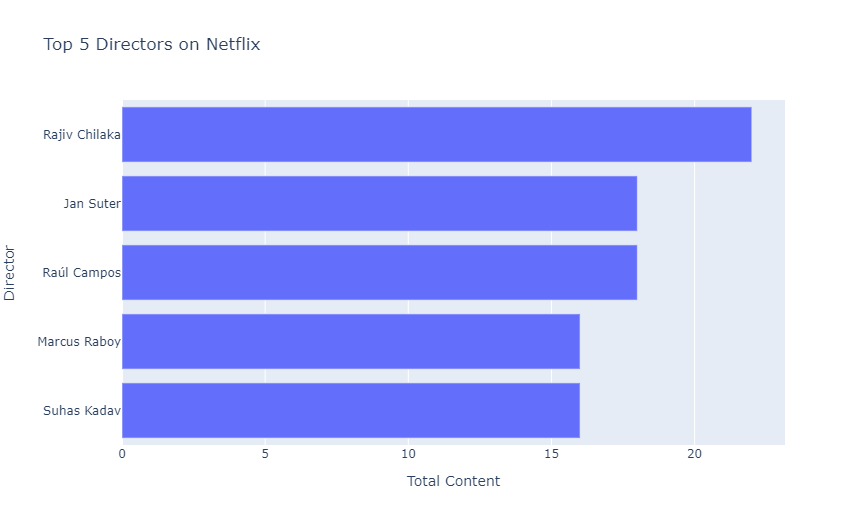


1. To begin the task of analysing Netflix data, I will start by looking at the distribution of content ratings on Netflix. The graph above shows that the majority of content on Netflix is categorized as “TV-MA”, which means that most of the content available on Netflix is intended for viewing by mature and adult audiences.



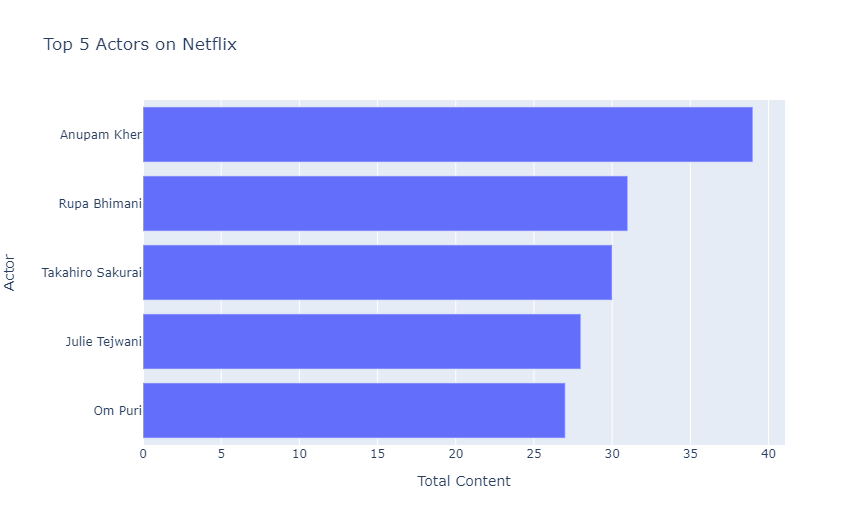
1. Now let us see the top 5 successful directors on this platform. From the above graph it is derived that the top 5 directors on this platform are:

* Raul Campos
* Jan Suter
* Jay Karas
* Marcus Raboy
* Jay Chapman

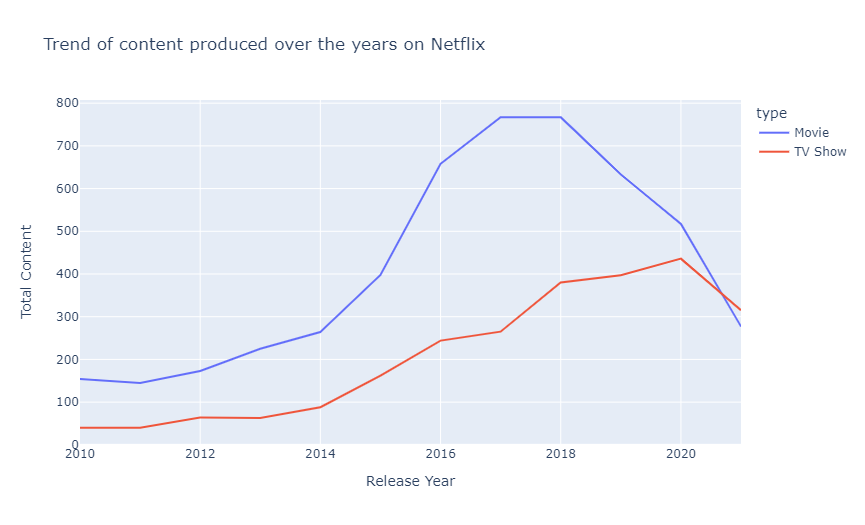


1. Now let us have a look at the top 5 successful actors on this platform. From the above plot, it is derived that the top 5 actors on Netflix are:

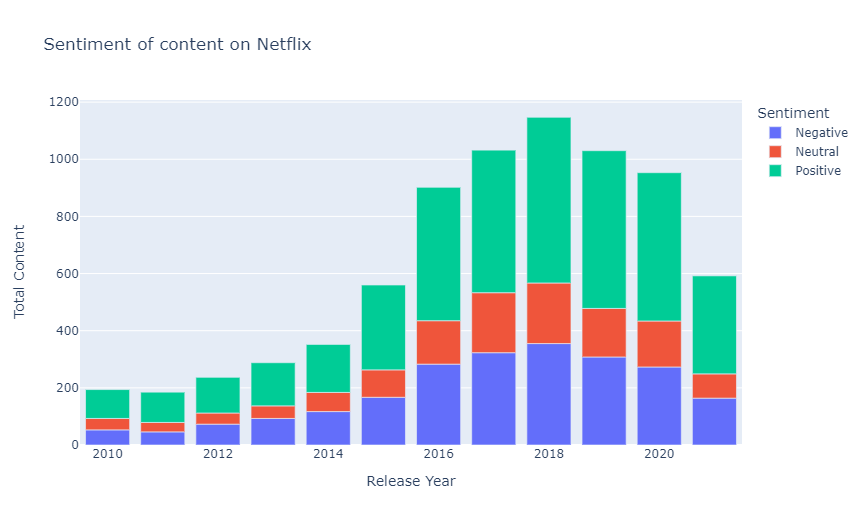
* Anupam Kher
* Om Puri
* Shah Rukh Khan
* Takahira Sakurai
* Boman Irani



1. The next thing to analyse from this data is the trend of production over the years on Netflix. The above line graph shows that there has been a decline in the production of the content for both movies and other shows since 2018.



1. At last, to conclude our analysis, I will analyse the sentiment of content on Netflix.



1. So, the above graph shows that the overall positive content is always greater than the neutral and negative content combined.